

ОБЪЯСНЕНИЕ
СОЦИАЛЬНОГО
ПОВЕДЕНИЯ:
*еще раз об основах
социальных наук*

ЮН ЭЛЬСТЕР

Перевод с английского
ИННЫ КУШНАРЕВОЙ



Издательский дом
Государственного университета — Высшей школы экономики
МОСКВА, 2011

EXPLAINING
SOCIAL
BEHAVIOR:
*More Nuts and Bolts
for the Social Sciences*

JON ELSTER

УДК 316.6
ББК 60.5
Э53

Составитель серии
ВАЛЕРИЙ АНАШВИЛИ

Научный редактор
АРТЕМ СМИРНОВ

Дизайн серии
ВАЛЕРИЙ КОРШУНОВ

Эльстер, Ю.

Э53 Объяснение социального поведения: еще раз об основах социальных наук [Текст]/пер. с англ. И. Кушнareвой; Гос. ун-т — Высшая школа экономики. — М.: Изд. дом Гос. ун-та — Высшей школы экономики, 2011. — 472 с. — (Социальная теория). — 1000 экз. — ISBN 978-5-7598-0821-3 (в пер.).

Эта книга, опубликованная в 2007 году издательством Cambridge University Press, является исправленным и дополненным изданием встретивших неоднозначные отклики «Основ социальных наук» (1989). Автор предлагает свой взгляд на природу объяснения в социальных науках; анализ психических состояний, которые предшествуют поступкам; систематическое сравнение моделей поведения, основанных на рациональном выборе, с альтернативными концепциями; исследование возможных заимствований социальных наук из нейронауки и эволюционной биологии; обзор механизмов ранжирования социальных взаимодействий от стратегического поведения до коллективного принятия решений. В провокационном заключении он настаивает на центральном значении качественной социальной науки в войне против мягких (литературных) и жестких (математических) форм обскурантизма.

УДК 316.6
ББК 60.5

First published by Cambridge University Press in 2007.

ISBN 978-5-7598-0821-3 (рус.)
ISBN 978-0-521-77179-5 (англ.)

© Jon Elster 2007

© Оформление. Издательский дом Государственного университета — Высшей школы экономики, 2011

СОДЕРЖАНИЕ

| | |
|--|-----|
| ПРЕДИСЛОВИЕ | 9 |
| ВВЕДЕНИЕ . | 12 |
| ЧАСТЬ ПЕРВАЯ. ОБЪЯСНЕНИЕ И МЕХАНИЗМЫ . | 17 |
| I. ОБЪЯСНЕНИЕ . | 21 |
| II. МЕХАНИЗМЫ. | 43 |
| III. ИНТЕРПРЕТАЦИЯ . | 62 |
| ЧАСТЬ ВТОРАЯ. РАЗУМ . | 77 |
| IV. МОТИВАЦИИ | 87 |
| V. ЭГОИЗМ И АЛЬТРУИЗМ | 106 |
| VI. БЛИЗОРУКОСТЬ И ПРОЗОРЛИВОСТЬ | 121 |
| VII. ВЕРОВАНИЯ. | 134 |
| VIII. ЭМОЦИИ | 154 |
| ЧАСТЬ ТРЕТЬЯ. ДЕЙСТВИЕ. | 171 |
| IX. ЖЕЛАНИЯ И ВОЗМОЖНОСТИ. | 175 |
| X. ЛЮДИ И СИТУАЦИИ. | 188 |
| XI. РАЦИОНАЛЬНЫЙ ВЫБОР | 201 |
| XII. РАЦИОНАЛЬНОСТЬ И ПОВЕДЕНИЕ | 201 |
| XIII. ОТВЕЧАЯ НА РАЦИОНАЛЬНОСТЬ. | 241 |
| XIV. НЕКОТОРЫЕ СЛЕДСТВИЯ ДЛЯ ИНТЕРПРЕТАЦИИ ТЕКСТОВ . | 255 |
| ЧАСТЬ ЧЕТВЕРТАЯ. УРОКИ ЕСТЕСТВЕННЫХ НАУК | 265 |
| XV. ФИЗИОЛОГИЯ И НЕЙРОНАУКА. | 271 |
| XVI. ОБЪЯСНЕНИЕ С ТОЧКИ ЗРЕНИЯ ПОСЛЕДСТВИЙ И ЕСТЕСТВЕННОГО ОТБОРА | 280 |
| XVII. ОТБОР И ЧЕЛОВЕЧЕСКОЕ ПОВЕДЕНИЕ | 295 |
| ЧАСТЬ ПЯТАЯ. ВЗАИМОДЕЙСТВИЕ | 307 |
| XVIII. НЕПРЕДНАМЕРЕННЫЕ ПОСЛЕДСТВИЯ . | 310 |
| XIX. СТРАТЕГИЧЕСКОЕ ВЗАИМОДЕЙСТВИЕ | 322 |

| | |
|---|---------|
| XX. ИГРЫ И ПОВЕДЕНИЕ | 340 |
| XXI. ДОВЕРИЕ. | 352 |
| XXII. СОЦИАЛЬНЫЕ НОРМЫ | 361 |
| XXIII. КОЛЛЕКТИВНОЕ ФОРМИРОВАНИЕ УБЕЖДЕНИЙ | 379 |
| XXIV. КОЛЛЕКТИВНОЕ ДЕЙСТВИЕ | 394 |
| XXV. КОЛЛЕКТИВНОЕ ПРИНЯТИЕ РЕШЕНИЙ | 406 |
| XXVI. ОРГАНИЗАЦИИ И ИНСТИТУТЫ | 431 |
| ЗАКЛЮЧЕНИЕ: ВОЗМОЖНА ЛИ СОЦИАЛЬНАЯ НАУКА? | 449 |

ПРЕДИСЛОВИЕ

Эта книга замышлялась как переработанное издание «Основ социальных наук» («Nuts and Bolts for the Social Sciences»), выпущенных мной в 1989 году. Однако в результате получилась гораздо более амбициозная работа. В ней поднято большее количество тем, и рассматриваются они более подробно и совсем в ином ключе. Хотя девять глав книги сохранили те же названия, что в предыдущем издании, без изменений в сущности остались только главы IX и XXIV.

При всем стремлении к полноте тематического охвата эта книга не стала трактатом, в чем-то превосходя этот жанр, а в чем-то до него не дотягивая. В ней дается простой, свободный и честный взгляд, обладающий, как я считаю, значительным потенциалом для объяснения социального поведения. Множество используемых мной примеров взяты из анекдотов или имеют литературные корни, лишь некоторые почерпнуты из более систематических исследований. Изредка применяемая здесь алгебра не выходит за рамки школьной программы. В то же время методологический и философский уклон этой книги обычно не свойствен учебным курсам. В ней предпринимается попытка поместить социальные науки в один ряд с другими науками — как естественными, так и гуманитарными. Книга также пытается заставить читателя постоянно помнить о том, какие ограничения налагают общие принципы научного анализа на любые объяснительные теории.

Стиль библиографических примечаний отражает развитие Интернета, в особенности Википедии, Google.com и Scholar.Google.com¹. Поскольку читатели в течение минуты могут найти источники любых приведенных в тексте высказываний, я опустил ссылки многих тезисов и утверждений. Вместо этого я стараюсь указать читателям на важнейшие работы, современные тексты, признанные классическими, книги и статьи, содержащие сведения, которые труднее отследить в Интернете, и на авторов, от которых я взял так много, что не упомянуть их означало бы оправдать одно из значений своей фамилии (*Elster*, по-немецки, — это «сорока, болтун»). Хотя основной текст содержит несколько ссылок на современных ученых, я обильно цитирую Аристотеля, Сенеку, Монтеня, Ларошфуко, Сэмюэля Джонсона, Ханса Кристиана Андерсена, Стендаля, Токвиля, Пруста и других классиков, остающихся в буквальном смысле неистощимым источником каузаль-

¹ Один из сервисов Google, позволяющий осуществлять широкий научно-библиографический поиск. — *Примеч. пер.*

ных гипотез. Мы лишились бы многих догадок, если бы игнорировали механизмы, подсказанные философией, литературой, драматургией и поэзией. Если бы мы отказались от размышлений о разуме, действии и взаимодействии, которые велись в течение двух с половиной тысяч лет, в пользу созданного за последние сто или десять лет, мы нанесли бы себе серьезный вред. Я цитирую этих авторов не столько ради апелляции к их авторитету, сколько для того, чтобы подчеркнуть, что широкая начитанность лучше узкого кругозора. Эта книга, находящаяся в прямой оппозиции к тому, что я воспринимаю как неумолимую профессионализацию социальных (особенно американских) наук, отвращающую студентов от изучения иностранных языков и чтения старых книг, является развернутым призывом к более широкому подходу к изучению общества.



При подготовке рукописи я получал подсказки и комментарии от многих людей. Прежде всего я должен поблагодарить моих студентов из Колумбийского университета за их острые вопросы и замечания к курсу, материалы которого легли в основу этой книги. Особенно полезными оказались предложения, сделанные Пабло Калмановичем (Pablo Kalmanovitz). Аананд Хилланд (Aanund Hylland) и Оле-Йорген Ског (Ole-Jorgen Skog) провели со мной три дня в Коллиуре, обсуждая черновик книги. С Хилландом, Карлом О. Мене (Karl O. Moene) и Джоном Ремером (John Roemer) мы также в течение полутора дней вели дискуссию в Осло. Их замечания не только спасли меня от многочисленных (!) ошибок, но и подсказали, как я могу дополнить и усилить изложение. Я особенно благодарен Ремеру за то, что он побудил меня написать заключение. Я получил письменные замечания по рукописи от Диего Гамбетты (Diego Gambetta), Раджа Сааха (Raj Saah) и от анонимного рецензента. Замечания Гамбетты были особенно подробными и полезными. У меня также состоялся полезный разговор с Уолтером Мишелом (Walter Mischel) об идеях, изложенных в главе X, которыми я во многом обязан ему. Кроме того, я получил ценные письменные замечания от Джорджа Эйнсли (George Ainslie) касательно проблем, изложенных в главе I, многие из которых поднял он сам. Бернар Манен (Bernard Manin) сделал конструктивные замечания к главе XXV. Робин Доуз (Robyn Dawes) представил острые замечания к главам VII и XII. Наконец, в течение последних нескольких лет я представлял черновики отдельных глав книги членам «Monday group», которая начиная с 1995 г. каждую неделю собиралась в Нью-Йорке осенью и изредка

весной, — Джону Ферджону (John Ferejohn), Ракель Фернандес (Raquel Fernandez), Расселу Хардину (Russell Hardin), Стивену Холмсу (Stephen Holmes), Стивену Льюксу (Steven Lukes), Бернару Манену, Паскуале Пасквино (Pasquale Pasquino), Адаму Пшеворскому (Adam Przeworski) и Джону Ремеру. Я благодарю их всех за дружеские и конструктивные возражения.

Я посвящаю эту книгу Джонатану и Джоанне Коул (Jonathan and Joanna Cole) — они знают, почему.

ВВЕДЕНИЕ

Эта книга об объяснении социального поведения. В первой части я изложу мою концепцию объяснения, а остальные четыре посвящу инструментарию из концептов и механизмов, которые можно применить к частным случаям. Излишне говорить, что я не претендую на полноту. Вместо того чтобы заполнять пробелы, которые, конечно же, будут, позвольте начать с перечисления фрагментов головоломок (puzzles), которые, как я считаю, могут быть прояснены принятым мною подходом. В заключении я вернусь к тем же головоломкам с короткими отсылками к объяснениям, которые были даны в предшествующих частях книги.

Для рассмотрения примеров и объяснений необходимо сделать две оговорки. Во-первых, я не утверждаю, что все объясняемое (explananda) является твердо установленным фактом. В самом объяснении это, конечно, важнейший первый шаг; бессмысленно объяснять то, чего не существует. Однако для формирования инструментария необязательно придерживаться строгих правил. Во-вторых, даже в случае хорошо задокументированного объясняемого я не утверждаю, что приведенные мною объяснения являются правильными. Я только утверждаю, что они удовлетворяют минимальным требованиям для объяснения, а именно имеют своим логическим следствием объясняемое. Предполагается, что головоломки и объяснения должны показать: если такого рода вещи происходят, вот механизм, который мог бы это объяснить; а также: если действует такой механизм, вот что он может произвести. С учетом данных предположений приведу головоломки, собранные в довольно произвольном порядке (поскольку многие из них подпадают сразу под несколько категорий) в соответствии с четырьмя основными частями книги¹.

I. РАЗУМ

- Почему одни игроки полагают, что если пять раз подряд выпадало красное, то более вероятно, что именно оно, а не черное, выпадет в следующий раз?
- Почему другие игроки полагают, что если пять раз подряд вы-

¹ Хотя этот список в какой-то мере пересекается со списком головоломок, представленных в главе XI, в целях проблематизации теории рационального выбора, он не имеет полемических целей и призван лишь возбудить любопытство читателя.

падало красное, то выше вероятность, что дальше выпадет черное, а не красное?

- Почему предпочтения иногда меняются с течением времени?
- Почему многие из тех, кто верит в жизнь после смерти, тем не менее хотят, чтобы смерть наступила как можно позже?
- Почему люди неохотно признаются себе и другим в том, что они завистливы?
- Почему люди неохотно признаются себе и другим в том, что они чего-то не знают?
- Почему вера в то, что человеку предопределено попасть в рай или ад, приносила новообращенным кальвинистам XVI века большее успокоение, чем вера в то, что спасения можно достичь добрыми делами?
- Почему иногда верно, что «тот, кто обидел, не может простить»?
- Почему в некоторых культурах стыд важнее вины?
- Почему победа французской сборной в 1998 году на ЧМ по футболу вызвала в стране такую радость, и почему тот факт, что французская сборная не смогла выйти из группы на ЧМ 2002 года, поверг страну в такое уныние?
- Почему женщины часто испытывают стыд, став жертвой изнасилования?
- Почему унижительные ритуалы инициации вызывают большую, а не меньшую, верность группе?

II. Действие

- Почему сегодня на бродвейских шоу аплодируют стоя чаще, чем двадцать лет назад?
- Почему наказания могут привести к увеличению, а не к уменьшению, количества преступлений, которые они призваны предотвратить?
- Почему люди так неохотно нарушают правила, которые сами себе установили, даже если их соблюдение не имеет смысла?
- Почему существует месть по модели «два oka за одно», а не по модели «око за око»?
- Почему долгосрочный доход по акциям гораздо выше, чем по облигациям (то есть почему стоимость акций не поднимается, чтобы сравняться доходами)?
- Почему количество самоубийств падает, когда опасные лекарства начинают продавать в блистерных упаковках, а не в бутылочках?

- Почему ни один из тридцати восьми очевидцев не вызвал полицию, когда убивали Китти Дженовезе¹?
- Почему были люди, скрывавшие и спасавшие евреев от нацистов?
- Почему президент Ширак в 1997 году объявил досрочные выборы, в результате только потеряв большинство в парламенте?
- Почему при разводе родители часто соглашаются на совместную опеку над ребенком даже в тех случаях, когда, обратившись в суд, один из них мог бы добиться исключительного права опеки?
- Почему вероятность эмиграции малоимущих ниже?
- Почему некоторые люди выбирают рождественские вклады, по которым не выплачивается процент и с которых нельзя снять деньги до Рождества?
- Почему люди занимаются такими изначально неприбыльными проектами, как создание авиалайнера «Конкорд»?
- Почему в юстиции переходного типа (когда агентов авторитарного режима отдают под суд после перехода к демократии), те, кого судят сразу после такого перехода, получают более суровые приговоры, чем те, кого судят позднее?
- Почему в шекспировской пьесе Гамлет откладывает месть до последнего акта?

III. УРОКИ ЕСТЕСТВЕННЫХ НАУК

- Почему родители чаще убивают приемных детей и пасынков, чем своих биологических детей?
- Почему так редок инцест между потомством одних родителей, учитывая возникающие соблазны и все имеющиеся для этого возможности?
- Почему люди вкладывают деньги в проекты под руководством третьих лиц, даже если те свободно могут оставить себе всю прибыль?
- Почему люди мстят, хотя месть приносит им материальные убытки и не приносит материальной выгоды?
- Почему люди торопятся с выводами, выходя за рамки имеющихся доказательств?

¹ Американка итальянского происхождения, забитая до смерти на глазах нескольких свидетелей, не предпринявших никаких попыток оказать помощь. Инцидент вошел в учебники социальной психологии под названием «синдром Дженовезе».

IV. ВЗАИМОДЕЙСТВИЕ

- Почему приверженцы социалистических партий иногда голосуют за коммунистов и тем самым препятствуют победе своей партии?
- Почему новые независимые государства принимают в качестве официального язык своих бывших империалистических угнетателей?
- Почему палатки с мороженым на пляже часто стоят рядом, тогда как и для покупателей, и для продавцов было бы удобнее, если бы они стояли на отдалении друг от друга?
- Почему люди голосуют на выборах, когда почти точно знают, что их голоса никак не повлияют на результат?
- Почему в современных западных обществах экономически успешные люди обычно стройнее остальных?
- Почему люди воздерживаются от некоторых сделок, которые могли бы всем принести выгоду (например, не просят вперёдистоящего продать свое место в очереди к автобусу)?
- Почему президент Никсон старался казаться Советскому Союзу человеком, склонным к нерациональному поведению?
- Почему военачальники иногда сжигают за собой мосты (или собственные корабли)?
- Почему люди всегда придают большое значение в сущности незначительным вопросам этикета?
- Почему пассажиры оставляют чаевые таксистам, а посетители ресторанов — официантам даже при посещении иностранного города, в который они не предполагают вернуться?
- Почему фирмы инвестируют в большие запасы товара на складе, даже когда не предвидят остановки производства?
- Почему в группе студентов каждый думает, что другие поняли сложный текст лучше, чем он?
- Почему на многих политических ассамблеях проводятся поименные голосования?
- Почему договоренности по принципу «Ты — мне, я — тебе» чаще встречаются в обычных законодательных органах, чем на конституционных ассамблеях?

Возможные объяснения этих явлений будут даны в разных частях этой книги и кратко резюмированы в Заключение. Здесь я хочу только сделать общее замечание о двух видах объяснения, которые не представляются полезными. Как читатель увидит в самой первой части и с учетом нескольких напоминаний по ходу дела, одна из целей этой книги — вну-

шить скептицизм по отношению к двум распространенным линиям рассуждений. Во-первых, за несколькими исключениями, социальные науки не могут полагаться на функциональное объяснение, которое оценивает действия или модели поведения через их следствия, а не через их причины. Разве система чаевых существует потому, что посетители могут проследить за официантами эффективнее, чем владелец ресторана? Я так не думаю. Во-вторых, сегодня я полагаю, что теория рационального выбора обладает меньшей объясняющей способностью, чем мне казалось ранее. Разве реальные люди действуют, исходя из расчетов, заполняющих многие страницы математических приложений к ведущим журналам? Я так не думаю.

Тем не менее как минимум по трем пунктам теория рационального выбора составляет значимую часть инструментария. Понятая в качественном ключе здравый смысл, она способна объяснить многое в повседневном поведении. Даже когда она мало что объясняет, она может иметь огромную концептуальную ценность. В частности, теория игр прояснила структуру социального взаимодействия, выйдя далеко за рамки прозрений, полученных в предшествующие столетия. Человек, наконец, *хочет* быть рациональным. Желание иметь прочные основания своих поступков, а не быть марионеткой неосознанно действующих психических сил, оказывает постоянное противодействие многим иррациональным механизмам, исследуемым мною в этой книге.

Хотя я критически отношусь ко многим объяснениям через рациональный выбор, я считаю концепцию *выбора* фундаментальной. В книге я рассматриваю несколько альтернатив объяснениям через рациональный выбор и прихожу к заключению, что хотя многие из них иногда являются полезным дополнением к этому подходу, полностью заменить его они не могут. Тот факт, что люди совершают поступки под действием различных *ограничений*, например, часто может объяснить значительное число вариаций в поведении. В некоторых случаях можно возразить, что за наблюдаемое поведение отвечает скорее *отбор агентов*, нежели их собственный *выбор*. Однако в общем и целом я полагаю, что субъективный фактор выбора обладает большей объясняющей способностью, чем объективные факторы ограничений и отбора. Очевидно, что это интуитивная догадка, которая не может быть доказана в строгом смысле слова. В любом случае представители социальных наук должны найти в своем инструментарии место для всех факторов.

ОБЪЯСНЕНИЕ И МЕХАНИЗМЫ

Ч А С Т Ь П Е Р В А Я

Основа этой книги — особый подход к проблеме объяснения в социальных науках. Не являясь трудом по философии социальных наук как таковой, книга заимствует и отстаивает определенные методологические принципы объяснения социальных явлений. В первых трех частях эти принципы изложены прямо. В оставшейся части книги они присутствуют в неявном виде, хотя время от времени, а именно в главах с XIV по XVIII и в Заключение, снова выходят на первый план.

Я утверждаю, что всякое объяснение каузально. Объяснить какое-то явление (*экспланандум*) — значит привести более раннее явление (*эксплананс*), которое стало его причиной. Ратуя за каузальные объяснения, я не отрицаю возможность интенционального объяснения поведения. Намерения могут выступать в качестве причин. Отдельная разновидность интенционального объяснения — *объяснение через рациональный выбор* — будет подробно рассмотрена в последующих главах. Многие интенциональные объяснения, однако, основываются на допущении, что агенты так или иначе *иррациональны*. Сама по себе иррациональность — просто негативная или остаточная идея, включающая все то, что нерационально. Чтобы эта идея имела какую-то объяснительную ценность, мы должны апеллировать к особым формам иррациональности со специфическими последствиями для поведения. Например, в XII главе я перечисляю одиннадцать механизмов, которые могут породить иррациональное поведение, и привожу примеры.

Иногда ученые объясняют явления через их *следствия*, а не через их *причины*. Они могут сказать, например, что кровопролитные распри объясняются необходимостью поддерживать на приемлемом уровне численность населения. Это может показаться метафизически абсурдным: как существование чего-то или какое-то событие, произошедшее в одной временной точке, может объясняться чем-то, что еще не возникло? В дальнейшем мы увидим, как можно сформулировать эту проблему, чтобы объяснение через следствие стало осмысленной концепцией. В биологических науках можно найти пример эволюционного объяснения. Однако в социальных науках успешные примеры таких объяснений встречаются реже. Пример с кровопролитной враждой точно не из их числа.

Естественные науки, особенно физика и химия, предлагают *объяснения при помощи законов*. Законы — это обобщенные положения, позволяющие делать вывод о том, что в какой-то момент одно утверждение верно, если верным было другое утверждение в какой-то предшествую-

щий момент времени. Так, когда мы знаем положение и скорость планет в один момент времени, законы их движения дают нам возможность рассчитать и предсказать их положение в любое другое время. Этот вид объяснения является *детерминистским*: если даны antecedенты, то возможен только один консеквент. Социальные науки редко (если вообще когда-либо) могут предложить объяснения такого рода. Отношения между экспланансом и экспланандумом — не «один к одному» или «много к одному», но «один ко многим» или «многие ко многим». Ученые нередко пытаются моделировать их с помощью методов *статистики*. Такие объяснения, однако, сами по себе неполны, поскольку вынуждены полагаться в конечном счете на интуитивные догадки относительно правдоподобных причинно-следственных механизмов.

I. Объяснение

ОБЪЯСНЕНИЕ: ОБЩИЕ ПОЛОЖЕНИЯ

Основная задача социальных наук — объяснение социальных явлений. Это не единственная, но самая важная задача, которой подчинены или от которой зависят все прочие. Основной разновидностью экспланандума является *событие*. Объяснить — значит рассказать, почему это произошло, сославшись на *предшествующее событие* как на причину. Так, мы можем объяснить победу Рональда Рейгана на президентских выборах в 1980 году провалом попытки Джимми Картера спасти американских заложников в Иране¹, а начало Второй мировой войны — любым количеством предшествующих событий, от Мюнхенского пакта до Версальского договора. Даже если в обоих случаях структура каузального объяснения будет более сложной, они все равно воплощают базовую модель объяснения по типу *событие-событие*. В традиции, идущей от Давида Юма, ее часто называют моделью бильярдного шара. Одно событие — шар А, ударяющий шар В, — является причиной и тем самым объясняет другое событие, а именно то, что шар В приходит в движение.

Те, кто знаком с типичным объяснением в социальных науках, могут не узнать эту модель или не заметить ее привилегированного статуса. Социальные науки так или иначе склонны фиксировать скорее *факты* или положение вещей, чем события. Предложение «В 9 утра дорога была скользкой» сообщает о факте. Предложение «В 9 утра автомобиль занесло» сообщает о событии. Как подсказывает этот пример, для аварии можно найти объяснение через *факт-событие*². И наоборот, для какого-то положения вещей можно предложить объяснение через *событие-факт*, как в случае утверждения, что атака на Всемирный торговый центр в 2001 году объясняет преследующее многих американцев чувство

¹ Чтобы предварить различие, обсуждаемое далее, заметим, что Картер не то чтобы не смог *предпринять попытку*, он ее *предпринял*, и она *провалилась*. Не-действие, такое как отказ от попытки, не может обладать каузальным эффектом, разве что только косвенным образом, если другие люди, осознав или заключив, что агент не может действовать, сами совершают поступки, которых они в противном случае не совершили бы.

² Другой иллюстрацией может послужить пример с количеством избирателей, пришедших на выборы, который мы разберем позднее.

страха. Наконец, часто в социальных науках даются стандартные объяснения по модели *факт-факт*. Пример, взятый наугад: высказывались предположения, что уровень образования женщин объясняет годовой доход на душу населения в развивающихся странах.

Давайте рассмотрим объяснение одного конкретного факта: 65% американцев выступают (или заявляют, что выступают) в поддержку смертной казни³. В принципе, эта проблема может быть сформулирована в категориях событий: как эти люди *пришли к одобрению* смертной казни? Какие именно значимые события — эпизоды с родителями, сверстниками или учителями — способствовали формированию такой позиции? На практике представители социальных наук редко интересуются этим вопросом. Вместо того чтобы объяснять голую статистику такого рода, они хотят понять *изменение* отношений во времени или *различия* в отношении в зависимости от социальной группы. Причина, возможно, в том, что они находят *голые* факты недостаточно информативными. Если спросить, много это или мало — 65%, ответ будет: «По сравнению с чем?». По сравнению с отношением американцев в 1990 году, когда за смертную казнь выступали 80%, это низкая цифра; по сравнению с отношением в некоторых европейских странах — высокая.

Исследования методом *продольных срезов* рассматривают временные вариации зависимой переменной, исследования методом *перекрестных срезов* — вариации среди разных групп населения. В обоих случаях экспланандум трансформируется. Вместо того чтобы попытаться объяснить явление «само по себе», мы пытаемся объяснить, как оно меняется во времени и пространстве. Успех объяснения частично измеряется тем, какую вариантность (*variance*) (техническая мера вариаций) оно допускает⁴. Полный успех объяснил бы все наблюдаемые вариации. Из общенационального исследования мы, например, могли бы выяснить, что процент людей, поддерживающих смертную казнь, строго пропорционален количеству убийств в расчете на 100 тыс. человек. Хотя эти результаты *не смогли бы* объяснить абсолютные цифры, они могли бы дать *прекрасное* объяснение разнице между ними⁵. На практике полный успех, конечно же, не достигим, но один пункт сохраняется. Объяснения вариантности ничего не говорят об экспланандуме «самом по себе».

³ Оценки колеблются. Кроме того, число людей, поддерживающих смертную казнь, резко сокращается, когда в качестве альтернативы выдвигается пожизненное тюремное заключение без права помилования.

⁴ Экономисты иногда говорят, что их интересует только то, что происходит в пределе (*at the margin*).

⁵ Строго говоря, причинная цепочка может идти и в другую сторону — от отношений к поведению, но в этом случае гипотеза недостоверна.

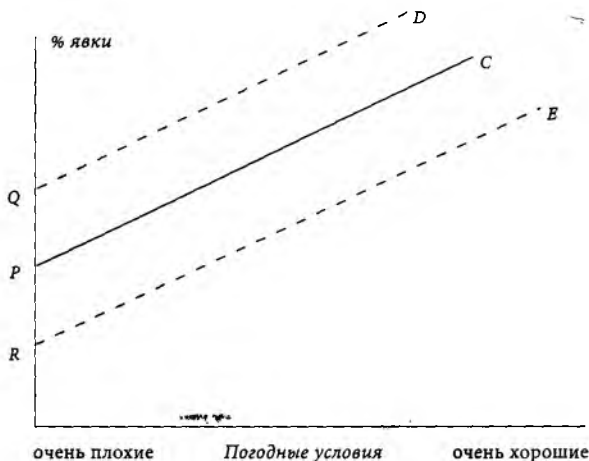


РИС. 1.1

Можно привести пример из исследований поведения избирателей. Как мы позднее увидим (глава XII), неясно, почему избиратели вообще участвуют в общегосударственных выборах, на которых, как можно с уверенностью сказать, один голос никогда ничего не решает. И все-таки значительная часть электората приходят на избирательные участки в день выборов. Почему бы им не махнуть на них рукой?

Вместо того чтобы попытаться решить эту задачу, эмпирические социологи обычно задаются другим вопросом: почему явка меняется от выборов к выборам? Есть гипотеза, что избиратели с меньшей вероятностью придут в плохую погоду, потому что, когда идет дождь или снег, приятнее сидеть дома. Если данные соответствуют этой гипотезе, как это показывает прямая С на рис. 1.1, можно утверждать, что варьирование явки объяснено (хотя бы частично). И все же нет *никакого* объяснения, почему прямая С пересекает ось ординат в точке Р, а не, например, в Q или R. Это как принять за данную первую десятичную дробь и сосредоточиться на объяснении второй. Для прогноза этого может быть достаточно, но для объяснения — нет. Голый факт, что 45 или более процентов электората обычно приходят на выборы, интересен и требует объяснения.

С точки зрения модели *событие-событие* идеальной процедурой была бы следующая. Рассмотрим два голосования, А и Б. Для каждого выявим условия, заставившие определенный процент избирателей прийти на выборы. Как только мы таким образом объяснили явку на выборах А и явку на выборах Б, объяснение разницы в явке (если таковая имеется)

последует автоматически как побочный эффект. Попутно мы также сможем объяснить, является ли одинаковая явка на выборах А и Б случайностью, то есть идет речь об уравнивающих друг друга различиях или нет. На практике эта процедура может оказаться очень сложной. Возможно, данные или имеющиеся теории не позволят нам объяснить явление «само по себе». Нужно, однако, сознавать, что прибегая к объяснениям вариантности, мы втягиваемся во второсортную объяснительную практику.

Иногда социологи пытаются объяснять *не-события* (non-events). Почему многие люди не обращаются за положенными им социальными льготами? Почему никто не вызвал полицию в деле Китти Дженовезе?⁶ В первом случае объяснением может служить то, что указанные лица *решили* не обращаться за льготами из-за стыда или страха за собственную репутацию. Поскольку принятие решения *является* событием, это дало бы совершенно удовлетворительное обоснование. В случае неудачи социологи снова обратятся к *различиям* между теми, кому положены льготы и кто их получает, и теми, кому они положены и кто за ними не обращается. Предположим, что они различаются только тем, что люди, не обращающиеся за льготами, не знают о том, что они им положены. В качестве объяснения это полезно, но недостаточно. Мы пойдем дальше и попробуем объяснить, *почему* те, кому положены льготы, не знают об этом. Выяснить, что они неграмотны, не могут прочесть писем, информирующих об их правах, будет полезно, но недостаточно. В какой-то момент в этом объяснительном регрессе мы должны придти к положительному событию, такому, например, как сознательное решение не учиться грамоте или сознательное решение официальных лиц скрывать информацию. Или мы должны заняться теми, кто обращается за положенными льготами. Как только мы объясним их поведение, объяснение, почему другие так и не сумели получить льготы, возникнет в качестве побочного продукта.

В случае с Китти Дженовезе нет никаких вариаций поведения, которые нужно объяснять, поскольку *никто* не позвонил в полицию. Отчеты об этой истории указывают на то, что несколько свидетелей *решили* не делать этого. С точки зрения непосредственных причин это вполне удовлетворительное объяснение, хотя мы, возможно, захотим узнать причины их решения. Произошло это, потому что они боялись ввязываться или потому что каждый решил, что в полицию позвонит кто-нибудь другой («У семи нянек дитя без глаза»)? Однако некоторые из на-

⁶ 27 марта 1964 года тридцать восемь уважаемых, законопослушных граждан в Квинсе, Нью-Йорк, более получаса наблюдали за тем, как убийца выслеживал женщину и ранил ее ножом, трижды напав на нее в Кью Гарденс. Дважды голоса и свет в квартирах его спугнули. И каждый раз он возвращался, находил жертву и снова наносил ранения. Ни один человек не вызвал полицию, ни один свидетель не позвонил после того, как женщина умерла.

блюдателей даже не думали вызывать полицию. Один мужчина и его жена наблюдали за сценой ради развлечения, другой сказал, что ему все надоело и он пошел спать. Чтобы объяснить, почему они не действовали решительнее, можно указать на отсутствие глубоких эмоций, но это снова будет объяснение негативного экспланандума через использование негативного эксплананса. Наше поведение в очередной раз будет истолковано лишь побочным, остаточным образом. Располагай мы удовлетворительным объяснением того, почему некоторые люди собирались позвонить в полицию, пусть и не решившись на это, у нас было бы единственное объяснение того, почему другие об этом даже не подумали.

В оставшейся части книги я часто буду применять не столь пуристский или ригористичный подход к тому, что считать релевантным экспланандумом и приемлемым объяснением. Настойчивое обращение к объяснениям, фокусирующимся на событиях, чем-то сродни принципу методологического индивидуализма, который является одним из исходных условий этой книги. В принципе, объяснения в социальных науках должны ссылаться только на индивидуумов и их действия. На практике социологи часто оперируют такими надиндивидуальными общностями, как семья, фирма или нация, как *безобидными условными обозначениями*, либо если отсутствие данных или надежной теории *вынуждают* к этому. Эти два оправдания применимы также к использованию фактов в качестве экспланандума или эксплананса, к объяснению скорее вариантности, а не явлений «самих по себе» и к анализу негативного экспланандума (не-событий и не-фактов). Цель предшествующих рассуждений — не навязать социологам бессмысленные или трудновыдерживаемые стандарты, но доказать, что на уровне основных принципов подход, основанный на событиях, по природе своей превосходит остальные. Принимая это в расчет, ученые чаще смогут находить точные и продуктивные объяснения.

Иногда у нас может возникнуть желание объяснить событие (скорее паттерн событий) через его последствия, а не причины. Я имею в виду не объяснение через *преднамеренные* (intended) последствия, поскольку намерение предшествует выбору или действию, которые он объясняет, а скорее идею, что события могут объясняться их *действительными* (actual), чаще всего *благоприятными* для кого-либо или чего-либо последствиями. Так как причина должна предшествовать следствию, эта идея может показаться внутренне противоречивой, и все же каузальное объяснение может принять такую форму при наличии обратной связи между последствиями и причинами. Изначально слезы ребенка могут быть просто вызваны болью, но если они привлекли внимание родителей, ребенок начнет плакать сильнее, чем если бы этого не произошло. В главах XVI и XVII я утверждаю, что этот род объяснения занимает до некоторой степени маргинальное место в изучении человеческого поведения. В этой книге я буду заниматься преимущественно простой разновидно-

стью каузального объяснения, где эксплананс, который может включать верования и обращенные в будущее намерения, предшествует появлению экспланандума⁷.

В дополнение к заслуживающему уважения функциональному объяснению, основанному на особых механизмах обратной связи, есть менее респектабельные формы, которые просто указывают на возникновение некоторых благоприятных последствий, а затем без каких-либо дополнительных доводов предполагают, что этого достаточно для объяснения поведения, которое их вызывает. Если экспланандум всего лишь экземпляр (token), такой как отдельный поступок или событие, то такой тип объяснения не срабатывает по чисто метафизическим причинам. Приведем пример из биологии: мы не можем объяснить появление нейтральной или вредной мутации, указав, что она была необходимым условием для дальнейшей, более благоприятной мутации. Когда экспланандум — это *тип*, как повторяющаяся модель поведения, он может быть действительным (valid) или недействительным. Но если он не поддерживается механизмом обратной связи, к нему следует относиться как к недействительному (invalid). Антропологи, например, утверждали, что мстительное поведение имеет множество благоприятных последствий — от контроля за численностью населения до децентрализованного принуждения к выполнению норм (в главе XII приведено много других примеров). Даже если допустить, что такие преимущества действительно обеспечиваются, их возникновение может оказаться случайным. Чтобы это опровергнуть, то есть показать, как некоторые преимущества поддерживают мстительное поведение, которое в свою очередь становится их причиной, необходимо продемонстрировать механизм обратной связи. И даже если такая демонстрация состоялась, изначальное появление такого экспланандума должно быть связано с чем-то еще.

СТРУКТУРА ОБЪЯСНЕНИЙ

Позвольте теперь перейти к более подробному рассмотрению темы объяснения в социальных науках (и, до некоторой степени, не только в них). Первый шаг легко пропускают: прежде чем попытаться объяснить факт или событие, мы должны установить, что факт *является* таковым, а событие действительно имело место. Монтень писал: «Я сознаю, что если

⁷ В ряде случаев полезно провести различие между каузальными, интенциональными и функциональными объяснениями. В физике используется только каузальное объяснение, биология помимо этого допускает функциональное объяснение, а социальные науки используют еще и интенциональные объяснения. Однако на базовом уровне все объяснения являются каузальными.

попросить людей изложить факты, они гораздо больше времени потратят на то, чтобы объяснить их причины, а не на выяснение того, правда это или нет... Они пропускают факты, но старательно делают выводы. Как правило, они начинают так: Как же так случилось? Но правда ли что-то случилось? Вот что им следовало бы спрашивать⁸.

Так, прежде чем объяснить, скажем, почему в одной стране происходит больше самоубийств, чем в другой, мы обязательно должны убедиться, что вторая страна не занижает количество самоубийств по, например, религиозным соображениям. Прежде чем пытаться объяснить, почему в Испании безработица выше, чем во Франции, нам придется удостовериться, что эти различия никак не связаны с расхождениями в определениях безработицы или со значительной долей «теневой» экономики в Испании. Если мы хотим объяснить, почему во Франции безработица среди молодежи выше, чем в Англии, нам нужно решить, что мы принимаем за экспланандум — уровень безработицы среди молодых людей, активно ищущих работу, или среди всей молодежи в целом, включая студентов. Если мы хотим сравнить безработицу в Европе и США, придется решить, берем мы в качестве экспланандума безработицу в широком понимании, то есть с учетом лиц, лишенных свободы, или же в техническом смысле, включая только тех, кто ищет работу⁸. Прежде чем пытаться объяснить, почему месть происходит по принципу «око за око» («я буду убивать одного из ваших всякий раз, как вы будете убивать одного из моих»), нужно проверить, что мы наблюдаем именно это явление, а не, скажем, «два ока за одно» («я буду убивать двоих из ваших всякий раз, как вы будете убивать одного из моих»). Науки, в том числе социальные, часто пытаются объяснить общеизвестные вещи, но не менее ценным было бы установление того, что некоторые вещи, представляющиеся нам таковыми, ими не являются. В этом случае социальные науки могли бы помимо прочего попытаться объяснить, почему мы уверены, что знаем что-то, хотя это не так, добавив тем самым новый фрагмент знания взамен отброшенного⁹.

⁸ В обоих случаях некоторые индивидуумы могут стать преступниками или студентами, потому что сомневаются, что смогут найти работу. Для одних целей их учитывают среди безработных, для других — нет.

⁹ Точно так же, как наука может помочь объяснить популярную веру в не-факты (non-fact), она помогает объяснить популярную веру в ложные объяснения. Например, большинство страдающих от артрита верят, что боли в суставах вызваны плохой погодой. Исследования показывают, однако, что такой связи нет. Возможно, нужно оставить поиски причинной связи между плохой погодой и артритом и вместо этого объяснить, почему больные артритом верят в такую связь. Скорее всего они когда-то услышали о наличии такой связи и с тех пор обращали больше внимания на случаи, которые подтверждали это представление, нежели на те, что его опровергали.

Теперь предположим, что мы имеем четко сформулированный экспланандум, для которого нет четко сформулированного объяснения, — *головоломку* (puzzle). Головоломкой может быть неожиданный или противоречащий здравому смыслу факт или просто необъяснимая корреляция. Один небольшой пример: почему из оксфордских библиотек чаще всего крадут книги по теологии, а не по каким-нибудь другим предметам? Еще один пример, на котором я вскоре остановлюсь подробнее: почему на бродвейских шоу теперь аплодируют стоя чаще, чем двадцать лет назад?

В идеале к объяснительной загадке нужно применять процедуру, состоящую из пяти шагов, которая будет изложена далее. Однако на практике шаги 1–3 не всегда идут по порядку. Мы можем «играть» с различными гипотезами до тех пор, пока одна из них не покажется нам наиболее убедительной, а потом начнем искать теорию, которая ее подтвердит. Если шаги 4–5 сделаны правильно, мы все еще можем полагаться на гипотезу, которой отдали предпочтение. Но по причинам, которые я буду обсуждать в конце следующей главы, у ученых может возникнуть желание ограничить свободу в выборе гипотез.

1. Выбрать теорию, набор взаимосвязанных каузальных положений, которая предоставляет самые большие перспективы успешного объяснения.
2. Подобрать гипотезу, которая согласует теорию с головоломкой, то есть чтобы экспланандум логически вытекал из гипотезы.
3. Установить или представить себе надежные основания, которые могут дать альтернативные объяснения, снова таким образом, чтобы экспланандум логически вытекал из каждого из них.
4. Опровергнуть все конкурирующие объяснения, указав для каждого на дополнительные, поддающиеся проверке следствия, которые в действительности *не* наблюдаются.
5. Усилить предложенную гипотезу, показав, что она имеет дополнительные, поддающиеся проверке следствия, желательны *новые факты*, которые находят свое подтверждение в действительности.

Эта процедура определяет то, что часто называют *гипотетико-дедуктивным методом*. В конкретном случае она может принимать форму, отраженную на рис. 1.2. Я проиллюстрирую ее головоломкой о росте числа стоячих оваций на Бродвее. Данное наблюдение основывается не на систематических замерах или контролируемых экспериментах, а на моих случайных впечатлениях, нашедших подтверждение в сообщ-

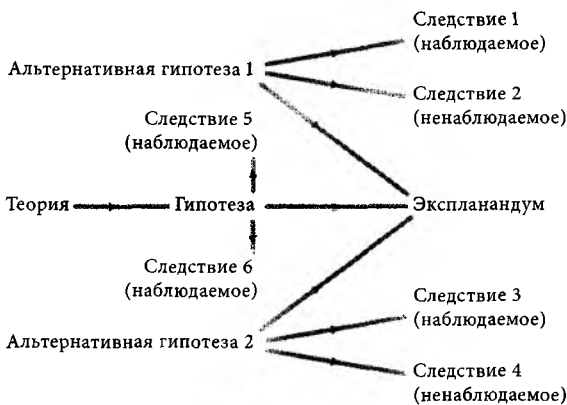


РИС. I.2

щения из газет. В данном случае шаткий статус экспланандума не имеет значения. Если на Бродвее сейчас действительно чаще аплодируют стоя, чем двадцать лет назад, то какое объяснение мы можем этому найти?

Я рассмотрю объяснение с точки зрения увеличения стоимости билетов на бродвейские спектакли. Журналист одной из газет сообщил, что драматург Артур Миллер якобы заявил: «Я думаю, что публика просто чувствует, что заплатила за то, чтобы сидеть, 75 долларов, и теперь пришло время постоять. Не хочу быть циничным, но положение изменилось, когда цена пошла вверх». Если люди заплатили 75 долларов или даже больше за билет, многие из них не могут себе признать, что шоу было плохим или посредственным и что они потратили деньги зря. Чтобы убедить себя в том, что они отлично провели время, они бурно аплодируют.

Для более формального объяснения можно использовать гипотезу «Чем больше денег и усилий затратили люди на то, чтобы что-то получить, тем в большей мере они склонны (при прочих равных условиях) оценивать полученное выше, чем когда они потратили меньше»¹⁰. Если рассматривать рост цен как исходную предпосылку, данное предполо-

¹⁰ Такого рода идею иногда используют для оправдания своих высоких гонораров психотерапевты: пациенты не поверят в терапию, если не заплатят за нее много денег. Но ни один известный мне терапевт никогда не утверждал, жертвует 50% своих гонораров Красному Кресту.

жение выдерживает минимальные требования, которым должна удовлетворять любая объяснительная гипотеза: если она истинна, из нее можно вывести экспланандум. Но требования и в самом деле настолько малы, что слишком многие предположения способны их выдержать¹¹. Чтобы подкрепить нашу веру именно в это объяснение, мы должны показать, что оно поддерживается снизу, сверху и сбоку.

Объяснение поддерживается *снизу* (from below), если мы можем вывести и проверить поддающиеся наблюдению факты, исходя из гипотезы, превосходящей факт, который она призвана объяснить. Гипотеза должна обладать избыточной объяснительной способностью. В случае с бродвейским шоу мы будем ожидать меньшего числа стоячих оваций на представлениях, цены на которые по каким-то причинам не выросли¹². Следует, кроме того, ожидать, что стоячих оваций будет меньше, если билеты на шоу закуплены компаниями и распространены среди сотрудников. (Это будет считаться *новым фактом*.) Даже если эти билеты дороги, зрители заплатили за них не из своего кармана и, следовательно, не обязаны доказывать себе, что потратили деньги не зря.

Объяснение поддерживается *сверху* (from above), если объясняющая гипотеза может быть выведена из более общей теории¹³. В данном случае объяснительное предположение является уточнением теории когнитивного диссонанса, предложенной Леоном Фестингером. Согласно этой теории, когда человек сталкивается с внутренним расхождением или расхождением между своими убеждениями и ценностями, можно ожидать своего рода ментальной перенастройки, которая устранил или смягчит этот диссонанс. Как правило, такая перенастройка идет по пути наименьшего сопротивления. Человеку, потратившему 75 долларов на неудачное, как оказалось, шоу, нелегко признать себе в том,

¹¹ Человеческий разум склонен к тому, чтобы представлять это минимальное требование как достаточное. Как только нам попалась гипотеза, которая может оказаться верной, мы часто не останавливаемся, чтобы дополнительно ее проверить или рассмотреть альтернативные варианты. Выбор гипотезы может быть обусловлен идеей *post hoc ergo, propter hoc* (после этого, следовательно, по причине этого) или заключением о том, что данная гипотеза более правдоподобна, чем остальные, из чего делается вывод, что она с большей вероятностью может оказаться верной.

¹² Совсем не обязательно, что люди на более дешевых местах в этом случае не будут вставать. Они могут почувствовать себя глупо, если будут сидеть, когда другие люди встали. Или они поднимутся, чтобы увидеть актеров, которых заслоняют от них вставшие зрители.

¹³ Точнее говоря, если это уточнение более общей теории. Отношение между общей теорией и конкретной объяснительной гипотезой редко бывает отношением выводимости. Во-первых, в самой теории может быть изъян (см. главу II); во-вторых, из одной теории может быть несколько следствий.

что деньги были пущены на ветер. Проще внушить себе, что представление на самом деле вполне удалось.

Хотя в ней есть свои недостатки, теория когнитивного диссонанса хорошо обоснована. Частично это обоснование опирается на примеры, существенно отличающиеся от тех, что мы здесь рассматриваем, как, допустим, в случае с покупателем автомобиля, который жадно ищет рекламу купленной марки, чтобы убедиться в правильности принятого решения. Иногда подтверждение предоставляют очень похожие ситуации, как например, болезненные и унижительные ритуалы инициации при вступлении в студенческие братства, воспитывающие в их членах глубокую преданность. Я не утверждаю, что люди сознательно говорят себе: «Я столько вынес, чтобы попасть в эту группу, значит, это хорошая группа». Механизмы формирования лояльности должны быть неосознаваемыми.

Объяснение может получить боковую опору (*lateral support*), если мы сможем придумать, а затем опровергнуть альтернативные объяснения, также удовлетворяющие минимальным требованиям. Возможно, стоячих оваций стало больше, потому что сегодняшняя публика, которую автобусы пачками привозят из Нью-Джерси, менее взыскательна, чем избалованные обитатели Нью-Йорка, или потому что шоу стали лучше, чем раньше. Для каждой из этих альтернатив мы должны найти, а затем опровергнуть дополнительные факты, которые позволили бы их верифицировать. Если стоячие овации участились оттого, что публика стала более впечатлительной, то следовало бы ожидать такой же частоты стоячих оваций и на выездных представлениях двадцатилетней давности. Если шоу стали лучше, чем двадцать лет назад, это должно отразиться и на рецензиях в прессе, и на том, как долго они остаются в репертуаре, прежде чем их закроют.

При такой процедуре защитнику начальной гипотезы приходится также выступать в роли адвоката дьявола. Он должен постоянно *опровергать себя*, максимально усложняя исследуемый вопрос. Нужно выбирать самые сильные и наиболее правдоподобные альтернативные объяснения, а не те, что можно легко опровергнуть. По сходным причинам, пытаясь продемонстрировать избыточность объяснительной силы гипотезы, мы должны вывести и подтвердить следствия, являющиеся новыми, непривычными, противоречащими здравому смыслу и максимально отличающимися от исходного экспланандума. Эти два критерия — опровергать наиболее правдоподобные альтернативы и порождать *новые факты* — являются решающими для достоверности объяснения. Поддержка сверху помогает, но не бывает решающей. В конечном счете, теорию поддерживают именно порождаемые ею успешные объяснения, а не наоборот. Эмилио Сегре, нобелевский лауреат в об-

ласти физики, сказал, что некоторые победители делают честь премии, тогда как другим делает честь она. Последние, таким образом, паразитируют на первых. Подобным образом теория паразитирует на всех порожденных ею успешных объяснениях. Если она способна оказать поддержку конкретному объяснению, то только потому, что сама получила поддержку от других, более ранних объяснений.

ЧЕМ ОБЪЯСНЕНИЕ НЕ ЯВЛЯЕТСЯ

Высказывания, имеющие своей целью объяснение события, следует отличать от *семи других типов высказывания*.

Во-первых, каузальные объяснения следует отличать от *подлинных каузальных утверждений* (true causal statements). Недостаточно привести причину: должен быть указан (или по крайней мере предложен) причинно-следственный механизм. В повседневном языке, хороших романах, исторических трудах и в многочисленных социологических исследованиях этот механизм не эксплицируется. Вместо этого его подсказывает способ описания причины. Каждое конкретное событие может быть описано самыми разными способами. В (хороших) нарративных объяснениях подразумевается, что для идентификации события используются его черты, которые релевантны с точки зрения причинности. Если сказано, что некто умер в результате употребления в пищу гнилых продуктов, мы понимаем, что таким механизмом было пищевое отравление. Если сказано, что он умер, съев пищу, на которую у него была аллергия, мы решаем, что в основе лежит аллергическая реакция. А теперь предположим, что он действительно умер из-за пищевого отравления, но кроме этого у него была аллергия на некоторые продукты, например на лобстера. Заявить, что он умер, потому что съел продукт, на который у него была аллергия, — значит сказать правду, но такое утверждение обманчиво. Утверждение, что он умер, потому что съел лобстера, также истинно, но неинформативно. Оно не содержит отсылки к причинно-следственному механизму и может подпадать под действие многих других; например, человек может быть убит кем-то, кто поклялся убить первого встреченного едока лобстера.

Во-вторых, каузальные объяснения следует отличать от высказываний о *корреляциях* (correlations). Иногда можно с полной уверенностью утверждать, что за событием одного типа всегда или как правило следует событие другого типа. Это не дает нам права говорить, что события первого типа вызывают события второго типа, потому что есть еще одна возможность: оба события могут быть общим следствием третьего. В «Жизни Сэмюэля Джонсона» Дж. Босуэлл описывает, как некий Маколей, хоть «и не имея никаких предрассудков против предрассудков»,

заявлял, что когда судно прибывает в Сент-Кильду на Гебридах, «у всех обитателей начинается насморк». В то время как некоторые предлагали каузальное объяснение этому (предполагаемому) факту, в одном из писем Босуэллу сообщили, что «учитывая географическое положение Сент-Кильды, для высадки чужеземцев на берег необходимо, чтобы дул северо-восточный ветер. Именно он, а не чужеземцы, вызывает эпидемию насморка». Или возьмем утверждение, что дети, дело об опеке над которыми рассматривалось в суде, чаще страдают от нервных расстройств, чем дети, чьи родители смогли достичь мирового соглашения в частном порядке. Возможно, различие объясняется самим фактом судебного процесса, причиняющим детям страдания и вызывающим у них чувство вины. Но может быть и так, что судебные разбирательства с большей вероятностью случаются там, где родители враждебно настроены друг к другу и именно их дети склонны к нервным расстройствам. Для различения этих двух интерпретаций нам пришлось бы замерить страдания до и после развода. Третьего варианта мы коснемся позднее.

Вот несколько более сложный, но любимый мною пример такого рода двусмысленности. В «Демократии в Америке» Алексис де Токвиль обсуждает предполагаемые каузальные связи между заключением брака по любви и несчастьем в браке. Он указывает на то, что эта связь может возникнуть только в обществах, где такого рода браки — исключение, а правилом является брак по расчету. Только упрямцы пойдут против течения, а двое упрямых людей вряд ли будут счастливы в браке¹⁴. Кроме того, к людям, поступающим вразрез с общепринятой моралью, хуже относятся их собратья, склонные к конформизму, что вызывает разочарование и горечь. Первый из этих аргументов основывается на некаузальной, вызванной третьим фактором корреляции между любовью и отсутствием счастья. Второй указывает на истинно каузальную связь, но не на ту, о которой думают критики браков по любви, к которым Токвиль адресует свое возражение. Брак по любви вызывает несчастье только в той среде, где такая практика является исключением. Биологи часто называют этот эффект частотной зависимостью (frequency dependent)¹⁵.

¹⁴ Здесь третий фактор — скорее черта характера (упрямство), чем событие.

¹⁵ Первый механизм — эффект отбора (a selection effect), второй — явление подлинного последствия (aftereffect). Это различие применимо довольно широко. Если мы спросим, почему человек в определенном состоянии (например, в положении безработного или госпитализированный по причине психического заболевания) склонен к тому, чтобы длить свое пребывание в нем, объяснением может послужить любой из этих механизмов (или оба). Например, люди, которые уже давно находятся без работы, могут образовывать подгруппу населения, чьи трудовые навыки не востребованы. И наоборот, все работающие лица могут потерять работу, но как только это про-

Помимо проблемы третьего фактора корреляция может вызвать у нас сомнения относительно *направления* каузальности. Рассмотрим старый анекдот:

ПСИХОЛОГ: Вы должны быть мягче с Джонни. Он из разбитой семьи.
УЧИТЕЛЬ: Ничего удивительного, Джонни сам какую хочешь семью разобьет.

Или, как говорил комик Сэм Левинсон, «Безумие — наследственная болезнь, его можно получить от собственных детей». Подразумевается, что психическое расстройство ребенка может разрушить брак, а не сам развод привести к такому расстройству. Аналогично негативная корреляция между тем, насколько родители осведомлены о том, чем занимаются их дети, и склонностью последних попадать в беду, не должна доказывать, что родительский присмотр эффективен; ведь подростки, намеренно ввязывающиеся в неприятности, вряд ли станут рассказывать, чем занимаются.

В-третьих, каузальное объяснение следует отличать от утверждений о *необходимости* (necessitation). Объяснить событие — значит показать, почему оно произошло именно так, как произошло. Тот факт, что обстоятельства могли сложиться иначе и сложились бы, не случись то, что случилось, не проясняет дела. Представим себе человека, больного раком поджелудочной железы в той стадии, которая через год скорее всего приведет к его смерти. Когда боль становится невыносимой, человек кончает жизнь самоубийством. В попытке объяснить, почему он умер в течение какого-то времени, бессмысленно указывать, что ему все равно пришлось бы умереть от рака¹⁶. Всего, что нам известно об этом случае, — возникновение раковой опухоли, предела срока жизни человека с таким диагнозом и факта смерти, достаточно для того, чтобы прийти к достоверному выводу: человек умер от рака. У нас есть предшествующее событие и причинно-следственный механизм, позволяющий объяснить последующее событие. Но механизм не является необходимым, его может упредить действие другого механизма. (В этом примере уп-

Окончание сн. 15

исходит, отсутствие работы меняет их самих (или меняет отношение к ним работодателей), так что вероятность их возвращения на рынок труда со временем падает. Теория «навешивания ярлыков», объясняющая психические заболевания или преступления, основывается на (сомнительном) допущении, что отложенные эффекты управляют эффектом отбора.

¹⁶ Джеймс Фитцджеральд Стифен (James Fitzjames Stephen) писал: «В законе совершенно четко указано, что если после нападения человек умер весной от болезни, которая должна была бы убить его, допустим, летом, причиной смерти является нападение».

реждающая причина сама является следствием упрежденной причины, но и это необязательно, так как человек мог погибнуть в автомобильной аварии.) Чтобы выяснить, что в действительности произошло, нам нужны более точные сведения. Поискам не будет конца: до самой последней секунды какие-то другие факторы могут упредить рак¹⁷.

Высказывания о необходимости иногда называют структурными объяснениями (structural explanations). Примером может служить анализ Великой французской революции, предпринятый Токвилем. В своей работе он выделяет целый ряд событий и тенденций, начиная с XV столетия вплоть до 80-х годов XVIII века, и утверждает, что на их фоне революция была неизбежна. По-видимому, он имел в виду, что (1) любое количество мелких и средних событий спровоцировало бы революцию, и что (2) была зыбкая уверенность в том, что *некоторые* вызывающие ее события произойдут, хотя и необязательно они будут именно теми и тогда, которые и когда в действительности произошли. Он также утверждает, что (3) после 1750 или, возможно, 1770 года было уже невозможно ничего сделать, чтобы предотвратить революцию. Хотя после Токвиля остались заметки для второго тома, в котором он намеревался описать революцию так, как она *в действительности* происходила, можно возразить, что если он успешно установил (1), (2) и (3), то в дальнейших шагах нет необходимости. Проблема такого рода рассуждений в том, что в случае многих интересных для социальных наук вопросов (и в отличие от примера с раком) такие утверждения, как (1), (2) и (3), очень трудно обосновать непредвзятыми методами¹⁸. Более сильный аргумент можно получить, если подобные события произошли в одно и то же время независимо друг от друга, что наводит на мысль, что они витали в воздухе. В качестве примера можно привести одновременные открытия в науке.

В-четвертых, каузальное объяснение следует отличать от *рассказывания историй* (storytelling). Подлинное объяснение передает то, что

¹⁷ Каузальное упреждение (preemption) следует отличать от каузальной сверхдетерминации (overdetermination). Последнюю можно проиллюстрировать случаем, когда в человека попадают сразу две пули, одной из которых достаточно, чтобы его убить; первое — случай, когда человека убивает одна пуля, упреждая тем самым действие второй, выпущенной на несколько секунд позже.

¹⁸ Возможно, самым убедительным кандидатом для структурного объяснения является американская революция. Такой нейтральный наблюдатель, как французский министр Шуасель, заметил еще в 1765 году, что независимость американских колоний неизбежна. Для беспристрастного французского комментатора, такого как Раймон Арон, независимость Алжира также была предрешена задолго до ее провозглашения. Французская революция напоминает скорее крушение коммунизма — неизбежная, главным образом, при взгляде в прошлое.

произошло в том виде, в каком оно произошло. Рассказать историю — значит объяснить то, что произошло, так, как оно *могло* произойти (и, возможно, произошло). Научные объяснения, как я ранее отмечал, отличаются от рассказов о том, что *должно было* произойти. Теперь же я заявляю, что они отличаются и от рассказов о том, что *могло* произойти. Это утверждение может показаться банальным или странным. Зачем кому-то выступать с чисто гипотетическим изложением событий? Уместны ли такие спекуляции в науке? Ответ: да, но не следует путать их с научным объяснением.

Рассказывание историй может подсказывать новые, то есть экономные, объяснения. Предположим, что кто-то утверждает, что самоотверженное поведение или стремление оказывать помощь являются решающими доказательствами того, что люди не всегда руководствуются эгоистическими побуждениями, а эмоциональные поступки — решающими свидетельствами того, что не все действия рациональны. Можно сделать вывод, что есть три не сводимые друг к другу формы поведения: рациональное и эгоистическое, рациональное и неэгоистическое и иррациональное. Характерное для серьезной науки стремление к экономии должно заставить нас усомниться в таком взгляде. Разве не бывает так, что люди помогают другим в расчете на вознаграждение, а приступы гнева, напротив, помогают им добиться своего? Рассказав историю о том, как рациональный эгоизм *может* порождать альтруистическое и эмоциональное поведение, можно перевести этот вопрос из области философии в область, доступную эмпирическому исследованию¹⁹. История по принципу «просто так» (just-so) может стать первым шагом на пути к выработке удачного объяснения. На самом деле, многие ответы на головоломки, представленные во введении, которые я даю в Заключение, имеют сильный привкус именно таких историй.

В то же время рассказывание историй может ввести в заблуждение и нанести вред, будучи по ошибке принятым за настоящее объяснение. За исключением двух случаев, описываемых в этом абзаце, объяснения по модели «как если бы» (as-if) в действительности ничего не объясняют. Рассмотрим для примера распространенное утверждение, что мы можем использовать модель рационального выбора для объяснения поведения, даже если *знаем*, что люди не могут проделывать в уме сложные вычисления, воплощенные в этой модели (или в математических приложениях к статьям, в которых излагается эта модель). Пока модель позволяет делать предсказания, в которые хорошо вписывается наблюдаемое поведе-

¹⁹ В этом отдельном случае истории по модели «просто так» представляются ложными, поскольку люди помогают другим и при разовом взаимодействии, а гнев может привести к тому, что остальные откажутся от взаимодействия.

ние, мы имеем право (как утверждается) говорить, что агенты действуют «как если бы» они были рациональны. Это операционалистский, или инструменталистский, взгляд, впервые предложенный физикой, а позднее использованный Милтоном Фридманом применительно к социальным наукам. Согласно этому взгляду, способом с большой точностью предсказывать и объяснять поведение искусного бильярдиста является предположение, что он знает законы физики и может совершать в голове сложные вычисления. Вопрос о том, верно ли такое допущение, к сути дела *не относится*.

Этот аргумент справедлив в отношении некоторых ситуаций, когда агенты учатся с течением времени методом проб и ошибок. Однако применим он именно потому, что мы можем указать на *механизм*, который неосознанно приводит к тому же результату, который сверхрациональный агент мог бы сознательно просчитать²⁰. При отсутствии такого механизма мы могли бы по-прежнему использовать инструменталистский подход, если бы это допущение позволяло предсказывать поведение с большой точностью. Закон всемирного тяготения долгое время казался таинственным, поскольку, как представлялось, основывался на непрозрачной идее действия на расстоянии. И все же, поскольку он давал возможность делать предсказания с точностью до десятых долей, теория Ньютона единогласно признавалась вплоть до появления общей теории относительности. Таинственные законы квантовой механики тоже принимаются, хотя и с некоторыми сомнениями, поскольку позволяют предсказывать с невероятной точностью.

Социальная наука, апеллирующая к рациональному выбору, не может полагаться ни на одно из этих двух оснований. *Нет общего неинтенционального механизма, который мог бы симулировать или имитировать рациональность*. В некоторых случаях этого можно добиться закреплением знаний (глава XVI), тогда как в других оно продуцирует систематические отклонения от рациональности. В иных обстоятельствах эту задачу в какой-то мере выполняет некий социальный аналог естественного отбора, если частота изменений среды превышает скорость адаптации к ним (глава XVII). Мне неизвестны механизмы, которые симулировали бы рациональность в единичных ситуациях или в быстро меняющейся среде. В то же время эмпирическое подтверждение анализа сложных

²⁰ При первом упоминании в книге слова *агент* (agent) следовало бы заметить, что многие исследователи предпочитают ему слово *актор* (actor). Возможно, термином «агент» чаще пользуются экономисты, тогда как социологи используют *актор*. Хотя не имеет большого значения, какой именно термин применяется, я предпочитаю понятие «агент», так как оно подразумевает некую инстанцию действия (agency), понятие «актор» в свою очередь предполагает публику, которая может присутствовать или отсутствовать.

явлений на основании теории рационального выбора, как правило, крайне слабо. Это, разумеется, огульное утверждение. Вместо того чтобы пояснить, что я понимаю под слабостью, позвольте просто указать на высокий уровень разногласий среди специалистов относительно объяснительной силы конкурирующих гипотез. Даже в экономике, в некотором смысле наиболее развитой общественной науке, существуют фундаментальные, устоявшиеся расхождения между школами. Нам *никогда* не достичь точности в нескольких десятых, которая положила бы этим спорам конец.

В-пятых, каузальное объяснение следует отличать от *статистических объяснений* (statistical explanations). Хотя многие объяснения в социальных науках носят характер статистических, они неудовлетворительны, поскольку не могут объяснить конкретные события. Применение статистических обобщений к частным случаям является грубой ошибкой не только в науке, но и в повседневной жизни²¹.

Предположим, что мужчины действительно агрессивнее женщин. Сказать разгневанному мужчине, что его гнев вызван мужскими гормонами, вместо того чтобы заявить, что его гнев не оправдан в данной ситуации, означает впасть в интеллектуальное и моральное заблуждение (fallacy). Предполагать, что обобщение, справедливое для большинства случаев, верно в каждом отдельном случае, — это интеллектуальное заблуждение. Моральное заблуждение — считать, что собеседник находится во власти биологических факторов, а не открыт для доводов и убеждения.

Хотя статистические объяснения — это всегда объяснения «второго сорта», на практике может оказаться, что у нас нет ничего лучше. Однако важно отметить, что в таких объяснениях неизменно руководствуются наилучшим идеалом каузального объяснения. То, что в демократических странах продолжительность жизни больше, чем при недемократических режимах, представляется статистическим фактом. Прежде чем заключить, что продолжительность жизни предопределяется политическим режимом, мы должны *проверить* другие переменные, которые могут оказывать влияние на результат. Может выясниться, что

²¹ Следует избегать и противоположной ошибки — использования частного случая для получения или обоснования обобщения. Пруст писал, что Франсуаза, служанка в семье рассказчика, «была в равной мере склонна путать частное с общим и общее с частным». Такое сочетание может оказываться губительным. Предположим, вы наблюдаете, как член группы X солгал. При обобщении у вас складывается убеждение, что члены группы X склонны ко лжи. Далее, наблюдая за другим членом группы, вы предполагаете, что он лжет. Наконец, (непроверенное) предположение используется в качестве дальнейшего основания для обобщений.

демократические страны в большей мере, чем недемократические, обладают некоторым качеством X и что на самом деле это качество X определяет продолжительность жизни. Но таких свойств бесконечное множество. Как узнать, какое именно нужно проверять? Ответ очевиден: мы должны руководствоваться каузальной гипотезой. Вполне правдоподобно, например, что граждане в индустриальных обществах живут дольше, чем граждане в менее развитых странах. Тот факт, что индустриальные общества бывают более демократическими, чем неиндустриальные, может объяснить наблюдаемые явления. Чтобы убедиться, что именно демократия, а не индустриализация, является каузальным фактором, нам следует сравнить демократические и недемократические режимы с равным уровнем индустриализации и посмотреть, сохранится ли различие. Как только мы будем иметь обоснованную уверенность в том, что проверили другие вероятные причины, можем попробовать выяснить, как, то есть при помощи каких причинно-следственных цепочек или механизмов, тип политического режима влияет на продолжительность жизни. Я буду рассматривать этот второй шаг в следующей главе. Здесь я только хочу заметить, что наша уверенность неизбежно основывается на каузальных догадках о том, что является (и что не является) правдоподобными третьими факторами, которые нуждаются в проверке²².

В-шестых, объяснения следует отличать от *ответов на вопрос «почему?»* (answers to why questions). Предположим, мы читаем научную статью и, к своему изумлению, обнаруживаем, что автор не ссылается на важную и релевантную статью. У нас возникает вопрос: «Почему он ее не упомянул?» Наше любопытство может быть полностью удовлетворено, если мы узнаем, что на самом деле автор ничего не знал о более ранней работе (хотя нам, возможно, захочется выяснить, почему он не изучил литературу по теме с большей тщательностью). Но утверждение: «Он не упомянул статью, потому что не знал о ней», — это не объяснение. Использование его в качестве объяснения означает, что для объяснения одного несовершившегося события ссылаются на другое. Предположим, однако, что автор знал о статье, но *решил* ее не упоминать, поскольку сам не был в ней упомянут. В этом случае ответ на вопрос «почему?» даст, помимо прочего, объяснение. Имеется событие — решение не упоминать о статье, вызванное предшествующим событием — негодованием из-за отсутствия упоминания.

²² Например, нет достоверного причинно-следственного механизма, который заставил бы нас проверять демократические и недемократические режимы на предмет численности населения. Хотя мы не можем исключать причинной связи между количеством жителей и средней продолжительностью жизни, социальная наука такой связи пока не установила; кроме того, мне она представляется надуманной.

Наконец, каузальные объяснения следует отличать от *предсказаний* (predictions). Иногда мы можем дать объяснение, не имея возможности предсказать, а иногда предсказать, не имея возможности объяснить. Верно, что во многих случаях одна и та же теория позволяет нам делать и то и другое, но я полагаю, что в социальных науках это скорее исключение, чем правило.

Отложим основное рассмотрение того, почему мы можем обладать способностью к объяснению, не обладая при этом способностью к предсказанию, до следующей главы. Дело в том, что во многих случаях мы можем выявить причинно-следственный механизм постфактум, но не в состоянии заранее предсказать, какой из нескольких возможных механизмов сработает. От этого до некоторой степени отличается особый случай биологического объяснения. Как будет показано в главе XVI, эволюция работает на двойном механизме — случайных мутаций и (более или менее) детерминированного отбора. Зная некоторые особенности, или модель поведения, организма, мы можем объяснить его происхождение, ссылаясь на случайные изменения в генетическом материале и на их *устойчивость* (persistence) ввиду благотворного влияния на репродуктивную способность. Но никто не может предсказать мутацию, прежде чем она произойдет. Более того, поскольку одна мутация задает возможные последующие, мы можем оказаться не в состоянии сказать, произойдет ли она в принципе. Таким образом, в биологии структурные объяснения малопродуктивны. Некоторый оттенок структурности есть в явлении *конвергенции*, когда разные виды приспосабливаются сходным образом, поскольку находятся в близких условиях, но этот феномен не позволяет сказать, что адаптация была неизбежна.

И наоборот, мы можем обладать способностью к предсказанию, не обладая способностью к объяснению. Чтобы предсказать, что потребители станут покупать меньше товаров, когда их цена возрастет, нет необходимости изобретать объясняющие их поведение гипотезы. Каковы бы ни были мотивы индивидуального действия — рациональные, традиционные или просто случайные, мы можем спрогнозировать, что люди станут покупать меньше товаров просто потому, что меньше смогут себе позволить (глава IX). Здесь есть несколько механизмов, которые вынужденно приводят к одному и тому же результату, так что не придется выбирать между ними, чтобы сделать прогноз. Но для объяснения важен сам механизм. Он дает понимание, тогда как предсказание в лучшем случае обеспечивает контроль.

Чтобы делать предсказания, кроме того, нет необходимости проводить различие между корреляцией, необходимостью и объяснением. Если есть близкая к закономерности повторяемость событий одного и другого типа, для предсказания неважно, связана она с каузальными

отношениями между ними или с тем, что они являются общим следствием третьей причины. В обоих случаях мы можем использовать факт первого события, чтобы предсказать второе. Никто не считает, что первые симптомы смертельной болезни в дальнейшем вызывают смерть, и все же они регулярно используются для такого предсказания. Точно так же если наши знания о состоянии здоровья человека позволяют сделать прогноз, что он не проживет и года, предсказание не становится ложным, если он погибнет в автокатастрофе или покончит с собой из-за невыносимых болей.

Библиографические примечания

Общие взгляды на объяснение и причинность, на которые я полагаюсь, более подробно изложены в: Эльстер Ю., Фёллесдал Д., Валлоэ Л. «Рациональная аргументация» (Elster J., Føllesdal D., Walløe L. *Rationale Argumentation*. Berlin: Gruyter, 1988) (английский перевод готовится). По вопросам применения этих взглядов к действиям человека отсылаю читателя к Д. Дэвидсон «Эссе о действиях и событиях» (Davidson D. *Essays on Actions and Events*. Oxford University Press, 1980). Мою критику функционального объяснения можно найти в различных работах, в частности в «Объясняя техническое изменение» (Explaining Technical Change. Cambridge University Press, 1983). Подробности по делу Китти Дженовезе можно найти в А. М. Розенталь «Тридцать восемь свидетелей» (Rosenthal A. M. *Thirty-Eight Witnesses*. Berkeley: University of California Press, 1999). Удачное изложение взглядов Фестингера приводится в «Раздвигая границы психологии: Избранные работы Леона Фестингера» под ред. Л. Фестингера, С. Шахтера и М. Гадзанига (Festinger L., Schachter S., Gazzaniga M. (eds). *Extending Psychological Frontiers: Selected Works of Leon Festinger*. New York: Russell Sage, 1989). Примеры воздействия детей на родителей взяты из двух увлекательных книг Дж. Р. Харриса «Предположения о воспитании потомства: почему дети становятся такими, какими становятся» (Harris J. R. *The Nurture Assumption: Why Children Turn Out the Way They Do*. New York: Free Press, 1998) и «Двух одинаковых нет» (No Two Alike. New York: Norton, 2006). Взгляды Токвиля на причинность я обсуждаю в работе «Модели каузального анализа в „Демократии в Америке“ Токвиля» (Patterns of causal analysis in Tocqueville's «Democracy in America» // *Rationality and Society*. 1991. No. 3. P. 277–97), а его взгляды на революцию в: «Токвиль о 1789 годе» (Tocqueville on 1789 // Welch C. (ed.). *The Cambridge Companion to Tocqueville*. Cambridge University Press, 2006). Защита Милтоном Фридманом рациональности по модели «как если бы» взята из «Методологии позитивной экономической науки» (Фридман М. *Методология позитивной экономической науки* // *THESIS*. 1994. No. 4. С. 20–52). Недавняя защита подхода «как если бы» в политических науках есть в работе Р. Мортон «Методы и модели: руководство по эм-

пирическому анализу формальных моделей в политологии» (Morton R. Methods and Models: A Guide to the Empirical Analysis of Formal Models // Political Science. Cambridge University Press, 1999). Как и большинство защитников этого подхода, она не указывает причину, по которой мы должны верить в вымысел «как если бы». Частичное объяснение есть в статье Д. Заца, Дж. Фереджона «Рациональный выбор и социальная теория» (Satz D., Ferejohn J. Rational Choice and Social Theory // Journal of Philosophy. 1994. No. 91. P. 71–87). Обсуждение вопросов «почему?» основано на книге Б. Хансона «Объяснения почему» (Hansson B. Why explanations), готовящейся к выходу в Theoria. Независимость закона спроса от мотивационных допущений отмечалась в статье «Иррациональное поведение в экономической теории» Г. Беккера (Becker G. Irrational behavior in economic theory // Journal of Political Economy. 1962. No. 70. P. 1–13).

II. Механизмы

ОТКРЫВАЯ ЧЕРНЫЙ ЯЩИК

Философы науки часто утверждают, что объяснение должно основываться на *общем законе* (general law). Объяснить событие — значит привести ряд начальных условий вместе с высказыванием, согласно которому, как только эти условия выполняются, происходит событие данного типа. В этой главе я выдвину два возражения против этой идеи, одно умеренное и относительно обтекаемое, второе — более радикальное и полемичное.

Первое возражение заключается в том, что даже если мы можем установить общий закон, из которого можем вывести экспланандум (второе возражение предполагает, что не всякий раз мы можем это сделать), эта процедура не всегда дает нам объяснение. Здесь мы снова можем сослаться на различие между объяснением, с одной стороны, и корреляцией и необходимостью — с другой. Общий закон, согласно которому за некоторыми симптомами болезни неизменно следует смерть, не объясняет, от чего умер человек. Общий закон, основывающийся на фундаментальной природе болезни, не объясняет смерть, если действие болезни упреждается самоубийством или автомобильной аварией.

Чтобы обойти эти проблемы, часто предлагают заменить идею общего закона идеей *механизма* (mechanism). Поскольку позднее я буду употреблять слово *механизм* как специальный термин, здесь я воспользуюсь выражением «причинно-следственная цепочка» (causal chain), чтобы передать, что я имею в виду¹. Вместо того чтобы объяснять событие *E* утверждением «всегда, когда происходят события C_1, C_2, \dots, C_n , следует событие типа *E*», можно попытаться установить причинно-следственную цепочку, которая ведет от причин C_1, C_2, \dots, C_n к *E*. Такой шаг часто называют «открытие черного ящика» (opening the black box). Предположим, мы знаем, что вероятность заработать рак легких для заядлых ку-

¹ Иногда в ранних работах я использовал понятие «механизм» в том значении, в котором сейчас использую выражение «причинно-следственная цепочка». В более поздних работах я стал использовать понятие «механизм» в значении, которое будет определено в этой главе. Вероятно, мне следовало бы выбрать иную терминологию, но уже слишком поздно.

рильщиков выше, чем для других. Этот факт вызван тем, что либо курение — причина рака легких, либо люди, предрасположенные к курению, предрасположены еще и к раку легких (возможно, гены, вызывающие предрасположенность к раку, связаны с генами, из-за которых у некоторых людей развивается никотиновая зависимость)². Чтобы обосновать первое объяснение, нам придется выявить цепочку физиологических причинно-следственных связей, которая начинается с заядлого курения и заканчивается раком легких. Окончательное объяснение будет более изящным, в нем будет больше каузальных звеньев, и оно будет более убедительным, чем непрозрачное утверждение «Курение вызывает рак».

Или предположим, что некто утверждает: высокий уровень безработицы ведет к агрессивным войнам, — и представляет доказательства напоминающей закон связи между этими двумя явлениями. И снова: откуда мы можем знать, что это причинная связь, а не простая корреляция? Возможно, высокий уровень рождаемости, вызывающий безработицу, помимо прочего, толкает политических лидеров на развязывание войн? Неудачные войны по крайней мере сократят численность населения, а успешные обеспечат новые территории для экспансии и миграции. Чтобы исключить такую возможность, мы должны прежде всего проанализировать рождаемость (и другие правдоподобные третьи факторы) и посмотреть, сохраняется ли подобная связь. Если да, мы все равно не должны останавливаться на достигнутом до тех пор, пока нам не удастся заглянуть в черный ящик и сказать, *каким образом* высокая безработица провоцирует войны. Происходит ли это потому, что безработица побуждает политических лидеров завоевывать новые рынки? Или потому, что они считают, что она создает социальное напряжение, которое следует канализировать и направить на внешнего врага, чтобы предотвратить возникновение революций у себя дома? Или потому, что они думают, что безработицу может абсорбировать военная промышленность? Или же это происходит потому, что безработные склонны голосовать за популистских лидеров, которые избегают дипломатии и предпочитают военные методы разрешения конфликтов?

Рассмотрим подробнее последнее предположение. Почему безработные должны голосовать за безответственных популистских лидеров, а не за политиков из уважаемой партии? Можно опять-таки представить несколько способов вскрытия этого черного ящика. Возможно, электорат популистских политиков чаще ходит на выборы, когда у него нет работы, потому что это обходится ему дешевле (ниже стоимости его времени), чем когда у него есть работа. Или, возможно, потому, что по-

² Как будет отмечено ниже, когда-то это объяснение рассматривали вполне серьезно.

литики-популисты чаще предлагают немедленные решения проблемы безработицы. Или же они предлагают политику, которая накажет тех, кто по мнению безработных виновен в их бедственном положении и извлекает из него выгоду, будь то капиталисты или экономически преуспевающие этнические меньшинства.

Рассмотрим более детально последнее предположение. Почему у безработных должно возникнуть желание наказать капиталистов или процветающие меньшинства? Не является ли это очередным утверждением из черного ящика? Его можно объяснить меркантильными интересами, которыми руководствуются безработные. Если бы государство конфисковало имущество элит, средства можно было бы перераспределить с выгодой для безработных. Возможно, они ведомы чувством мести, которое побуждает их наказывать элиты, даже если это не принесет им материальной выгоды. Если считается, что богатые безжалостно сокращают персонал для увеличения своих прибылей, потерявшие работу могут воспользоваться избирательной урной для того, чтобы с ними покончить. Или же безработные могут попросту испытывать зависть к ловким представителям меньшинств, добившимся успеха там, где они сами потерпели неудачу, и использовать выборы для того, чтобы поставить их на место.

Насколько мне известно, высокий уровень безработицы не вызывает войн. Это упражнение чисто гипотетическое. И все же я полагаю, что оно подтверждает идею о том, что правдоподобие объяснения растет по мере того, как общие законы истолковываются в категориях причинно-следственной цепочки. На уровне общих законов мы никогда не можем твердо знать, что проверили все третьи факторы, относящиеся к делу. Где-то на периферии всегда может притаиться факт, объясняющий экспланандум и его предполагаемую причину. Увеличивая число звеньев в цепочке, мы уменьшаем эту опасность.

Тем не менее опасность не может быть устранена полностью. Установить причинно-следственную цепочку — не значит совершенно отказаться от общих законов, но лишь перейти от их высокого абстрактного уровня к уровню более низкому. Мы, например, можем заменить универсальный закон «Высокий уровень безработицы вызывает войны» менее абстрактными — «Популистские лидеры склонны к войнам» или «Безработные голосуют за популистских лидеров». Последний в свою очередь может быть замещен соединением законов «Безработные испытывают зависть к богатым меньшинствам» и «Те, кто завидует богатым меньшинствам, голосуют за популистских лидеров». Как и в случае любого другого закона, все это может оказаться простой корреляцией. Если зависть к меньшинствам и отсутствие работы — следствия общей причины, успех на выборах агрессивных лидеров будет связан не с безработицей, а с фактором, каузально коррелирующим с ней. И все же на более

глубоком уровне меньше факторов, нуждающихся в проверке. Чем лучше мы сформулируем каузальную задачу, тем легче будет удостовериться в том, что перед нами не просто корреляция.

Объяснения через (очень) общие законы неудовлетворительны еще и потому, что они слишком непрозрачны. Даже если нам предоставят неопровержимое доказательство универсальной связи между безработицей и агрессивными войнами и убедительное свидетельство того, что были проверены все сколько-нибудь правдоподобные третьи факторы, мы все равно захотим узнать, *как именно* безработица вызывает войны. Мы можем верить в то, что объяснение правильное, и все же считать его недостаточным. Как я отмечал в предыдущей главе, таков был статус объяснений, основанных на законах гравитации, до появления общей теории относительности. Действие на расстоянии было столь загадочным, что многие отказывались верить в научную решенность этого вопроса. Поскольку этот закон позволял делать предсказания с точностью до десятых, скептикам пришлось согласиться, что все происходит так, «как если бы» это была правда, хотя они и не принимали существование силы, действующей там, где ее нет.

МЕХАНИЗМЫ

Читатели могут указать, что приведенные в этом упражнении примеры якобы универсальных законов крайне неправдоподобны. Соглашусь. Частично это происходит из-за того, что мне не хватило воображения при подборе примеров, но я полагаю, что есть и более глубокие причины. В социальных науках попросту слишком мало твердо установленных (well-established) общих законов. Закон спроса — при росте цен покупательная способность падает — хорошо обоснован (well supported), но как закон довольно слаб³. Закон всемирного тяготения, например, не только утверждает, что по мере увеличения расстояния между объектами сила притяжения между ними уменьшается; он позволяет сказать, *насколько* она *уменьшается* (обратно пропорционально квадрату расстояния между ними). В социальных науках ничего подобного нет⁴.

³ Более того, спрос на некоторые виды товаров растет при повышении цены на них. Потребителей может привлечь дороговизна товара (эффект Веблена) или они могут начать меньше покупать, например, хлеб, когда цены на него падают, потому что могут заменить его товарами более высокого качества.

⁴ Разумеется, часто говорят, что сила альтруизма по отношению к другим обратно пропорциональна социальной дистанции, отделяющей этих других от агента. Однако идея социальной дистанции здесь скорее метафора, чем концепция, и в любом случае формулировка «обратно пропорциональна» лишена точности формулы «обратно пропорционально квадрату расстояния».

Закон спроса и закон Энгельса, согласно которому доля доходов, идущая на еду, уменьшается по мере их роста, являются тем, что можно назвать *слабыми законами* (weak law). Для любого изменения (увеличения или уменьшения) независимой переменной они позволяют предсказать *направление*, знак изменения (увеличения или уменьшения) в зависимой переменной, однако не позволяют предсказать *величину* этого изменения. Будучи слабыми, эти законы все же обладают некоторым содержанием, так как позволяют нам исключить целую область возможных значений зависимой переменной, но не позволяют нам выделить значение, которое будет реализовано в неисключенной области.

Закон спроса не только слаб, но и плохо приспособлен для целей объяснения. Как мы видели в главе I, он совместим с несколькими предположениями относительно того, как поведут себя потребители. Чтобы *объяснить*, почему потребители меньше покупают товар по мере роста его цены, нам пришлось *бы* применить и проверить специальное предположение, касающееся реакций индивидуального потребителя на ценовые колебания. Ключевое слово здесь *индивидуальный*. В социальных науках удовлетворительное объяснение должно быть привязано к гипотезам об индивидуальном поведении. Этот принцип, известный под названием «методологический индивидуализм», является исходным предположением этой книги. Он гласит, что психология и, возможно, биология имеют фундаментальную важность для объяснения социальных явлений. Если у меня есть сомнения насчет биологии, то не потому, что я считаю ее неспособной объяснять аспекты человеческого поведения в принципе, но потому, что я думаю, что во многом она недостаточна разработана для такой задачи.

Для объяснения индивидуального поведения мы должны, главным образом, полагаться на то, что я называю *механизмами*. Грубо говоря, механизмы — это *часто повторяющиеся или легко узнаваемые каузальные модели* (causal patterns), которые *срабатывают при неизвестных в целом условиях или с неопределенными последствиями*. Они позволяют объяснять, но не предсказывать. Например, утверждалось, что на каждого ребенка, вырастающего в среде алкоголиков и становящегося им, приходится один его сверстник, который избегает алкоголя в результате реакции на то же самое окружение. Обе реакции воплощают следующие механизмы: делать то же самое, что делали родители; действовать противоположным образом. Мы не можем определить заранее, что станет с ребенком алкоголика, но если он станет трезвенником либо алкоголиком, мы предположительно знаем, почему.

Я не утверждаю, что здесь работает какая-то объективная неопределенность. В действительности эта концепция не имеет особого смысла вне рамок квантовой механики. Я утверждаю, что часто мы можем объ-

яснить поведение, показав, что оно является следствием общей каузальной модели, даже если не в состоянии объяснить, почему срабатывает эта модель. И механизм конформизма (например, делать то же, что делают ваши родители), и механизм нонконформизма (делать нечто противоположное тому, что делают они) являются очень общими. Если мы можем показать, что поведение ребенка родителя-алкоголика является проявлением того или иного механизма, значит, у нас уже есть объяснение поведения. Могут возразить, что если мы не показали, почему ребенок становится, скажем, алкоголиком, а не трезвенником, мы ничего не объяснили. Я, конечно, соглашусь, что объяснение, показывающее, почему результат был именно таким, а не иным, лучше, и не стану отрицать, что иногда мы в состоянии дать такое объяснение. Но подвести конкретный пример под более общую каузальную модель — тоже означает дать объяснение. Знать, что ребенок станет алкоголиком под действием конформизма, — значит до некоторой степени прояснить этот результат, хотя неясность сохранится до тех пор, пока мы не объясним, почему ребенок выбрал тот или иной путь.

Как я сказал, механизм — это часто повторяющаяся и легко узнаваемая каузальная модель. Народная мудрость в пословицах выявила много таких моделей⁵. В моем любимом определении это звучит так: «Пословица передается от поколения к поколению. Она резюмирует в одной короткой фразе общий принцип или распространенную ситуацию, и когда вы ее произносите, все точно знают, что вы имеете в виду». Более того, пословицы часто устанавливают механизмы (в том смысле, в котором мы употребляем это слово), но не общие законы. Рассмотрим, в частности, поразительную склонность пословиц образовывать взаимоисключающие пары. С одной стороны, «В разлуке чувство крепнет», с другой — «С глаз долой, из сердца вон». С одной стороны, мы считаем, что «Запретный плод сладок», с другой — «Видит око, да зуб неймет». С одной стороны, «Подобное привлекает подобное», с другой — «Противоположности сходятся». С одной стороны, «Яблоко от яблони недалеко падает», с другой — «У отца-скряги сын — мот». С одной стороны, «Поспешешь — людей насмешишь», с другой — «Промедление смерти подобно». С одной стороны, «Помяни лихо, и оно тут как тут», с другой — «Воспоминания о прошлых опасностях приятны». (Как отмечалось выше, два последних высказывания не являются взаимоисключающими.) Можно привести множество других примеров.

Не все пары противоположных механизмов отражены в пословицах. Рассмотрим, например, так называемую, пару «избыток — компенса-

⁵ Как мы увидим в главе XX, пословицы, тем не менее, не всегда являются мудрыми.

ция» (spillover — compensation). Если человек, который очень много работает, отправляется в отпуск, будет он отдыхать в таком же лихорадочном темпе (эффект избытка) или же, напротив, совершенно расслабится (эффект компенсации)? Будут граждане в демократических странах тянуться к религии или, наоборот, питать к ней отвращение? Если их привычка принимать самостоятельные решения переносится из политической сферы в религиозную (избыток), можно ожидать, что религиозная вера будет слабой. Если нехватка верховной власти в политике толкает их на поиски власти где-нибудь еще (компенсация), демократический политический режим будет скорее благоприятствовать религии. Современный вопрос, на который до сих пор так и не найдено ответа: насилие на экране стимулирует насилие в реальной жизни (перелив) или же смягчает его (компенсация)?

Сходные механизмы применимы к отношениям между индивидами. Рассмотрим случай благотворительности. Кого-то может больше всего занимать эффективность пожертвования. Если другие жертвуют мало, его дар будет более весомым, поэтому он с большей вероятностью сделает взнос. Если же другие дают много, его пожертвование будет менее значимым, поэтому он может отказаться в этом участвовать. Другой даритель может быть озабочен равенством вкладов. Он не понимает, почему должен давать больше в сравнении с остальными. И наоборот, если другие делают крупные пожертвования, он может последовать их примеру. Та же самая пара механизмов может сработать в ситуациях коллективного действия. По мере роста популярного движения некоторые индивиды могут от него отойти, потому что полагают, что их участие уже ничего не изменит, тогда как другие, наоборот, присоединятся из чувства, что не могут оставаться в стороне от общего дела (глава XXIV).

Даже пословицы, не имеющие противоположной пары, зачастую отражают механизмы, а не законы. Пословица «Чаще тонут лучшие пловцы» абсурдна, если понимать ее как констатацию того, что вероятность утонуть пропорциональна уровню навыков пловца. Но может быть и так, что уверенность пловцов в своих навыках опережает в росте сами эти навыки, тем самым побуждая идти на неоправданный риск («Гордыня до добра не доведет»). Или рассмотрим пословицу, к которой я еще не раз вернусь в этой книге, — «Мы легче всего верим в то, на что надеемся, и в то, чего боимся»⁶. Хотя эта пословица неправдоподобна, если толковать ее

⁶ Хотя эта пословица и не задает пару противоположных положений, мы могли бы вообразить пословицы, гласящие, что «Мы легко можем поверить в то, на что надеемся» и «Мы легко теряем веру в то, на что надеемся». Мы действительно очень часто испытываем предубеждение против того, на что надеемся, потому что делаем поправку на возможное принятие желаемого за действительное.

буквально, как универсальный закон, она является хорошим напоминанием о том, что в дополнение к общеизвестному феномену самообмана, когда желаемое принимается за действительное (*wishful thinking*), существует менее известная склонность к тому, что можно назвать *контрмотивированным самообманом* (*countermotivated thinking*)⁷. Наконец, рассмотрим ставшее пословицей утверждение «У семи нянек дитя без глаза».

Повторюсь: ценность пословиц не в установлении универсального закона, а в выявлении механизмов. Пословица может быть верна, если каждая из нянек полагает, что за ребенком присматривает кто-то другой (вспомним случай с Кити Дженевезе).

Давая определение механизмам, я также отметил, что они срабатывают при неизвестных в целом условиях или с неопределенными последствиями. Большинство механизмов пословиц, которые я перечислил, попадают в первую категорию. Мы не знаем, какие условия вызывают конформизм или неконформизм, принятие желаемого или нежелаемого (контрмотивированного) за действительное, адаптивные («Видит око, да зуб неймет») или контрадаптивные предпочтения («У соседа трава зеленее»). Мы знаем, что будет реализована в лучшем случае одна часть пары, но не можем сказать, какая именно. Оговорка «в лучшем случае» связана с тем, что некоторые люди не подвержены воздействию ни одной из составляющих пару механизмов. Настоящая автономия предполагает, что человек не является ни конформистом, ни неконформистом. Убеждения могут быть независимыми от желаний, а желания — независимыми от возможности их исполнения.

В других случаях пословицы предполагают одновременный запуск двух механизмов, которые оказывают разнонаправленное действие на результат. В этом случае неопределенность связана с определением *результатирующего эффекта* механизмов, а не в выяснении, какой из них сработает (если сработает). Рассмотрим для примера пословицу «Голь на выдумку хитра» и «Быть бедным — дорого». Первая пословица постулирует наличие причинно-следственной связи между бедностью и *стремлением* к инновациям, вторая — между бедностью и *недостаточными возможностями* для инноваций. Так как поведение формируется как желаниями, так и возможностями (глава IX), мы не можем сказать, является результирующее влияние бедности на инновации не-

⁷ Можно дать расширенное толкование этой идеи, включающее одновременно и неверие в то, на что надеются, и веру в то, чего боятся. В предыдущем примечании я указал на то, из чего может возникнуть первая разновидность. Что касается второй разновидности, то возможно, что желание подготовиться к худшему усиливает веру в то, что произойдет худшее. Сам процесс мысленного прокручивания сценария может отпечататься в мозге и поднять статус с отдаленно возможного до правдоподобного и даже вероятного.

гативным или позитивным. Или рассмотрим пару пословиц, упоминавшуюся ранее, — «Помяни лихо, и оно тут как тут» и «Воспоминания о прошлых опасностях приятны». Первая пословица опирается на то, что называли эффектом вклада (endowment effect): память о неприятном опыте неприятна⁸. Вторая основана на эффекте контраста (contrast effect): память о неприятном опыте повышает ценность настоящего⁹. Мы не можем сказать, позитивным или негативным окажется результирующий эффект предшествующего неприятного опыта на будущем благополучии.

Оговоримся еще раз: мы не ограничиваемся пословицами. Рассмотрим, к примеру, два не упоминаемых в пословицах механизма, задействованных в так называемой психологии тирании. Если тиран усиливает давление на подданных, возможны две реакции. С одной стороны, ужесточение наказаний может отвратить подданных от сопротивления или бунта. С другой стороны, чем жестче тирания, тем сильнее ненависть подданных. Тиран, как любой уличный задира, может внушать как ненависть, так и страх. Если ненависть преодолет страх, угнетение обернется против тирана. В странах, оккупированных Германией во время Второй мировой войны, члены движения Сопротивления иногда пользовались этим механизмом, когда убивали немецкого солдата, чтобы спровоцировать карательные акции, предполагая, что эффект тирании возьмет верх над эффектом устрашения¹⁰. Рассмотрим также в чем-то похожий случай человека, который встречает препятствие на пути к достижению цели. Эта угроза свободе действий может вызвать то, что психологи называют реактивным сопротивлением (reactance), попыткой вернуть или восстановить свободу. Эффект препятствия и последующего реактивного сопротивления противоположны друг другу, и обычно мы не можем сказать, какой из них окажется силь-

⁸ И наоборот, память о приятном опыте сама по себе приятна. Отсюда фраза Теннисона «Лучше любить и потерять, чем не любить никогда» («Tis better to have loved and lost than never to have loved at all»).

⁹ И наоборот, память о приятном опыте снижает ценность настоящего. Отсюда фраза Донна «Лучше быть подлым сейчас, чем честным когда-то» («Tis better to be foul than to have been fair»).

¹⁰ Иногда ненависть находит себе другую цель. В трех деревнях центральной и северной Италии, где немцы в 1944 году провели жестокие карательные акции, некоторые жители даже пятьдесят лет спустя сохраняли неприязнь к партизанам, так как считали их косвенно виновными в резне. Когда А вынуждает Б убить В, друзья и родственники В могут направить свой гнев скорее на А, чем на Б. В движении Сопротивления сражавшихся с нацистами можно было наблюдать оба механизма.

нее¹¹. В качестве иллюстрации представьте, какой эффект может вызвать поведение родителей, прячущих от маленького мальчика барабан, чтобы он на нем не играл¹².

Даже зная результирующий эффект, мы можем оказаться не в состоянии его объяснить. Предположим, мы получили возможность каким-то образом оценить и измерить нулевое результирующее воздействие эффекта вклада и эффекта контраста по отношению к приятному опыту в прошлом. Результат может быть двояким. Хотя обед в трехзвездочном французском ресторане в прошлом году может уменьшить мое удовольствие от последующих посещений более скромных французских заведений, это негативное впечатление заслоняется памятью о том, каким замечательным был обед. Однако наблюдаемый нулевой результирующий эффект совершенно не противоречит нулевым эффектам вклада и контраста, как и в случае, если оба эффекта окажутся одинаково сильными. Пока мы не знаем, с какой ситуацией имеем дело, мы не можем претендовать на объяснение результата. Чтобы оценить силу каждого эффекта, мы можем взглянуть на результат в ситуации, для которой известно, что другой эффект в ней действовать не может. Если на мое удовольствие от греческой кухни скорее всего никак не влияет обед в трехзвездочном французском ресторане, мы можем оценить силу эффекта вклада в чистом виде.

Связанная с этим неопределенность может возникнуть по отношению к первому типу механизмов, которые запускаются при обычно неизвестных условиях. Вернемся к случаю с родителем-алкоголиком. Если рассматривать общее число алкоголиков (или большой репрезентативный срез этой группы), то можно предположить, что их дети в среднем пьют не больше и не меньше, чем дети непьющих родителей. Если для простоты не учитывать влияние факторов наследственности, эти результаты могут толковаться двояко. С одной стороны, может быть так, что дети алкоголиков не являются ни конформистами, ни нонконформистами, то есть их поведение может определяться теми же причинами, что и поведение детей непьющих родителей. С другой стороны, возможно, половина детей алкоголиков — конформисты, а другая половина — нонконформисты, что дает результирующий эффект, равный нулю.

¹¹ Особенность этого примера в том, что один из двух конкурирующих эффектов (реактивное сопротивление) спровоцирован другим (препятствием). В другом примере два эффекта продуцируются одновременно общей причиной (например, тираническим гнетом).

¹² Еще один пример реактивного сопротивления: дети (а также взрослые) могут отказаться от любых предложений, касающихся их времяпровождения, даже от тех, *которые им нравятся*, если чувствуют, что на них давят.

Сходным образом теории поведения избирателей выявили эффект «темной лошадки» (*underdog mechanism*) и эффект «стадности» (*bandwagon mechanism*). Те, кто подвержен воздействию первого механизма, склонны голосовать за кандидата, который отстает по результатам предвыборных опросов, тогда как подверженные воздействию второго голосуют за лидера гонки. Если два этих типа равномерно распределены, заметного результирующего эффекта может и не быть, так что опрос в этом случае позволит легко предсказать реальные результаты. Недостаток влияния опроса на исход голосования в свою очередь не смог бы показать, что индивиды не подвержены влиянию опросов. Слабое совокупное влияние телевизионного насилия на насилие в реальной жизни могло бы маскировать силу его воздействия на отдельные социальные группы. Во всех этих случаях нейтральная сумма отражала бы либо однородность общества, состоящего из индивидов, не подверженных влиянию вообще, либо в той же мере гетерогенное, сильно подверженное влиянию, но в противоположном направлении. Необходимость снять эту двусмысленность дает еще один аргумент в пользу методологического индивидуализма. Для объяснения поведения на уровне целого мы должны смотреть на поведение индивидуальных составляющих.

МОЛЕКУЛЯРНЫЕ МЕХАНИЗМЫ

Я рассматривал то, что может быть названо атомарными механизмами — элементарными психологическими реакциями, которые не могут быть редуцированы к другим механизмам на том же уровне. Может возникнуть вопрос: как далеко нас могут завести эти психологические механизмы при объяснении социальных явлений? Ответ: мы можем использовать атомарные механизмы как кирпичики в более сложных молекулярных механизмах. Попробуем снова начать с пословиц. Одна из них — «У страха глаза велики» — означает, с одной стороны, что страх часто превосходит опасность, с другой — он же ее преувеличивает. Пословица, таким образом, предполагает, что чрезмерный страх воспроизводит сам себя. Английская пословица гласит, что «В семье не без урода» (*дословно: «В каждом стаде есть паршивая овца»*. — *Примеч. пер.*), французская — что «паршивая овца все стадо портит». Соединив их вместе, мы можем сделать вывод, что каждое стадо будет испорчено¹³.

Рассмотрим для примера самоуправляемый коллегиальный орган — факультет университета или рабочий кооператив. Есть довольно четкий

¹³ Я позволяю себе некоторую свободу применительно к этим пословицам. Буквально французское словосочетание «Une brebis galeuse» отсылает к овце с кожной болезнью, вызываемой паукообразным паразитом.

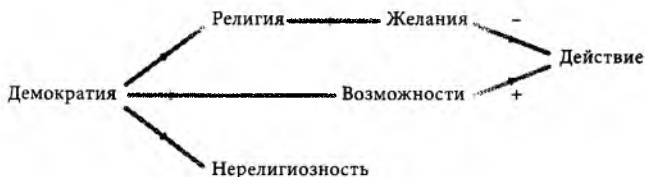


РИС. II.1

и распространенный сценарий, который по праву можно отнести к молекулярным механизмам. Во-первых, по теории вероятности, в любой группе из двадцати и более человек может оказаться один болтливый и своевольный человек, «в каждой бочке затычка» (spoiler), определяемый словарями как «тот, кто портит шансы для победы оппонента, сам не являясь потенциальным победителем». Во-вторых, присутствие такого человека в группе затрудняет самоуправление. Дискуссии продолжаются до бесконечности, принятые решения постоянно оспариваются, неформальная коллегиальность подменяется духом формализма, возникает враждебность и так далее. В конце концов группа приветствует переход к управлению небольшим исполнительным комитетом или даже отдельным индивидом.

Оставив в стороне пословицы, обратимся к другому молекулярному механизму. Многие столетия или тысячелетия элиты с недоверием относились к демократии как к форме правления, считая, что она допускает всевозможные виды опасного поведения или распушенности. Однако возможности для опасного поведения сами по себе его не вызывают, должен быть какой-то мотив. Могут ли демократические режимы каким-то образом сдерживать желания граждан делать то, что позволяет им демократия? В этом заключалась претензия Токвиля. Он думал, что для того чтобы удовлетворить потребность в авторитете, не удовлетворенную политикой, демократические граждане обратятся к религии, имеющей predisposition ограничивать и сдерживать их желания. Он утверждал, что критики демократии ошибаются, потому что сосредотачиваются только на возможностях, игнорируя желания. Хотя он излагал этот аргумент в форме универсального закона, ему можно дать более достоверное объяснение в категориях механизмов. С одной стороны, если работает эффект избытка, а не эффект компенсации, недостаток политического авторитета будет скорее ослаблять, а не усиливать религию; с другой — даже если работает эффект избытка, мы не можем ничего сказать о результирующем воздействии. Если набор возможно-

стей сильно расширен, а желания сдерживаются слабо, результирующий эффект демократии может привести к увеличению, а не к уменьшению распространения данного поведения.

Две пары механизмов суммарно представлены на рис. II.1. Если влияние демократии на религию опосредуется эффектом компенсации, а не эффектом избытка, то демократические общества будут религиозными. Если негативное воздействие демократии на желания (опосредованное религией) достаточно сильно, чтобы свести на нет позитивное воздействие демократии на возможности, то демократические граждане будут вести себя сдержанно¹⁴.

МЕХАНИЗМЫ И ЗАКОНЫ

Часто объяснение через механизмы — это все, что мы можем сделать, но иногда мы все же в состоянии добиться большего. Выявив механизм, запускающийся при неизвестных в целом условиях, мы можем определить сами эти условия. В таком случае механизм будет заменен законом, хотя и слабым в том смысле, который мы указывали выше.

Здравый смысл подсказывает, что получивший подарок должен испытывать благодарность. Если он ее не испытывает, он сам виноват. Классические моралисты — от Монтеня до Лабрюера — утверждали, что подарки вызывают у их получателей скорее обиду, нежели благодарность. Похоже, что моралисты что-то поняли, но не говорят нам, когда следует ожидать тот или иной результат. Классический древнеримский моралист Публий Сир установил условия, *приводящие в действие такой механизм*: небольшой дар создает обязательства, большой — врага¹⁵. Апеллируя к размерам дара как запускающему механизм условию, мы (до некоторой степени) превращаем пару механизмов в высказывание, напоминающее закон¹⁶. Приведу еще один пример. Мы можем быть

¹⁴ Токвиль, кроме того, утверждал, что сдерживающее действие будет оказывать тирания мнения большинства. «Некоторые правительства стремятся беречь нравы, осуждая авторов развратных книг. В Соединенных Штатах никого не приговаривают за такие книги, но ни у кого и не возникает соблазна их написать».

¹⁵ Здесь я немного жульничаю, чтобы получить подходящий пример, поскольку Сир имеет в виду ссуды, а не дары. Хотя то, что и ссуда, и дарение большой суммы денег могут вызвать обиду у их получателя, верно, они скорее всего делают это по-разному.

¹⁶ Только «до некоторой степени», поскольку точный размер дара или ссуды, при котором благодарность превращается в обиду, в значительной мере зависит от контекста. Это замечание относится также к другим обсуждаемым в тексте примерам.

в состоянии указать, когда напряжение между желанием и убеждением (когнитивный диссонанс) разрешается при помощи модификации убеждения, а когда — при модификации желания¹⁷. Чисто фактические убеждения могут оказаться слишком устоявшимися, чтобы их можно было легко изменять (глава VII). Человек, заплативший 75 долларов за билет на бродвейское шоу, не может внушить себе, что заплатил всего 40. Но, как правило, он может найти в шоу привлекательные стороны и убедить себя, что они важнее его недостатков.

Ранее я упомянул контраст между механизмом «запретного плода» («forbidden fruit» mechanism) и механизмом «зеленого винограда» («sour grapes» mechanism). В некоторых случаях у нас есть возможность предсказать, какой механизм сработает. В эксперименте испытуемым в одном случае предлагали дать оценку четырем записям с точки зрения их привлекательности, сказав, что на следующий день им дадут одну выбранную наугад запись. В другом случае предлагали оценить записи, сказав, что на следующий день они смогут сами выбрать одну из них. На следующий день всем испытуемым сообщали, что запись, которую они поставили на третье место, отсутствует, и просили их заново оценить записи, пытаясь среди прочего посмотреть, как повторное прослушивание записи влияет на оценку. Как предсказывала теория реактивного сопротивления, в первом случае испытуемые продемонстрировали эффект «зеленого винограда», снизив оценку отсутствующей опции, тогда как во втором случае проявился эффект «запретного плода», повысивший оценку этой записи (контрольная группа, которой ничего не сказали об исключении одной из записей, не показала никаких изменений). Ключевое различие в том, что вторая группа испытала угрозу своей свободе, а первая — нет.

Однако давайте рассмотрим более сложный пример. Для пословиц «В разлуке чувство крепнет» и «С глаз долой — из сердца вон» есть третья, подсказывающая условие запуска механизма: «Короткая разлука любви порука». Ларошфуко предложил другое условие: «Отсутствие ослабляет умеренные страсти и усиливает более значительные, подобно ветру, задувающему свечу, но раздувающему пожар». Эти правдоподобные положения не являются сильными законами. Чтобы предсказать развитие страсти, нам пришлось бы установить, что считается короткой разлукой (три недели?) и сильной страстью (та, что не дает спать по ночам?). Кроме того, нам пришлось бы определить, как *взаимодействуют* длительность разлуки и сила страсти, усиливая или ослабляя последнюю. Попробуем разобраться в этом вопросе.

¹⁷ Нужно, однако, не забывать, что такое напряжение может остаться и неразрешенным.

ВЗАИМОДЕЙСТВИЕ ПРИЧИН ДРУГ С ДРУГОМ

Социальные науки, как правило, не слишком сильны в объяснении того, как причины взаимодействуют между собой, производя общее следствие. Чаще всего утверждается, что каждая причина влияет на следствие самостоятельно (аддитивная модель). Чтобы раскрыть доход, например, можно принять в расчет, что частично он определяется заработком родителей и частично их образованием, а затем использовать статистические методы, чтобы определить соответствующую долю этих двух причин. Для обсуждавшихся мною примеров эти методы могут оказаться непригодными. Продолжительность разлуки может сама по себе не оказывать воздействия на силу чувств после нее, ее действие может зависеть скорее от интенсивности изначальных эмоций. Этот эффект взаимодействия продемонстрирован на рис. II.2.

Некоторые ученые утверждают, что мир (или по крайней мере, та его часть, которую они изучают) демонстрирует не так уж много примеров подобных взаимодействий. Редко бывает, говорят они, чтобы зависимая переменная Z возрастала (падала) при низком уровне независимой переменной X вместе с зависимой переменной Y , тогда как при высоком уровне X возрастание Y вызывает падение (рост) Z .

Предположительное отношение, представленное на рис. II.2, было бы (если бы существовало) исключением. В лучшем случае, утверждают они,

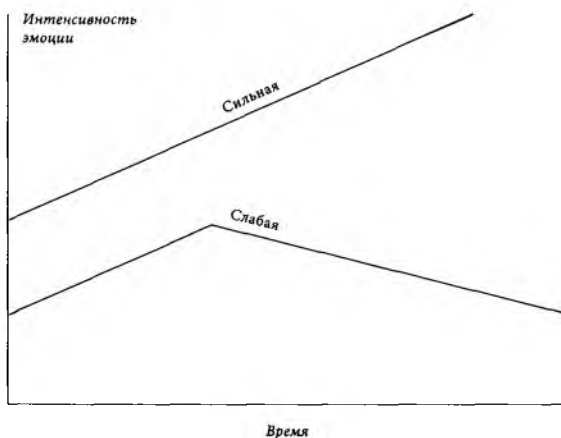


РИС. II.2

мы установим, что на низком уровне XU не оказывают на Z большого воздействия, но отражается на нем при более высоком уровне X . При объяснении дохода, например, можно предположить, что заработок родителей оказывает большее или меньшее влияние при разном уровне родительского образования. Такого рода взаимодействие можно уловить через категорию мультипликативного взаимодействия (multiplicative interaction); таким образом, Z является функцией X , U и XU . Обратное каузальное воздействие — U на Z при высоком уровне X — не может быть передано подобным способом. Однако если мы верим, что такие *перемены* направления происходят редко, нам не приходится слишком о них беспокоиться.

Существование эффекта взаимодействия может быть подвержено того же рода неопределенности, с которой мы обычно сталкиваемся в случае механизмов. Рассмотрим взаимодействие между возрастом и базовыми политическими взглядами как причинами экстремизма. Можно предположить, что молодежные организации будут в большей степени склоняться влево, чем партии, к которым они относятся. Таким образом, молодые консерваторы приобретут светлый оттенок синего, а молодые социалисты — более яркий оттенок розового. Оба предположения кажутся правдоподобными, в реальности наблюдались обе модели. Или рассмотрим взаимодействие между настроением перед приемом наркотиков и самим приемом наркотиков как причинами, определяющими настроение после употребления оных. Можно предположить, что такое вещество, как алкоголь или кокаин, *поднимает настроение*, смягчая депрессию и превращая удовлетворение в эйфорию. Но можно выдвинуть предположение, что сильнодействующие вещества *интенсифицируют настроение*, ухудшая плохое и улучшая хорошее. Опять-таки оба предположения правдоподобны, в реальности наблюдаются обе модели. В обоих случаях первый механизм совместим с аддитивной моделью, тогда как второй предполагает наличие обратного эффекта (reversal effect).

В случае устойчивых данных добавление члена, характеризующего взаимодействие, или подгонка кривой — не единственное возможное решение. Существует альтернативная стратегия — анализ данных (data mining). При начертании кривой зависимые и независимые переменные фиксируют и подыскивают математическую функцию, которая даст хорошее статистическое соответствие при таких показателях. При анализе данных фиксируют математическую функцию (обычно это простая аддитивная модель) и подыскивают независимые переменные, которые хорошо сочетаются с зависимой переменной. Предположим, что хорошим сочетанием мы полагаем корреляцию, для которой вероятность случайного совпадения составляет 5%. В исследовании любого сложного социального явления, такого как доход, можно с легкостью перечислить

дожину переменных, которые могут оказывать свое влияние¹⁸. Кроме того, есть, наверно, полдюжины способов концептуализировать доход. Крайне маловероятно, чтобы при каком-нибудь из определений дохода ни одна из независимых переменных не показала корреляцию на 5%-м уровне¹⁹. Теория вероятности гласит, что самым невероятным совпадением было бы полное отсутствие невероятных совпадений²⁰.

Как только ученый подобрал подходящую математическую функцию или подходящий ряд зависимых или независимых переменных, он может подыскивать каузальную историю, которая подскажет, как обосновать результаты его изысканий. Когда он описывает результаты для научной публикации, зачастую он представляет все в обратном порядке. Он пишет, что начал с каузальной теории, затем стал искать наиболее достоверный путь ее трансформации в формальную гипотезу и нашел ей подтверждение в данных²¹. Это все лженаука. В естественных науках нет необходимости в том, чтобы логика оправдания соответствовала логике открытия. Как только гипотеза сформулирована в окончательной форме, ее генезис уже не имеет значения. Значение имеют последствия гипотезы.

¹⁸ Так, в одном продольном исследовании отношений между практиками материнства и доходами детей только 35 из 552 корреляций были статистически значимыми на уровне $p < 0,05$ (что означает, что был 1 шанс из 20, что они получены случайно); факт, очевидный только тем, кто прочел приложение к книге. При переиздании эта часть была исключена.

¹⁹ Теория может предполагать, что плохая погода депрессивно действует на трейдеров, заставляя их продавать. Ученые, однако, сообщают об обратном эффекте, если под плохой погодой понимается стопроцентная облачность. При изменении определения плохой погоды на 80%-ю облачность знак корреляции волшебным образом меняется на противоположный.

²⁰ У меня было два случая. Первый раз приехав в Нью-Йорк, я купил билеты на два бродвейских шоу. Одно было выстроено вокруг музыки Фэтс Уоллера, а другое — Дюка Эллингтона. Билетов оставалось мало, и мне пришлось брать то, что есть; в обоих случаях это был ряд Н, место 130. Это было в высшей степени странно, но другое совпадение кажется еще более значительным. Это два других случая, которые со мной однажды произошли: один заключался в том, что я был приглашен на званный ужин и забыл об этом; другой — в том, что я был приглашен на ужин, а потом за полчаса до начала хозяин дома позвонил мне и сообщил, что все отменяется из-за его болезни. Совпадение, заставившее меня во второй раз подумать, что кто-то следит за мной, заключалось в том, что это был один и тот же званный ужин.

²¹ Таким образом, есть три проблемы, связанные с использованием корреляции как пути к достижению каузальности. Во-первых, корреляции могут возникать случайно и не иметь причинно-следственной интерпретации. Во-вторых, корреляция может иметь косвенную причинно-следственную интерпретацию, если два коррелирующих явления — общие следствия третьего фактора. В-третьих, направление каузальности может быть неоднозначным.

зы, а не истоки. Так происходит потому, что гипотеза может быть проверена на бесконечном числе наблюдений, далеко выходящих за рамки тех, что первоначально подсказали ученому саму эту гипотезу. В социальных (и в гуманитарных) науках для большинства объяснений используется конечный ряд данных. Поскольку процедура сбора данных зачастую не стандартизирована, ученые могут быть лишены возможности проверять гипотезы, используя новые данные²². И даже если процедуры действительно стандартные, данные могут не отразить меняющуюся реальность. Невозможно, например, объяснить модели потребления, не принимая в расчет новые продукты и изменение цен на старые.

Без сомнения, практики такого рода имеют место. Я не знаю, насколько широко они распространены, но они встречаются довольно часто, чтобы обеспокоить представителей социальных наук. Главная причина этой проблемы, возможно, — наше неадекватное представление о каузальности со многими ее факторами. Если бы у нас было глубокое понимание того, как при производстве следствия могут взаимодействовать несколько причин, нам не пришлось бы полагаться на механическую процедуру добавления члена, характеризующего взаимодействие, когда не работают аддитивные модели. Но из-за слабости понимания мы в самом деле не знаем, что искать, и тогда возня с моделями кажется единственной альтернативой — по крайней мере пока мы сохраняем амбициозную цель дать объяснения, подобные законам. Учитывая опасности такой возни, нам, возможно, следует умерить амбиции.

Библиографические примечания

Многие идеи, излагаемые в этой главе, взяты из главы 1 моей книги «Алхимия разума» (*Alchemies of the Mind*. Cambridge University Press, 1999). Там я также ссылаюсь на работы Раймона Будона (Raymond Boudon), Нэнси Картрайт (Nancy Cartwright) и Поля Вена (Paul Veyne), защищающие сходные положения. Недавно эти проблемы поднимались в книге Питера Хедстрема «Вскрытие социального» (*Hedström P. Dissecting the Social*. Cambridge University Press, 2005). Полезные размышления

²² Можно обойти эту проблему или по крайней мере смягчить ее, практикуя самодисциплину. Если есть большой набор данных, можно сосредоточиться на репрезентативном примере и игнорировать все остальные. Как только сделано все, чтобы объяснить подгруппу наблюдений, можно распространить объяснение на весь набор данных и посмотреть, работает ли оно. Если да, то легче избежать подделок. Кроме того, сохранению научной честности способствовал бы отказ научных журналов принимать статьи для публикации, если проверяемые гипотезы вместе с методами их проверки не были переданы редактору, например, за два года до публикации. Такую практику применяет исследователь проблем алкоголя Кэтил Бруун (Ketil Bruun).

о психологических механизмах есть в работах Ф.Хейдера «Психология межличностных отношений» (Heider F. The Psychology of Interpersonal Relations. Hillsdale, NJ: Lawrence Erlbaum, 1958) и Р.Абельсона «Статистика как научно-теоретический аргумент» (Abelson R. Statistics as Principled Argument. Hillsdale, NJ: Lawrence Erlbaum, 1995). В последней содержатся проницательные и остроумные замечания о подводных камнях статистического анализа. Стандартное краткое введение в идею о том, что наука дает объяснения через общие законы, можно найти у Карла Гемпеля в «Философии естествознания» (Hempel C. Philosophy of Natural Science. Englewood Cliffs, NJ: Prentice-Hall, 1966). Принцип методологического индивидуализма подробно изложен в части 4 «Материалов по философии социальных наук» под редакцией М. Бродбека (Brodbeck M. (ed.). Readings in the Philosophy of the Social Sciences. London: Macmillan, 1969) и в части 6 «Материалов по философии социальных наук» под редакцией М. Мартина и Л. Макинтайра (Martin M., McIntyre L. (eds). Readings in the Philosophy of Social Science. Cambridge, MA: MIT Press, 1994). См. также К.Эрроу «Методологический индивидуализм и социальное знание» (Arrow K. Methodological individualism and social knowledge // American Economic Review: Papers and Proceedings. 1994. No. 84. P.1–9). Я писал о поговорах более систематично в работе «Наука и мудрость: роль поговорок в познании человека и общества» («Science et sagesse: Le rôle des proverbes dans la connaissance de l'homme et de la société») и в книге «Агент и его мотивы: сборник Раймона Будона» под редакцией Ж. Бешлера (Baechler J. (ed.). L'acteur et ses raisons: Mélanges Raymond Boudon. Paris: Presses Universitaires de France, 2000). Идея психологии тирании взята из Дж. Ремера «Рационализируя революционную идеологию» (Roemer J. Rationalizing revolutionary ideology // Econometrica. 1985. No. 53. P. 85–108). Эксперимент, в котором участникам были обещаны музыкальные записи, описывается в Дж. Брем и др. «Привлекательность альтернативы с отсутствием выбора» (Brehm J. et al. The attractiveness of an eliminated choice alternative // Journal of Experimental Social Psychology. 1966. No. 2. P. 301–313). Общее введение в теорию реактивного сопротивления есть в работе Р. Уиклунд «Свобода и реактивное сопротивление» (Wicklund R. Freedom and Reactance. New York: Wiley, 1974). Скептические замечания по поводу взаимодействия, которое вызывает обратный эффект, можно найти в Р.Хейсти, Р.Доус «Рациональный выбор в неопределенном мире», глава 3 (Hastie R., Dawes R. Rational Choice in an Uncertain World. Thousand Oaks, CA: Sage, 2001). Приведенная в примечании история о 6% значимых корреляций изложена в «Парадоксе родительского влияния» Р.Р.Маккрай и П.Т.Косты (McCrae R.R., Costa P.T. The paradox of parental influence // Perris C., Arrindell W.A., Eissmann M. (eds). Parenting and Psychopathology. New York: Wiley. P.113–14). Еще один пример из примечаний — влияние плохой погоды на биржевых трейдеров — взят из статьи П.Кеннеди «О нет! Я поставил неверный знак! Что же мне делать?» (Kennedy P. Oh no! I got the wrong sign! What should I do? // Journal of Economic Education. 2005. No. 36. P. 77–92), в которой также содержатся полезные замечания более общего характера о затратах на анализ данных (и о его преимуществах).

III. Интерпретация

ИНТЕРПРЕТАЦИЯ И ОБЪЯСНЕНИЕ

Во многих работах по гуманитарным наукам основное внимание уделяется интерпретации, а не объяснению. В немецкой традиции существует противопоставление гуманитарных наук (*Geisteswissenschaften*) естественным (*Naturwissenschaften*). Утверждается, что для первых соответствующая процедура — интерпретация, или понимание (*Verstehen*), а вторые говорят на языке объяснения (*Erklären*). Макс Вебер, например, писал, что естественные науки не ставят своей целью понимания поведения клеток.

В таком случае мы можем спросить: на чем основываются общественные науки — на понимании или на объяснении? Я полагаю, это неправильная постановка вопроса. По моему мнению, интерпретировать — значит объяснять. Интерпретация — это не что иное, как особый случай гипотетико-дедуктивного метода (глава I). Ученые-гуманитарии, например, не могут использовать эмпатию как удобный короткий путь к интерпретации поведения, поскольку она допускает разность толкований. Выбирая из противоречащих друг другу интерпретаций, они должны сопоставить эти интерпретативные догадки, или гипотезы (а именно этим они и являются), с *опытом*. Как я утверждал в главе I, опыт включает не только факты, которые мы пытаемся понять, но также *новые факты*, которые в противном случае нам не пришлось бы в голову исследовать¹.

Интерпретация направлена на деятельность человека и ее плоды, такие, например, как произведения искусства. В главе XIV я займусь проблемой интерпретации литературных произведений, а именно тех из них, в которых нам необходимо понимать действия персонажей, а также авторский выбор. При попытке понять другие литературные произведения, а также бессловесные искусства, такие как живопись, скульптура

¹ В экспериментальных науках понятие «новые факты» может означать в буквальном смысле новые факты, как например, когда человека или крысу подвергают воздействию, которому они не подвергаются в естественных условиях. В гуманитарных и неэкспериментальных общественных науках определение «новый» употребляется скорее в эпистемологическом смысле «ранее неизвестный», нежели в онтологическом смысле «ранее не существовавший».

или инструментальная музыка, подобные двухзвенные проблемы не возникают. Однако и в этих видах искусства к авторскому выбору применим анализ, который я собираюсь предложить для авторских решений. Художники делают выбор в соответствии с некоторым критерием предпочтительности (*betterness*), который ни мы, ни они, вероятно, не сможем эксплицитно сформулировать, но который раскрывается в практике, когда они выбирают тот или иной набросок, эскиз или запись и отбраковывают другие. И все же отношения между критерием предпочтительности и человеческой психологией в бессловесных искусствах сложнее, чем в (классической) литературе. За них я браться не буду.

РАЦИОНАЛЬНОСТЬ И ИНТЕЛЛИГИБЕЛЬНОСТЬ

Остальная часть этой главы будет посвящена интерпретации действия. При этом мы должны объяснить действие с точки зрения предшествующих убеждений и желаний (мотиваций) агента. Более того, мы должны объяснить сами эти ментальные состояния так, чтобы они обрели осмысленность, полностью поместив их в рамки комплекса *желание-убеждение* (*desire-belief complex*). Изолированное желание или убеждение, которое не солидаризируется с другими ментальными состояниями, — это просто голый факт, который позволит нам объяснить поведение, но не понять его.

Парадигматический способ объяснения действия — продемонстрировать, что оно было совершено потому, что *рационально* (глава XI). Чтобы это сделать, недостаточно показать, что действие имело для агента благоприятные последствия; оно должно быть понято как оптимальное с его точки зрения. Известно, например, что если люди придают большую ценность последствиям их настоящего поведения, то есть имеют низкую степень скидки на время (*time discounting*) (глава VI), их жизнь складывается лучше. Кроме того, вполне вероятно, что подобные предпочтения может формировать высокий уровень образования. Однако эти две предпосылки недостаточны, чтобы считаться объяснением того, почему люди решают получить образование, в категориях рационального выбора. Чтобы дать объяснение, нужно показать, что люди обладают необходимым убеждением, что образование накладывает отпечаток на способность откладывать удовлетворение и что у них есть субъективная заинтересованность в приобретении этой способности². Попытка объяснить выбор его благотворными последствиями является формой

² В главе XI я утверждаю, что идея заинтересованности в заинтересованности в долгосрочных последствиях является концептуально противоречивой. Но это отдельный разговор.

функционализма рационального выбора (соединяющего в себе два подхода, против которых я предостерегал во Введении), никаким образом не проясняющего смысл поведения.

Если поведение рационально, оно в силу самого этого факта является интеллигибельным. Иррациональное поведение, однако, также может быть интеллигибельным. Я покажу различия трех разновидностей интеллигибельного, но иррационального поведения и сопоставлю их с некоторыми случаями неинтеллигибельного поведения.

Первая возникает, когда механизм принятия решения (рис. XI.1) тем или иным образом *усекается* (truncated). В силу своей особой настойчивости сильная эмоция может помешать агенту оглядеться (то есть собрать информацию), прежде чем действовать. Вместо того чтобы избрать выжидательную стратегию, подобно древнеримскому генералу Фабию Максиму Кунктатору (букв.: «медлителю»), агент бросается действовать, не остановившись, чтобы обдумать последствия. Другая форма усечения возникает из-за безволия, традиционно понимаемого как действие вопреки трезвому расчету (глава VI). У человека, решившего бросить курить и принимающего предложение выкурить сигарету, есть причина для действия, а именно желание курить. Но чтобы действие было рациональным, оно должно быть оптимальным с точки зрения всего комплекса причин, а не только одной из них. У меня еще будет случай поставить под вопрос такое понимание безволия.

Вторая разновидность возникает при «коротком замыкании» (short-circuiting) механизма принятия решения, когда на формирование убеждений оказывают давление желания агента. Например, самообольщение, принимающее желаемое за действительное, иррационально, но полностью интеллигибельно. Более тонкая форма мотивированного формирования представлений возникает, когда агент перестает собирать информацию, а уже накопленной информации достаточно, чтобы поддержать убеждение, которое он хотел бы считать истинным³. Эти формы мотивированного формирования убеждений (motivated belief formation) в свою очередь являются оптимизирующими процессами. Они максимизируют удовольствие, которое агент извлекает из своих представлений о мире, а не удовольствие, которое он ожидает от своих встреч с миром.

Третья разновидность — это то, что мы можем назвать «соединение проводов» (wire-crossing) в механизме принятия решений. Мы можем легко понять, почему разум может включиться в редукцию когнитивного диссонанса (одной из разновидностей которой является принятие же-

³ Во времена бóльшого статистического простодушия метод «бросить, когда ты уже много добился» явно практиковал в своих экспериментах Грегор Мендель, открывший законы наследственности.

лаемого за действительное). По почему он должен стремиться к *производству* диссонанса? Примером может послужить идея, упоминавшаяся в главе II, — страх заставляет легко поверить в то, чего бояться. Почему страх плохого конца заставляет нас считать его более вероятным, хотя это не явствует из имеющихся у нас свидетельств? Если представление не поддерживается ни свидетельствами, ни нашими желаниями, то почему мы его принимаем? Ясно, что в данном случае ничего не оптимизируется. В определенном смысле такое поведение труднее понять, чем усечение или «короткое замыкание», поскольку для агента *в нем ничего нет*. Нет ни частичной, ни краткосрочной цели, которую такое поведение удовлетворяло бы. И тем не менее оно интеллигибельно (в моем понимании этой идеи), потому что возникает из системы *убеждений-желаний* агента.

К ускользающим от интерпретации относятся также действия, вызванные навязчивыми влечениями и одержимостями, фобиями, самоистязаниями, анорексиями и тому подобным. Несомненно, такое поведение имеет определенный эффект, объясняющий, почему оно возникает, — оно ослабляет тревогу, которую испытывает агент, если не совершает подобных действий. И все же пятьдесят раз в день мыть руки или пройти пятьдесят этажей, чтобы не садиться в лифт, — это не прием транквилизаторов. Прием валиума может быть таким же рациональным и интеллигибельным, как прием аспирина, но навязчивые поступки и фобии неинтеллигибельны, потому что не являются частью взаимосвязанной *системы* убеждений и желаний. Или возьмем пример из Джона Ролза: нам трудно будет понять поведение человека, занятого подсчетом травинки, если оно не связано с какой-то целью, например пари.

Самообольщение, принимающее желаемое за действительное, интеллигибельно точно так же, как его противоположность — принятие нежелаемого за действительное (*counterwishful thinking*); вера психически нездорового индивида в то, что стоматолог из соседнего здания направляет на него рентген, чтобы разрушить его мозг, — нет. В политике же, наоборот, параноидальные убеждения интеллигибельны, поскольку укоренены в желаниях агента. Ярый антисемит заинтересован в том, чтобы поддерживать абсурдные представления о всемогуществе и злой природе евреев (см. главу VII). Дело не в том, что он хочет, чтобы у них *были* эти черты, а в том, что у него есть мотивация, чтобы в это *верить*, потому что такая вера может рационализировать его желание их уничтожить. Даже противоречивые представления могут быть интеллигибельными. Антисемит в одном случае может характеризовать евреев как паразитов, а в другом заявлять об их всемогуществе. Один и тот же человек, говорящий, что евреи всегда пытаются пробиться туда, где их не желают видеть, также полагает, что евреи склонны к клановости, всегда держатся

друг за друга. Один и тот же мусульманин может заявлять, что за атаками на Всемирный торговый центр 11 сентября 2001 года стояла израильская разведывательная служба «Моссад» и *в то же время* гордиться, этим событием.

ПОНИМАНИЕ ГРАЖДАНСКИХ ВОЙН

Позвольте привести два пространных примера, в которых идеи интеллигибельных убеждений и желаний обрастают плотью. Оба примера взяты из исследований, посвященных гражданским войнам, прошлым и настоящим. Далее я буду использовать эти и некоторые другие исследования, обратившись к ключевому вопросу герменевтики — как мы можем вменять или создавать мотивации и убеждения.

Рассмотрим сначала веру в предопределение, которая была основным предметом разногласий между католиками и кальвинистами во время Религиозных войн. У истоков проблемы стоит сильная религиозная тревога, которую испытывали многие верующие в эпоху, предшествовавшую Реформации, из-за неопределенности перспективы личного спасения. Можно ли быть уверенным в том, что поступков человека будет достаточно для его достижения? Вспоминая о годах своей молодости, Кальвин писал в 1539 году, что даже выполняя требования Церкви признаваться в своих грехах и стирать их из Божественной памяти посредством добрых дел и покаяния, «я был далек от уверенности, и совесть моя не была спокойна. Ибо всякий раз, как я погружался в себя или в сердце своем возносился к Тебе, меня поражал такой ужас, что ни очищение, ни отпущение грехов не могли меня исцелить».

От тревоги его освободил переход от концепции бога, имманентного миру, грозного и гнетущего присутствия, к концепции бога, абсолютно трансцендентного. По сути эта идея была связана с доктриной двойного предопределения: поскольку бог уже давно сделал выбор, кто спасется, а кто будет проклят, для своего спасения *ничего нельзя сделать*, а следовательно, не стоит беспокоиться о том, что сделанного недостаточно. Ключевой для толкования вопрос касается связи между верой в предопределение и избавлением от тревоги. Априори такое действие доктрины кажется неинтеллигибельным. Согласно учению Кальвина, избранные были небольшим меньшинством, количество которых варьирует (в его высказываниях) от одного на сотню до одного из пятерых. Что, как не вера в то, что человек скорее всего окажется среди проклятых и нельзя ничего сделать, чтобы избежать вечных мучений в аду, может вызвать большую тревогу? Не попадает ли человек при обращении из католицизма в кальвинизм в буквальном смысле слова из огня в полымя?

Ответ, вероятно, нужно искать в строчках, которые впервые набросал Макс Вебер. Исходя из веры в предопределение, кальвинисты не могли утверждать, что рациональные и систематические усилия принесут им спасение, но они могли утверждать (и утверждали), что она даст им субъективную *уверенность* в спасении. Сам Кальвин писал, что «призвание избранных само как демонстрация и свидетельство их избранности». И обращение в кальвинизм, как кажется, действительно эффективно устранило неуверенность в спасении. Я вернусь к этой форме магического мышления в главе VII. Здесь я только хочу подчеркнуть, как двойной механизм — самообман и магическое мышление — делает интеллигибельной веру в предопределение.

Далее давайте рассмотрим интеллигибельность мотиваций. Почему юные палестинцы готовы отдать жизнь, выполняя самоубийственные задания? Их основной мотив — отвоевать или защитить свою родину — понять нетрудно⁴. Их дело может быть таким же привлекательным, как в свое время защита демократии в борьбе против Гитлера. Что может поставить в тупик, так это *сила* этой мотивации. Чтобы сделать ее интеллигибельной, могут потребоваться дополнительные каузальные факторы. Я рассмотрю подюжины таких факторов и сделаю заключение в пользу одного из них.

До 11 сентября существовало распространенное мнение, что типичный террорист-самоубийца с Ближнего Востока — это неженатый молодой безработный, возможно, испытывающий недостаток секса, для которого религиозное движение может заполнить пустоту, которая при иных обстоятельствах была бы заполнена семьей и работой. После атаки на Всемирный торговый центр за одну ночь эксперты по терроризму решили, что «книгу нужно переписать». Однако колеблющийся, но высокий даже до этого процент использования женщин-террористок должен был побудить ученых усомниться в данном стереотипе. Во время второй интифады участие террористок-смертниц, некоторые из которых имели детей или хорошее образование, было еще более поразительным фактом.

Часто упоминаемые факторы бедности и неграмотности тоже, как представляется, имеют ограниченную каузальную эффективность, по крайней мере в качестве характерных черт конкретных террористов-смертников. Доход и уровень образования среди палестинских террори-

⁴ Утверждая, что это их главный мотив, я не отрицаю, что могут быть другие, такие как желание посмертной славы, материальные выгоды, которые получит семья террориста-самоубийцы, месть Израилю за погибшего друга или родственника либо социальное давление, принуждающее вызваться добровольцем. Как я отмечаю во введении ко второй части, у меня скептическое отношение к мотивационной силе религиозных «льгот» в форме привилегированного «доступа» в рай.

стов на самом деле выше, чем у обычного населения. Объяснения посредством указания на бедность также неудовлетворительны, оставляя без ответа вопрос, как бедность порождает подобную мотивацию. Есть распространенное мнение, что выгоды от самоубийства нужно сравнивать с ценой самоубийства — жизнью. Если жизнь ценится невысоко, цена самоубийства уменьшается. Согласно такому подходу, жизнь в нищете и убожестве так мало ценна для индивида, что цена самоубийства становится несущественной. Я скептически отношусь к этому аргументу, поскольку считаю, что для бедняка его жизнь не менее ценна, чем для любого другого человека. То, что люди приспособливают свои амбиции к обстоятельствам так, чтобы поддерживать более или менее постоянный уровень удовлетворенности, — это почти точно установленный факт.

Более достоверный фактор, чем абсолютная депривация, — *депривация относительная*, то есть разрыв между ожиданиями и реальностью, с которым на собственном опыте сталкиваются многие образованные палестинцы, лишённые в настоящей момент перспективы найти приличную работу. Тот же самый эффект может иметь нисходящая социальная мобильность. Но все же наиболее релевантными чертами могут оказаться постоянно испытываемое чувство *неполноценности* и *рессентимент*⁵. Первая из этих эмоций основана на *сравнении* себя с другими людьми, вторая — на *взаимодействии* с ними. По большому счету, эмоции, основанные на взаимодействии, сильнее эмоций, основанных на сравнении. Многие из пишущих о палестинских террористах-самоубийцах ставят акцент на остром чувстве обиды, вызываемом ежедневными унижениями, возникающими при взаимодействии с израильскими военными. Помимо унижительных проверок, которым подвергаются палестинцы, есть также осознание того, что многие израильтяне считают всех арабов ленивыми, трусливыми и жестокими, как двадцать лет назад сказал мне один иерусалимский водитель.

Если это объяснение правильное, сильный *рессентимент в отношении тех*, кто оккупировал их родную землю, позволяет понять готовность палестинских террористов-самоубийц умереть. Желание бороться с Израилем черпает силу из глубоко заложенного мотивационного комплекса. Есть, однако, и альтернативный взгляд. Покровители обычно держат палестинских террористов, что называется, на коротком поводке, будучи готовыми в любой момент усилить давление, если первичная мотивация не сработает в момент осуществления акции. Один из потенциальных террористов-смертников, которого задержали и разоружили, так как внешне он был слишком нервным, сказал, что за три

⁵ От франц. *ressentiment* — зависть, обида, озлобленность. Качество рабской морали в философии Фридриха Ницше. — Примеч. ред.

дня до этого его заперли в комнате с муллой, который говорил о рае и кормил его «специальным супом, делавшим его сильнее». Ментальное состояние, в действительности провоцирующее взрыв бомбы, может, следовательно, оказаться эфемерным или быть чем-то вроде артефакта, а не постоянной чертой человека. Хотя такие выражения, как «гипноз» или «промывка мозгов», могут показаться слишком сильными, есть свидетельства того, что за минуты до своей гибели многие террористы были в состоянии транса. Когда, как в этом случае, интенция изолирована от системы желаний и убеждений, интерпретация невозможна. Поведение покровителей и, шире, организаторов террористических актов, конечно же, интерпретации поддается.

ГЕРМЕНЕВТИЧЕСКАЯ ДИЛЕММА

Разумеется, можно сколько угодно долго требовать объяснения поведению в категориях предшествующих ментальных состояний, желаний и убеждений, которые его породили. Но как мы установим эти исходные причины? Во избежание порочного круга мы не можем использовать как доказательство само поведение. Мы должны искать другие свидетельства, такие как высказывания агента о его мотивации, соответствие между его высказываниями и невербальным поведением, мотивы, приписываемые ему другими людьми, и соответствие между их заключением и его собственным невербальным поведением. Но как исключить возможность, что эти вербальные и невербальные формы поведения не были использованы намеренно, чтобы заставить аудиторию поверить в эти мотивы? Признания и приписывание мотиваций могут сами по себе быть мотивированными. Это центральный вопрос для коллективного принятия решений. Как я утверждаю в главе XXV, все методы консолидации индивидуальных предпочтений в социальное решение создают для участников процесса стимулы к представлению их предпочтений в ложном свете.

Рассмотрим в качестве примера мотивы лидеров гражданских войн и их последователей. Партии признавали, или же их оппоненты приписывали им один из трех мотивов: *религия, власть и деньги*. Тех, кто заявляет о религиозных мотивах, часто обвиняют в использовании религии для маскировки реальных интересов, будь они политическими или материальными. Во время французских религиозных войн (1562–1598) противоборствующие стороны постоянно обвиняли друг друга в использовании религии в политических или корыстных целях. Для такого обвинения были некоторые основания. Генрих Наваррский (будущий Генрих IV) за свою жизнь менял религию шесть раз, его последнее обращение в 1593 году давало серьезные основания подозревать его в оп-

портунизме. Его отец герцог Антуан де Бурбон дал понять, что его веру может купить тот, кто заплатит самую высокую цену. Он сопровождал королеву-регентшу к мессе, а свою жену-протестантку — к причастию. На смертном одре он искал утешения у обеих религий. Ведущий реформатор Оде де Колиньи кардинал де Шатийон, женился после перехода в протестантизм, но сохранил кардинальский титул и доходы от епископата. Еще один прелат, Антуан Карачиоло, епископ Труа, также хотел совместить протестантское пастырство и доход от епископата. Вождь католиков герцог де Гиз был готов заключить союз с кальвинистами против короля Генриха III.

В современном мире религия также используется как предлог для политики, а политика — как предлог для богатства. Цели чеченских сепаратистов и некоторых палестинских организаций, таких как ФАТХ, первоначально были чисто политическими. Когда они приобрели религиозную окраску, это было сделано в основном для привлечения большего числа последователей. В Палестине из-за соперничества с безусловно религиозной ХАМАС это стало необходимым условием выживания организации. На Филиппинах террористическая группа «Абу Сайяф» использовала требование создать независимое исламское государство как предлог для похищений людей и требований большого выкупа. В Колумбии уже трудно понять, революционные Вооруженные силы Колумбии (FARC) остались верны своей первоначальной цели борьбы с социальной несправедливостью или же превратились в мафию. Во всех этих случаях, как и в случае французских религиозных войн, приписывание мотивов часто довольно неопределенно. В частности, трудно узнать, согласуются ли между собой мотивы вождей и их последователей.

Существует множество причин, по которым у людей может возникнуть желание представить в ложном свете мотивы своих оппонентов и свои собственные. С одной стороны, каждое общество имеет нормативную иерархию мотиваций (глава IV), которая заставляет делать вид, что человеком движут благородные, а не низменные интересы, последние же приписывать оппоненту. Во французских религиозных войнах, так же как в гражданской войне в Англии, каждая сторона хотела показать, что ею движут истинно религиозные мотивы, тогда как противная сторона ведома лишь жадной власти. С другой стороны, если человек может заставить других поверить своим заявлениям о той или иной мотивации, возможно, ему будет легче достичь своей цели. Поскольку образ террориста внушает больший страх, чем образ обычного преступника, похитители людей могут повысить свои шансы на получение выкупа, размахивая знаменем какого-либо движения. В Колумбии многие похищения совершают обычные преступники, которые, чтобы напугать семьи жертв, заявляют о принадлежности к партизанским груп-

пам. Похищения становятся страшнее, если считается, что террористы готовы пойти на крайние меры в случае ошибки и не склонны вести переговоры о сроках и деньгах. Если они не могут получить то, чего требуют, то по крайней мере могут «сделать заявление», убив свои жертвы.

Проблема своекорыстной предвзятости в высказываниях социальных агентов о своих намерениях является серьезной, но все-таки преодолимой. Можно выйти из положения простым путем, рассмотрев *объективные интересы* агента и заключив, что в отсутствие убедительных доказательств в пользу обратного они совпадают с его субъективной мотивацией, независимо от того, что он о них говорит. И наоборот, можно идентифицировать *реальные последствия* его действия и заключить, что в отсутствие убедительных доказательств в пользу обратного они соответствуют исходному намерению. (К обсуждавшемуся выше выбору в пользу более высокого уровня образования применимы обе идеи.) Однако тот факт, что существует *две* подобные процедуры для переноса бремени доказательства, подсказывает, что ни одна из них неприемлема. И объективные интересы, и реальные последствия могут подсказывать полезные гипотезы, касающиеся субъективных мотивов, но ни одна из процедур не обладает презумпцией в свою пользу.

Историки и специалисты по социальным наукам разработали другие подходы к этой проблеме, которые (в особенности при совместном применении) позволяют делать надежные заключения. Один подход заключается в том, чтобы обойти заявления, сделанные на публике, и обратиться к заявлениям, которые с меньшей вероятностью могут быть мотивированы желанием ввести в заблуждение. Письма, дневники, переданные разговоры, черновики могут в данном случае послужить бесценным источником. Из их писем женам мы знаем, что некоторые делегаты Национального учредительного собрания Франции в 1789 году голосовали против двухпалатной системы и королевского права вето, потому что полагали, что в противном случае их жизнь может оказаться в опасности. На самих заседаниях они оправдывали свой поступок, апеллируя к общественным интересам. Пытаясь выяснить мотивации, стоявшие за резней в Варфоломеевскую ночь 1572 года, историки сочли полезным отложить предвзятые отчеты о событиях, данные их участниками, и полагаться на сообщения иностранных дипломатов, которые были заинтересованы в достоверном изложении событий. В Англии XIX столетия заявления, сделанные на смертном одре, не подпадали под действие правил, касавшихся свидетельств, основанных на слухах. Первая редакция того или иного документа может рассказать об убеждениях и мотивах его автора больше, чем опубликованная позднее работа. Так, полезно сравнить черновики «Гражданской войны во Франции» Карла Маркса или его письмо Вере Засулич с официальными версиями.

Помимо этого может существовать разительный контраст между тем, что агенты говорят на публике, и тем, что они говорят за закрытыми дверями. Хотя опубликованные дебаты Национального собрания 1789–1791 годов совершенно очаровательны, но совместное действие двух факторов делает их не слишком надежным источником сведений о ментальных состояниях. С одной стороны, обстановка публичности вынуждала делегатов прибегать только к аргументам, связанным с общественными интересами; голые групповые интересы были недопустимы; с другой — их честолюбие распалялось выступлением перед тысячей делегатов и тысячей слушателей на галереях. В обоих отношениях американский Конституционный конвент в большей степени поощрял искренность. Поскольку число делегатов было небольшим (55; для сравнения: в Париже их было 1200) и Конвент проходил в обстановке секретности, могли произойти (и происходили) сделки, основанные на корыстном интересе. В то же время, как много лет спустя писал Джеймс Мэдисон, «если бы члены конгресса сначала взяли на себя обязательства публично, впоследствии из соображений последовательности в убеждениях им пришлось бы придерживаться той же позиции, тогда как при тайном обсуждении каждый считал себя обязанным придерживаться своих убеждений только до тех пор, пока считал их истинными и уместными, и оставался открытым воздействию убеждения». Дебаты не могли охладить даже страх последующих разоблачений, поскольку предполагалось, что режим секретности будет сохраняться до бесконечности. Действительно, он был нарушен, когда много лет спустя были опубликованы заметки Мэдисона. Стратегические причины для введения в заблуждение отходят на второй план, если за искренность не приходится расплачиваться.

Социологи могут добиться искренности, создав искусственную завесу неведения (*veil of ignorance*). Предположим, ученый хочет изучить отношения между сексуальной ориентацией и некоторыми переменными, представляющими интерес. Возможно, будет сложно побудить людей давать правдивые ответы на вопросы о том, вступали ли они когда-либо в связь с людьми одного с ними пола, даже если заверить их, что ответы останутся анонимными. Чтобы справиться с этой проблемой, исследователь может проинструктировать опрашиваемых отвечать честно в случае, если у них был такой опыт; если же его не было, то бросить монетку, чтобы решить, отвечать «да» или «нет». Если они согласятся (а у них нет причины не пойти на это) и выборка будет большой, данные будут ничуть не менее объективными, чем если бы все отвечали правдиво.

Еще один метод проверить, не противоречит ли невербальное поведение агентов декларируемым мотивам, — спросить, готовы ли они ответить за свои слова? Когда в 2003 году Администрация Буша с уве-

ренностью ссылалась на наличие у Саддама Хусейна оружия массового поражения как на причину вторжения в Ирак, приняла ли она все необходимые меры для того, чтобы защитить американских солдат от этой угрозы? Некоторые поведенческие модели могут приоткрыть истинные мотивы похищения людей. В 1996 году в Коста-Рике похитители (главным образом, бывшие контрасты из Никарагуа) потребовали выкуп 1 миллион долларов вместе с гарантиями занятости для рабочих, снижением цен на продовольствие, увеличением минимальной заработной платы и освобождением их товарищей из тюрем. Когда им предложили 200 тысяч долларов, они удовлетворились этой суммой и не стали настаивать на политических требованиях. Факт убедил власти в том, что их заявления в духе Робин Гуда были уловкой и что их целью всегда были только деньги. Или возьмем поведение французских аристократов, эмигрировавших во время Французской революции в Лондон. В этом очаге слухов о грядущей реставрации монархии и постоянного соревнования по роялизму было жизненно необходимо выразить свою готовность служить делу контрреволюции. Словесных заверений было недостаточно. На всякого, кто снимал квартиру более чем на месяц, смотрели с подозрением. Предпочтительнее было арендовать понедельно, дабы не оставлять сомнений в готовности в любую минуту вернуться во Францию по призыву роялистов.

Не только современники, но и историки регулярно используют эти поведенческие модели как индикатор искренности публичных заявлений о лояльности. Например, к концу Второй мировой войны в оккупированной Франции присутствовала значительная доля скептицизма относительно победы Германии. Выражать такую позицию было небезопасно, но она отражалась в поведении. Процент учащихся средней школы, выбравших в качестве иностранного языка немецкий (или тех, чьи родители выбрали его для них), удвоился с 1939 по 1942 год, а затем стремительно упал. Многие издатели, с готовностью купившие права на перевод немецких книг, так и не воспользовались ими.

Судьи и присяжные также нередко прибегают к этому методу. Иногда они спрашивают: «Был ли у обвиняемого мотив для совершения X?» — в надежде, что ответ поможет им решить, действительно ли обвиняемый совершил X. В этом случае наличие мотива является объективным свидетельством того, было ли обвиняемому выгодно совершение X. В других, более характерных случаях установлено, что обвиняемый совершил X, после чего встает вопрос, каков был его мотив. Чтобы установить, было убийство совершено в состоянии аффекта или было преднамеренным действием, судьи и присяжные смотрят не только на объективные выгоды, но стараются определить субъективное состояние обвиняемого. Если обвиняемый утверждает, что действовал под влиянием гнева

или ревности, а позднее выясняется, что он заблаговременно приобрел орудие убийства и не сразу пустил его в ход⁶, доверие к нему снижается.

В отдельности эти подходы могут не сработать. Депутат может не захотеть признаться своей жене, что он боялся за свою жизнь, или может утверждать, что боялся, чтобы на самом деле скрыть мотив, недостойный уважения (например, взятку). В Индии XIX столетия заявления, сделанные на смертном одре, считались ненадежными, потому что люди иногда использовали этот момент для того, чтобы поквитаться со своими врагами. В примере с французскими эмигрантами и те, кто верил в реставрацию, и те, кто в нее не верил, имели основания, чтобы арендовать жилье по неделям: первые для того чтобы упростить свое возвращение во Францию, когда наступит время, вторые — во избежание обвинений в пораженчестве. Однако у способности людей плести паутину обмана, не раскрывая истинных мотивов, есть свои пределы. Лицемерие, как сказал Сомерсет Моэм, — это профессия, которая занимает все время. Даже Тартюф в конце поскользнулся. Чтобы доказать искренность религиозной веры Генриха IV, его биограф не только ссылается на положительные свидетельства «множества эпизодов, когда его религиозный дух проявлял себя без декларации своих намерений», но также утверждает: «Если здесь и было какое-то лицемерие, то оно бы проявилось по тому или иному подходящему поводу». В том же духе высказывался и Монтень:

Те, кто, вопреки моему мнению о себе, имеют обыкновение утверждать, будто то, что я в своей натуре называю искренностью, простотою и непосредственностью, на самом деле ловкость и тонкая хитрость и что мне свойственны скорее благоразумие, чем доброта, скорее притворство, чем естественность, скорее умение удачно рассчитывать, чем удачливость, не столько бесчестят меня, сколько оказывают мне честь. Но они, разумеется, считают меня чересчур уж изворотливым, и того, кто понаблюдал бы за мной вблизи, я охотно признаю победителем, если он не вынужден будет признать, что вся их мудрость не может предложить ни одного правила, которое научило бы воссоздавать такую же естественную походку и сохранять такую же непринужденность и беспечный облик — всегда одинаковый и невозмутимый — на дорогах столь разнообразных и извилистых; если он не признает также, что все их старания и уловки не сумеют научить их тому же⁷.

⁶ В решении британского суда по делу Макферсона от 1957 года лорд Годдard задал риторический вопрос: «Как можно утверждать, что лицо, подавшее апелляцию, действовало в порыве страсти, если он сделал не один, а четыре выстрела, после каждого из которых требовалось четыре раза подряд перезаряжать ружье, извлекая патрон?»

⁷ Монтень М. Опыты: в 3 кн. Кн. III. Гл. 1: О полезном и честном. М.: Голос, 1992. — Примеч. пер.

Хотя выгода от представления в ложном свете может быть значительной, его цена может оказаться слишком высокой. До некоторой степени сознательная декларация своих мотивов имеет пределы. Потому что каждый мотив оказывается включенным в широкую сеть других мотивов и представлений, и для того чтобы поддерживать лицемерие, необходимым станет согласование такого количества вещей, что это может оказать парализующее действие. Довольно одной фальшивой ноты, чтобы обрушилась вся конструкция. Есть множество пословиц, свидетельствующих о том, что разрушение доверия необратимо. И хотя утверждать вслед за народной мудростью: «Раз солгал, а навек лгуном стал», — можно только с серьезными оговорками (глава X), это непрофессиональное убеждение в действительности очень распространено и в какой-то степени является сдерживающим средством против лжи. Помимо прочего это одна из причин, по которой можно согласиться с Декартом, сказавшим, что «самая большая уловка из всех — не пользоваться никакими уловками».

Библиографические примечания

Споры по поводу понимания (*Verstehen*) и объяснения (*Erklären*) излагаются в третьей части «Сборника по философии социальных наук» под ред. М. Мартина и Л. Макинтайра (Martin M., McIntyre L. (eds). *Readings in the Philosophy of Social Science*. Cambridge, MA: MIT Press, 1994). В этом издании глава, написанная Дагфином Фёллесдалом (Dagfinn Føllesdal), «Герменевтика и гипотетико-дедуктивный метод» защищает позицию, близкую моей собственной. Цитата из Вебера взята из его эссе «Интерпретативное понимание социального действия» (The interpretive understanding of social action // Brodbeck M. (ed.). *Readings in the Philosophy of the Social Sciences*. London: Macmillan, 1969, p. 33). Непоследовательность антисемитской позиции затрагивается в Дж. Телушкин «Еврейский юмор» (Telushkin J. *Jewish Humor*. New York: Morrow, 1992), который является источником других замечаний, сделанных о евреях и евреями о приписываемых им характерных чертах. Замечания об образовательном выборе являются косвенной полемикой с «Эндогенной детерминацией временных предпочтений» Г. Бекера и К. Маллигана (Becker G., Mulligan C. *The endogenous determination of time preferences* // *Quarterly Journal of Economics*. 1997. No. 112. P. 729–58). Свидетельская ценность предсмертных признаний обсуждается в Дж. Ф. Стифен «История английского уголовного права» (Стифен Д. *Уголовное право Англии в кратком начертании*. СПб., 1865). В работе Б. Н. Сасса «Преступления в состоянии аффекта» (Sass B. N. *Affektdelikte* // *Nervenarzt*. 1983. No. 54. S. 557–572) перечисляется тринадцать причин, почему заявление о том, что преступление совершено в состоянии аффекта, лишено достоверности. Выдающее-

ся интерпретативное обсуждение французских религиозных войн содержится у Дени Крузе в «Воины Божьи» (Crouzet D. Les guerriers de Dieu. Paris: Champ Vallon, 1990). Интерпретативный анализ мотиваций и убеждений террористов-смертников можно найти в эссе, написанных мною и С. Холмсом, Л. Риколфи в «Объясняя самоубийственные задания» (Holmes S., Ricolfi L., Elster J. Making Sense of Suicide Missions / D. Gambetta (ed.). Oxford University Press, 2005). Замечания касательно религиозных верований Генриха IV содержатся в Ж.-П. Бабелон «Генрих IV» (Babelon J.-P. Henri IV. Paris: Fayard, 1982. P. 554). Чрезмерный скептицизм по отношению к мотивациям человека обсуждается в Дж. Макки «Все ли люди лжецы?» (Mackie G. Are all men liars? // Elster J. (ed.). Deliberative Democracy, Cambridge University Press, 1998).

РАЗУМ

Ч А С Т Ь В Т О Р А Я

Эта книга построена вокруг принципа действия, основанного на модели убеждения-желания (belief-desire model). Чтобы понять, как действуют и взаимодействуют люди, мы должны сначала понять, как работает их разум. В основном это вопрос интроспекции и фолк-психологии, очищенной и откорректированной более систематическими исследованиями психологов и все чаще экономистов-бихевиористов. Модель играет важнейшую роль не только в объяснении поведения, но и в разграничении одобрения, вины или наказания. Вина обычно предполагает mens rea (лат. «виновная воля»), то есть намерения и убеждения. Объективное вменение (strict liability), то есть вина, признаваемая на основании одних только реальных последствий действия, встречается редко. На практике мы часто признаем людей виновными лишь на основании намерений, даже если никакие последствия не наступили. «Ведьмы, — заявлял Джон Донн, — иногда думают, что убивают, когда не убивают, и следовательно, так же виновны, как если бы убили». «Что касается ведьм, — писал Гоббс, — я полагаю, что их колдовство не есть реальная сила, и тем не менее я думаю, что их справедливо наказывают за их ложную уверенность, будто они способны причинять подобное зло, — уверенность, соединенную с намерением причинить это зло, будь они на то способны».

Хотя без модели убеждение-желание не обойтись, она очень хрупка. Методы, которыми мы пользуемся, приписывая людям разные ментальные состояния, не всегда дают стабильные результаты. Если необходимо измерить высоту здания, неважно, откуда начинать измерение — с крыши или от земли. При вычислении убеждений и желаний результат, напротив, может зависеть от не относящихся внешне к делу факторов. Рассмотрим, например, утверждение, что люди стремятся к максимизации предполагаемой выгоды (глава XI). Чтобы прояснить его, нам придется допустить, что у людей есть четкое и стабильное представление о ценности каждого возможного последствия поступка, а также о вероятности его наступления. Часто такое предположение оправданно, но иногда это не так.

Вначале разберем убеждения агента. Для выявления субъективных оценок вероятности, примеряемых индивидом к событию, как правило, применяется следующая стандартная процедура. Беря для начала число p , мы спрашиваем индивида, что он предпочтет — лотерею, в которой выиграет определенную сумму денег с вероятностью p , или лотерею, в которой получит денежный приз при наступлении некоего события¹.

¹ Мы должны предположить, что означенное — это такое событие, которое не затрагивает агента лично, как например, открытие жизни на других пла-

Эта книга построена вокруг принципа действия, основанного на модели убеждения-желания (belief-desire model). Чтобы понять, как действуют и взаимодействуют люди, мы должны сначала понять, как работает их разум. В основном это вопрос интроспекции и фолк-психологии, очищенной и откорректированной более систематическими исследованиями психологов и все чаще экономистов-бихевиористов. Модель играет важнейшую роль не только в объяснении поведения, но и в разграничении одобрения, вины или наказания. Вина обычно предполагает mens rea (лат. «виновная воля»), то есть намерения и убеждения. Объективное вменение (strict liability), то есть вина, признаваемая на основании одних только реальных последствий действия, встречается редко. На практике мы часто признаем людей виновными лишь на основании намерений, даже если никакие последствия не наступили. «Ведьмы, — заявлял Джон Донн, — иногда думают, что убивают, когда не убивают, и следовательно, так же виновны, как если бы убили». «Что касается ведьм, — писал Гоббс, — я полагаю, что их колдовство не есть реальная сила, и тем не менее я думаю, что их справедливо наказывают за их ложную уверенность, будто они способны причинять подобное зло, — уверенность, соединенную с намерением причинить это зло, будь они на то способны».

Хотя без модели убеждение-желание не обойтись, она очень хрупка. Методы, которыми мы пользуемся, приписывая людям разные ментальные состояния, не всегда дают стабильные результаты. Если необходимо измерить высоту здания, неважно, откуда начинать измерение — с крыши или от земли. При вычислении убеждений и желаний результат, напротив, может зависеть от не относящихся внешне к делу факторов. Рассмотрим, например, утверждение, что люди стремятся к максимизации предполагаемой выгоды (глава XI). Чтобы прояснить его, нам придется допустить, что у людей есть четкое и стабильное представление о ценности каждого возможного последствия поступка, а также о вероятности его наступления. Часто такое предположение оправданно, но иногда это не так.

Вначале разберем убеждения агента. Для выявления субъективных оценок вероятности, примеряемых индивидом к событию, как правило, применяется следующая стандартная процедура. Беря для начала число p , мы спрашиваем индивида, что он предпочтет — лотерею, в которой выиграет определенную сумму денег с вероятностью p , или лотерею, в которой получит денежный приз при наступлении некоего события¹.

¹ Мы должны предположить, что означенное — это такое событие, которое не затрагивает агента лично, как например, открытие жизни на других пла-

Если он выберет первое, мы предложим ему новый выбор — с более низкой вероятностью выигрыша; если он предпочтет второе, мы повысим вероятность. Действуя в таком ключе, мы в конце концов достигнем вероятности p^* , при которой подопытному будет все равно, выиграет он деньги в лотерею с вероятностью p^* или если произойдет определенное событие. В принципе, p^* должно быть независимым от изначального p , то есть найденная вероятность должна быть независимой от процедуры ее выявления. На практике это не так: более высокая p влечет более высокую p^* . Результаты показывают (по крайней мере в известных пределах), что *объекта поиска*, стабильного ментального состояния, которое выявила бы данная процедура².

Другие процедуры еще менее надежны. Ученые часто предполагают в агентах наличие субъективных оценок вероятности; исходя из того, что если те мало знают о ситуации, они склонны оценивать как одинаковую вероятность достижения любого положения вещей. В обосновании этой процедуры лежит принцип недостаточного основания (principle of insufficient reason): если у вас нет твердых оснований считать некоторое положение вещей более вероятным, чем другое, логика заставляет вас приписывать им одинаковую вероятность. Но абстрактное состояние дел может концептуализироваться и описываться разными способами. Предположим, вы преследуете вора и попадаете на развилку трех дорог, две из которых идут вверх и одна — вниз. Поскольку у вас нет оснований полагать, что вор предпочел одну дорогу другой, вероятность того, что он пошел вниз, должна, согласно этому принципу, быть один к трем. Но поскольку у вас также нет причины думать, что он пошел вверх, а не вниз, то та же самая вероятность должна быть один к двум. В этом случае принцип недостаточного основания слишком неопределен, чтобы как-то помочь в вычислении или определении вероятности.

Перейдем к рассмотрению проблемы предпочтений. В экспериментах испытуемых спрашивали, купят ли они товар (компьютерные аксессуары, вино и тому подобное) по цене, которая соответствует двум последним цифрам номера их полиса социального страхования. Затем их про-

Окончание сн.1

нетах. Если это событие будет победой его любимой спортивной команды, он может поставить деньги на то, что она проиграет, так что независимо от исхода у него будет что-то, чем он сможет удовлетвориться.

² Это утверждение, вероятно, слишком категорично. Применяя эту процедуру, можно добиться любого результата в диапазоне вероятности между 50 и 80 %, не более и не менее. В данном случае мы можем оправдаться, если скажем, что подопытный верит, что событие скорее произойдет, чем нет, но до конца в этом не уверен. Однако такая оценка будет гораздо более приблизительной, чем это допускают стандартные модели принятия решений.

силы установить максимальную сумму, которую они готовы заплатить за этот товар. Выяснилось, что номер полиса существенным образом повлиял на их платежеспособность. К примеру, те, у кого номера полиса были в верхнем квинтиле, согласились заплатить в среднем 56 долларов за беспроводную компьютерную клавиатуру, тогда как те, чьи номера были в нижнем квинтиле, готовы были заплатить только 16. Хотя предполагалось, что процедура должна помочь уловить или выявить предварительные предпочтения, результаты показали, что выявлять нечего, объект исследования отсутствует. Цифры определялись привязкой к номерам полиса социального страхования, а не с какими-то реальными предпочтениями.

Существуют также свидетельства того, что баланс разных человеческих ценностей крайне нестабилен и может в равной степени быть обусловлен как процедурными особенностями, так и лежащей за ним психической реальностью. Такого рода баланс можно выявить через *выбор* (choice) или *подбор соответствия* (matching) в экспериментах. Участникам эксперимента могут предложить выбрать между спасением множества жизней при высокой цене за каждую спасенную жизнь (А) или спасением меньшего числа жизней при более низкой цене (Б). В противоположность этому испытуемых могут попросить назвать цену, при которой им будет все равно, выбрать спасение множества жизней по этой цене (опция В) или опцию Б. Предположим, что некий участник устанавливает значение ниже цены А. Поскольку этот человек одинаково относится к В и Б и, как предполагается, предпочитает вариант В варианту А (потому что В позволяет спасти то же самое число жизней по более низкой цене), он должен предпочесть и выбрать вариант Б, а не вариант А. Подавляющее большинство участников действительно выбрали цену для В ниже, чем для А, однако две трети заявили, что выберут А, а не Б. Более высокая ценность спасения жизни явственнее выделяется при выборе, а не при поиске баланса, хотя логически эти процедуры должны быть эквивалентными.

Есть и другие причины, почему мы не должны принимать утверждения об установках и других ментальных состояниях за чистую монету. Особую проблему в этом отношении составляют религиозные убеждения. В Англии начала XVII столетия такие деятели Церкви, как епископ Эндрюс, могли утверждать, что чума послана богом, чтобы покарать грешников, и в то же время бежать из Лондона в сельскую местность. Вера в то, что в силу их божественного происхождения французские короли могли своим прикосновением исцелять от золотухи, явно ослабела к концу XVIII столетия, когда традиционная формула «Король касается тебя, Господь тебя исцеляет» была заменена на условное наклонение — «Король тебя касается, да исцелит тебя Господь». Рвение, с которым ко-

ролевский двор искал документальные свидетельства успешных исцелений, также указывает на недостаток веры.

В качестве современного примера можно указать, что поведение исламских террористов-самоубийц может быть отчасти объяснено их верой в загробную жизнь, к которой они получают привилегированный доступ благодаря своему мученичеству. Можно задаться вопросом: такая же это вера, как и наша вера в то, что солнце завтра снова взойдет, то есть используется ли она с равной уверенностью как предпосылка к действию? Это не вопрос уверенности и вероятности, а вопрос веры и ее отсутствия. Я могу обладать твердой уверенностью (и даже буду готов на этом основании держать пари) в вероятности, основанной на многочисленных событиях в прошлом. Но вера в загробную жизнь, разделяемая многими людьми, вероятно, совсем иного рода. Скорее это некая призрачная квазивера, разделяемая в силу ее потребительской ценности, а не как предпосылка к действию. Если бы все, кто декларирует веру в загробную жизнь, придерживались ее в полной мере, как достоверной возможности, то нам встречалось бы больше мучеников, чем мы наблюдаем сегодня. Хотя некоторые верующие относятся именно к такому типу и террористы-самоубийцы могут преимущественно рекрутироваться из их числа, я подозреваю, что для многих религия служит не побуждением к действию, а утешением, когда решение уже принято³.

Подобным образом люди склонны испытывать квазиэмоции, отличающиеся от настоящих тем, что не выражаются ни в каких действиях. Некоторые негодующие по поводу нищеты в странах третьего мира и никогда не берушиеся за бумажник, могут наслаждаться своим негодованием как потребительским товаром, потому что это позволяет им хорошо о себе думать. Точно так же очевидное наслаждение, с которым некоторые скорбели о смерти принцессы Дианы, не имело ничего общего с чувством настоящей скорби. Думаю, что подходящим определением этому была бы *сентиментальность* (немецкое *Schwärmerei* еще ближе подходит к сути). Оскар Уайльд определял сентиментального человека как «желающего наслаждаться роскошью испытывать эмоции, ничем за это не заплатив». Принимает такая расплата форму пожертвования в Оксфордский комитет помощи голодающим или форму страдания — их отсутствие позволит сделать вывод о том, что мы имеем дело с ненастоящей эмоцией⁴.

³ Идея религии как опиума для народа также характеризует ее скорее как потребительский товар, нежели как предпосылку к действию. В то же время она может быть предпосылкой и для бездействия.

⁴ Сходное явление имеет место, когда люди иногда смотрят на себя со стороны. Пословица гласит: «Добродетель себя не знает». Точно так же невозможно заявлять о собственной наивности, поскольку понятие предполагает

Сюда же относится проблема силы самовнушения. Как только мы узнаём, что *X* — это предположительно *Y*, мы начинаем утверждать и *верить*, что это действительно *Y*. Величайшие эксперты по Вермееру поверили очевидным (как *сейчас* представляется) подделкам ван Меегерена. Пруст упоминает «способность, дающую возможность раскрыть содержание симфонического фрагмента после ознакомления с программкой и обнаружить внешнее сходство в ребенке, когда вы знаете, из какой он семьи». Один любитель джаза из Европы полностью изменил свое мнение о Джеке Тигардене, когда узнал, что тот не был черным. Если мы расположены к какому-то автору, мы можем приписывать глубокий смысл тому, что непредвзятому наблюдателю может показаться не более чем банальностью. Мы проецируем на мир наши ожидания, а потом заявляем, что мир подтверждает и оправдывает наши чаяния.

Смысл этих замечаний в том, что нам не следует рассматривать убеждения, желания, предпочтения, эмоции и тому подобные вещи как стабильные и устойчивые сущности наравне с яблоками и планетами. В последующих главах будет много примеров ускользающей, неустойчивой и зависимой от контекста природы ментальных состояний. Учитывая все сказанное, читатели смогут найти подтверждения псевдоточности или вымышленной строгости, от которой я предостерегаю. Чтобы пойти дальше констатаций, не имеющих определенных выводов (например, «Агенты придают большее значение благосостоянию в настоящем, чем в будущем»), нам придется сказать что-нибудь о том, *насколько* большее значение они придают настоящему. Сделав это, мы неминуемо придем к заключениям, проводящим более четкие различия, чем те, что мы наблюдаем в поведении агентов. Фокус (скорее искусство, чем наука) в том, чтобы знать, когда упрощение проясняет вопрос, а когда затемняет.

В этой книге я нередко ссылаюсь на бессознательную работу мозга. Редукция диссонанса (глава I), принятие желаемого за действительное (глава IV) и трансмутация мотивов (глава IV), например, связаны с работой бессознательных механизмов. Мы можем не до конца понимать, как они действуют, но их существование невозможно отрицать. Многие выступали в пользу существования бессознательных психических состояний. Самообман, в отличие от принятия желаемого за действительное, предполагает, что есть бессознательные установки. Фрейд считал, у всех нас есть бессознательные желания, в которых мы не можем признаться. Кроме того, могут быть бессознательные эмоции и предрассудки.

В той мере, в которой бессознательные психические состояния обладают каузальной силой, их можно идентифицировать по их воздей-

неосознанность. Ниро Вульф в одном из романов Рекса Стаута сказал: «Заявлять о достоинстве — значит подделывать его».

ствию. Если, например, высказывание «Эта женщина слишком щедра на уверенья» отрицают с непропорциональной силой, мы можем заключить, что именно в него верит указанный индивид, пусть и бессознательно⁵. О Зигмунде Фрейде рассказывают историю (источник которой я не смог найти), что его пригласили познакомиться со знаменитым деятелем международного еврейского движения, доктором Х. Во время разговора доктор Х спросил: «Скажите, доктор Фрейд, кто, по-вашему, сегодня самый важный представитель еврейской нации в мире?» Фрейд вежливо ответил: «Я думаю, что это, должно быть, вы». А когда доктор Х ответил: «Нет, нет», Фрейд спросил: «А разве одного нет было недостаточно?» Двойное отрицание может быть равносильно утверждению.

Бессознательные предрассудки можно идентифицировать по их внешним проявлениям. В экспериментах участников просили быстро поделить (постучав себе по левому или по правому колену) все имена, приведенные в списке, на те, что считаются именами чернокожих (такие как Малик или Лашонда), и те, что чаще всего рассматриваются как имена белых (такие как Тиффани или Питер). Затем их попросили быстро классифицировать каждое из слов списка как приятное по значению (например, *любовь* и *младенец*) и неприятное (например, *война* или *рвота*). После этого они классифицировали наугад составленный список, включавший все имена чернокожих, белых, приятные и неприятные слова. Сначала испытуемых просили постучать по левому колену, когда они услышат любое имя чернокожего или неприятное слово, и по правому колену, если услышат приятное слово. Затем инструкции изменились. Их просили стучать по левому колену в случае имен белых людей и неприятных слов, и по правому колену — в случае имен чернокожих и приятных слов. Для выполнения второй задачи понадобилось в два раза больше времени, хотя объективно это были задачи одинаковой сложности.

Бессознательные эмоции часто могут быть идентифицированы наблюдателями, которые делают вывод об их существовании из характерных физиологических или поведенческих реакций. Многие из нас слышали и с раздражением произносили фразу «Я не раздражен». Зависть может проявляться в резкости тона и в презрительном взгляде, которые очевидны наблюдателю, но не самому субъекту. В «Красном и черном» мадам Реналь раскрывает свои чувства к Жюльену Сорелю, только когда у нее возникает подозрение, что у него роман с ее горничной; одна эмоция (ревность), таким образом, вскрывает наличие другой (любви).

Самообман (см. главу VII) в этом отношении более проблематичен. Предположим, у меня сложилось позже вытесненное представле-

⁵ Шекспир У. Гамлет. — Примеч. пер.

ние, что у моей жены роман с моим лучшим другом. Хотя и неосознанно, но убеждение, что они любовники, может по-прежнему руководить моими действиями, например, заставляя избегать той части города, в которой живет мой друг и в которой я могу столкнуться с женой, посетившей его. Это может показаться правдоподобной историей, но насколько я знаю, нет доказательств того, что бессознательные установки имеют каузальное действие. Многие аргументы в пользу существования самообмана основываются на (1) столкновении человека с убедительными доказательствами того, во что он или она не желают верить, и (2) на факте, что агент проявляет себя и действует на основании иных, более позитивных побуждений. Чтобы добыть прямые улики сохранения в бессознательном негативных установок, необходимо показать (3), что оно способно руководить действиями, как в вышеприведенном гипотетическом примере. Повторюсь, я не знаю ни одного доказательства этого эффекта.

Мне представляется, что такого явления не существует. Приятный самообман — вообразить, будто мои бессознательные убеждения стоят на службе сознательных и ограждают меня от доказательств, которые могли бы подорвать эти сознательные убеждения, но это не более чем безосновательная история в духе «просто так». Рассуждая в этом же ключе, можно было бы вообразить, что бессознательное способно порождать не прямые стратегии (один шаг назад, два шага вперед), например, заставляя ребенка причинять себя боль, чтобы привлечь внимание родителей. Такие предположения делают бессознательное слишком похожим на сознательную психическую деятельность, приписывая ему представления о будущем (глава VI) и о действиях и намерениях других людей (глава XIX). Не осознаваемые нами психические состояния могут вызывать спонтанные реакции, такие как два «нет» вместо одного, но мне неизвестны свидетельства того, что они могут вызывать инструментально рациональное поведение.

Библиографические примечания

Свидетельства укорененности оценок вероятности приводятся в «*Принятии решений в неопределенности: Правила и предубеждения*» Д. Канемана и др. (Канеман Д., Словик П., Тверски А. Принятие решений в неопределенности: Правила и предубеждения. М.: Генезис, 2005). Свидетельства укорененности предпочтений можно найти в статье «*Последовательная произвольность*» Д. Ариэли, Дж. Левенштейна и Д. Прелека (Ariely D., Loewenstein G., Prelec D. Coherent arbitrariness // Quarterly Journal of Economics. 2003. No. 118. P. 73–105). Сведения о эпизопе Эндрюсе взяты из книги «*Секретари бога*» А. Николсона (Nicolson A. God's Secretaries. New York: HarperCollins, 2003), а информация об исцеляющем прикосновении у М. Блока в «*Королях-чудотворцах*»

(Блок М. Короли-чудотворцы. Очерк представлений о сверхъестественном характере королевской власти, распространенных преимущественно во Франции и в Англии. М.: Языки русской культуры, 1998). Хорошее обсуждение сентиментальности есть в статье «Сентиментальность» М. Таннера (*Tanner M. Sentimentality // Proceedings of the Aristotelian Society. 1976–1977. No. 77. P. 127–47*). О масштабном (Интернет) эксперименте рассказывается в «Сборе сведений об имплицитных групповых отношениях и убеждениях на демонстрационном вебсайте» Б. Нозека, М. Банаджи и А. Гринвальда (*Nosek B., Banaji M., Greenwald A. Harvesting implicit group attitudes and beliefs from a demonstration website // Group Dynamics. 2002. No. 6. P. 101–115*). Что касается аргумента о том, что Фрейд сделал бессознательное слишком похожим на сознание, см.: «Новое бессознательное» Л. Наккаша (*Naccache L. Le nouvel inconscient. Paris: Odile Jacob, 2006*).

IV. Мотивации

Эта и две последующие главы будут посвящены разновидностям мотиваций. В настоящей главе представлены довольно общие замечания по теме. В дальнейшем я сосредоточусь на двух специфических вопросах: эгоизм против альтруизма и временная близорукость против прозорливости. Эти вопросы дополняют друг друга, последний является вневременным вариантом первого, внеличностной оппозицией. Как мы увидим, они по сути связаны друг с другом в том смысле, что прозорливость может имитировать альтруизм.

Набор мотиваций человека — пирог, который можно нарезать по-разному. Хотя ни один из них не может претендовать на статус канонического, существует четыре подхода, которые представляются мне важными. Первый предполагает континуум мотиваций, второй и третий — трихотомию, четвертый — дихотомию. Классификации одновременно в чем-то сходные и любопытным образом различающиеся, что позволяет нам рассмотреть одно и то же поведение под разным углом зрения.

ОТ ИНТУИТИВНОГО К РАЦИОНАЛЬНОМУ

11 сентября 2001 года люди бросались из окон Всемирного торгового центра из-за непереносимой температуры. «Это вряд ли можно счесть добровольным выбором, — сказал Луис Гарсия, начальник пожарной охраны Нью-Йорка. — Если подвести человека к окну и создать такую температуру, есть веские основания полагать, что большинство людей будут вынуждены прыгнуть». Реальной альтернативы не было. Субъективно это можно сравнить с опытом тех, кто в отсутствие пресной пьет морскую воду. Они могут знать, что даже если морской воды выпить немного, все равно вступаешь на опасный путь: чем больше пьешь, тем сильнее жажда. И все-таки некоторые не могут устоять перед соблазном. Тяга к веществам, вызывающим привыкание, может рассматриваться в этом же ключе. Писатель XVIII века Бенджамин Раш предложил драматический пример. «Когда друзья стали уговаривать горького пьяницу бросить пить, он ответил: Если бы в одном конце комнаты стоял бочонок рома, а из другого постоянно стреляла бы пушка, я все-таки бы прошел прямо перед пушкой, чтобы заполучить ром». Сексуальное желание также может быть настолько сильным, чтобы заглушить соображения осторожности.

Интенсивность некоторых эмоций может быть такой, что они вытесняют остальные соображения. Например, чувство стыда может быть совершенно непереносимым, как это показывает пример американского морского адмирала, в 1996 году совершившего самоубийство накануне дня, когда должны были предать огласке, что он незаслуженно носил некоторые свои медали. Или же шесть случаев самоубийства среди французов после того, как было сообщено, что они потребляют порнографию для педофилов. Гнев тоже может стать непреодолимо сильной эмоцией, как это случилось 9 июля 2006 года, когда Зинедин Зидан на последних минутах финального матча ЧМ по футболу боднул своего итальянского оппонента головой в живот в ответ на провокацию на глазах 70 тысяч человек, присутствовавших на стадионе, и приблизительно миллиарда телезрителей по всему миру. Если бы перед этим он остановился и хотя бы на секунду задумался, то понял, что такое действие может стоить его команде победы и навредить личной репутации.

За исключением, возможно, только потребности прыгнуть из горящего здания Всемирного торгового центра, сомнительно, чтобы все эти желания были в буквальном смысле неодолимы, подобно тому как невозможно остановить камень, катящийся со склона. (Позыв ко сну может быть неодолимым, но засыпание не является действием, вот почему все попытки заставить себя заснуть обречены на неуспех.) Люди с продуктовой зависимостью иногда восприимчивы к ценам: когда те вырастают, они снижают потребление продукта¹. Люди в спасательных лодках иногда могут помешать друг другу пить соленую воду. Сексуальный соблазн и позыв к убийству безусловно можно преодолеть. Из-за своей интенсивности эти инстинктивные стремления, тем не менее, занимают крайнюю часть спектра человеческих мотиваций. У них есть потенциал (пусть даже не всегда реализуемый) блокировать сознательные установки, компромиссы и даже выбор.

В другой части спектра мы видим эталон (или скорее карикатуру) рационального агента, который не подвержен воздействию инстинктивных факторов, включая эмоции. Он действует только после того как тщательно (но не тщательнее, чем того требуют обстоятельства) взвесит последствия каждой имеющейся альтернативы. Рациональный генерал, менеджер высшего звена или врач обеспокоены исключительно тем, чтобы найти наилучшие средства для достижения объективной цели, такой как победа в войне, максимизация прибыли или спасение жизни. Субъективные желания или составляющие их основу инстинкты не включаются в это уравнение. Хотя в главе XI мы увидим, что кон-

¹ Это может происходить потому, что бюджет не позволяет им поддерживать прежний уровень потребления (глава IX).

цепция рациональности гораздо шире, чем эта бледная идея, она все же может послужить нам ориентиром.

Примером расхождения между инстинктивной (visceral) и рациональной мотивациями может послужить различие между инстинктивным и благоразумным *страхом*. Хотя страх часто называют эмоцией, его можно считать только комплексом убеждений и желаний. Когда я говорю: «Боюсь, что сейчас пойдет дождь», — я *всего лишь* делюсь предположением, что будет дождь, и своим нежеланием, чтобы он пошел. Если страх побуждает действовать, как в случае, когда я беру с собой зонт, он является примером рационального поведения (глава XI), в нем нет ни одной из характерных черт эмоции (глава VII). В свою очередь инстинктивный страх может побудить к действию, которое не является инструментально рациональным. Например, было подсчитано, что 350 американцев, которые в противном случае ~~остались бы живы~~, погибли в автомобильных авариях, потому что боялись летать на самолетах после 11 сентября 2001 года. Испанцы же, похоже, напротив, не продемонстрировали роста смертности, пересев с поезда на автомобиль после теракта на железной дороге в Мадриде 11 марта 2004 года. Возможно, что из-за продолжительности террористических атак баскской организации ЭТА население стало испытывать к ним благоразумный, а не инстинктивный страх. Для них риск террористического акта стал одним из многих, подобным, хотя и якобы более опасным, риску того, что пойдет дождь.

Между крайностями инстинктивного-рационального континуума мы находим поведение, которое частично мотивировано инстинктивными факторами и в то же время восприимчиво к экономическим соображениям. Человек может стремиться отомстить (инстинктивное желание) и при этом тянуть время до тех пор, пока не застанет врага врасплох (соображения благоразумия). Если он вызывает своего врага на дуэль (как того требует кодекс чести), он может тайно брать уроки фехтования (нечестная, но полезная практика). Если человеку сделали предложение одновременно нечестное и выгодное в том смысле, что если он его примет, ему будет лучше, чем если не примет, то он может принять или отвергнуть его в зависимости от силы эгоизма или энергии протеста². В более сложных случаях один инстинктивный фактор может противодействовать другому. Желание адюльтера может быть нейтрализовано чувством вины. Вызванное страхом желание спастись бегством может уравниваться или замещаться желанием сражаться, спровоцированным гневом.

² В подробно рассмотренном примере (см. главу XX) в результате отказа должна будет пострадать также сторона, сделавшая предложение, — фактор, усиливающий энергию протеста.

ИНТЕРЕС, РАЗУМ И СТРАСТЬ

В своем анализе человеческих мотиваций французские моралисты XVII века ввели плодотворное различие интереса, разума и страсти. Интерес — это погоня за личной выгодой, будь то деньги, слава, власть или спасение души. Любое действие, направленное на то, чтобы помочь нашим детям, преследует личный интерес, поскольку наша судьба тесно связана с их судьбами. Родитель, отправляющий детей учиться в дорогую частную школу, не жертвует своими интересами, а действует им во благо. Можно сказать, что страсти включают эмоции так же, как другие инстинктивные побуждения, например голод, жажду, сексуальные и другие влечения. Древние включали в эту общую категорию душевное расстройство, потому что оно, подобно эмоциям, непроизвольно, неконтролируемо и подрывает рациональные соображения. В некоторых случаях мы можем относить к страстям состояние интоксикации. С точки зрения закона гнев, алкогольное опьянение и безумие часто рассматривались почти на равных.

Разум — более сложная идея. Моралисты часто использовали его (как я буду его здесь использовать) в связи с желанием поддерживать общий, а не частный интерес. Изредка они пользовались им, ссылаясь на долгосрочные (благоразумные) мотивации, в противоположность краткосрочным (близоруким) соображениям. Обе идеи можно подвести под категорию *беспристрастности*. При разработке общественной политики к индивидам нужно относиться беспристрастно, а не защищать интересы одних групп и индивидов в ущерб другим.

Индивиды также могут действовать, исходя из подобной мотивации. Родители могут жертвовать своими интересами, послав детей в государственную школу, потому что исповедуют равенство возможностей. В то же время политики, как и частные лица, должны в равной мере беспристрастно оценивать долгосрочные последствия, придавая им одинаковый с краткосрочными результатами вес в процессе принятия решения. Действительно, некоторые моралисты утверждали, что озабоченность долгосрочным интересом способствует общественному благу. На Филадельфийском конституционном конвенте Джордж Мейсон, в частности, заявил:

Мы должны заботиться о правах людей каждого класса. Равнодушные высших классов общества к этому велению гуманизма и политики поразительно, учитывая, что, какими бы зажиточными они сейчас ни были или какое бы высокое положение ни занимали, возможно (более того, наверняка), их потомки в течение нескольких лет распространятся среди самых низших классов общества. Таким образом, каждый эгоистический мотив, каждая семейная привязанность должны отстаивать такую политическую систему, которая заботилась бы о пра-

вах и благополучии нижних слоев граждан с неменьшим тщанием, чем о верхних слоях.

Каждая форма беспристрастности имеет свою градацию. Сила беспокойства за других имеет тенденцию варьировать обратно пропорционально не только генеалогической, но географической удаленности. Сходным образом даже благоразумные индивиды до некоторой степени придают больше значения ближайшему будущему, чем более отдаленному, — факт, который только частично можно объяснить их знанием о том, что они могут и не дожить до этого отдаленного будущего.

В качестве примера того, как понимать поведение в категориях любой из этих трех мотиваций, мы можем процитировать письмо канцлера Нью-Йорка Роберта Ливингстона Александру Гамильтону, датированное 1783 годом, в котором он комментирует преследования тех, кто выступал в войне за независимость на стороне британцев. -

Я хочу серьезно пожаловаться вам на дух жестоких преследований, который преобладает здесь и последствий которого для благосостояния, торговли и будущего я так страшусь. Он особенно причиняет мне боль, потому что почти лишен каких бы то ни было *более чистых патриотических мотивов*. У немногих это слепой дух мести и негодования, но у большинства он представляет собой *самый низменный интерес*.

Фразы, выделенные мною курсивом, соответствуют разуму, эмоциям и интересу. Эпитеты говорят сами за себя: разум — *чистый*, страсть — *слепая*, интерес — *низменный*. Я вернусь к некоторым следствиям таких оценок.

ОНО, Я, СВЕРХ-Я

В анализе человеческих мотиваций Фрейд подсказал три базовые формы, каждая из которых связана с отдельной психической подсистемой. Эти три системы — «Оно», «Я» и «сверх-Я», соответствующие принципу удовольствия, принципу реальности и совести. «Оно» и «сверх-Я» включают в себе импульсы и контроль над ними, тогда как «Я», «беспомощное в обоих отношениях... тщетно защищает себя как от подстрекательств смертоносного Ид, так и от упреков наказывающей совести». В еще более показательном высказывании из того же эссе («Я и Оно») Фрейд писал, что «Я» — «несчастное существо, исполняющее три рода службы и вследствие этого страдающее от угроз со стороны трех опасностей — внешнего мира, либидо Оно и суровости сверх-Я». Но даже эта формулировка не полностью передает то, что, по моему мнению, явля-

ется сутью фрейдовской идеи. Это предположение о том, что в процессе существования «Я» во внешнем мире (принцип реальности) ему также приходится вести борьбу на два фронта — против импульсов, исходящих от «Оно» (принцип удовольствия), и с предположительно строгим контролем за этими импульсами, осуществляемым «сверх-Я» (совесть)³.

Это оригинальное, глубокое и верное предположение. Единственное, чего ему не хватает, так это механизма. Почему «Я» само не может осуществлять необходимое управление всевозможными импульсами? Почему мораль и совесть так часто принимают форму жестких правил? Следует нам постулировать существование отдельных или квазиавтономных психических функций? Чтобы получить удовлетворительные ответы на эти вопросы, потребовались новаторские работы Джорджа Эйнсли (George Ainslie). Я обсуждаю его взгляды в главе XIII. Здесь я только хочу привлечь внимание к тому факту, что многие импульсы необходимо держать в узде по причине *совокупного* ущерба, который они могут нанести, если их не сдерживать⁴. В отдельно взятом случае чрезмерное потребление алкоголя или переизбыток, мотовство или прокрастинации (откладывание выполнения домашнего задания, например) совсем необязательно влекут значительный ущерб для агента. Ущерб возникает после повторяющихся злоупотреблений (или повторяющихся неудач). Таким образом, управление импульсами не должно фокусироваться на индивидуальных случаях, так как человек всегда может сказать себе, что с завтрашнего дня он начнет новую, более правильную жизнь. Управление импульсами должно учитывать то, что они предсказуемо возникают в неопределенном количестве случаев. Решение можно найти, если переформулировать проблему так, что неспособность контролировать импульс в данном конкретном случае становится индикатором неспособности справиться с ним в дальнейшем. «Да, я могу отложить сдерживание этого импульса до завтра, но чем завтрашний день будет отличаться от сегодняшнего? Если у меня ничего не получится сегодня, то и завтра ничего не получится». Запустив эффект внутреннего домино и повысив ставки, агент может получить мотивацию побороть импульсы, которая у него отсутствовала бы, если бы он

³ Если соединить две фрейдовские метафоры, это как всадник, сидящий на неуправляемой лошади («Оно»), которым одновременно понукает инкуб («сверх-Я»).

⁴ Помимо этого существует фактор суммарного риска. Вероятность нежелательных последствий от секса без предохранения может быть мала в данном конкретном случае, но совокупный риск в течение всей жизни может быть значителен. Во время отдельно взятой поездки на автомобиле риск получить повреждения, не пристегнув ремень безопасности, невелик, но совокупная вероятность в течение всей жизни составляет 1 к 3.

рассматривал каждый день в отдельности. Обратной стороной медали может стать неослабность контроля. Как говорили викторианские моралисты, «*никогда не знает исключений*».

УЧИТЫВАЯ ПОСЛЕДСТВИЯ

Наконец, мотивации могут быть консеквенциалистскими (consequentialist) или неконсеквенциалистскими (nonconsequentialist), то есть ориентированными на результат действия или на само действие. Экономическое поведение по большей части ориентировано на результат. Когда люди откладывают деньги на старость, а брокеры покупают и продают акции, они не придают какой-либо существенной — позитивной или негативной — ценности самим этим действиям; их интересует только результат. Безусловный пацифист, отказывающийся идти на военную службу даже для борьбы с самыми злыми врагами, не принимает в расчет последствий своего поведения. Самое главное для него то, что некоторые действия, такие как убийство человека, подлежат безусловному запрету. Дело не в том, что он *не осознает* последствия, как в случае эмоционального действия; просто последствия не имеют значения для того, что он делает.

В публичной политике также действует два вида мотиваций. Тот, кто разрабатывает политику, может применить принцип «Находка принадлежит нашедшему» (как в патентном законодательстве), предположив, что если человек, открывший ценный ресурс, получит право собственности на него, то будет открываться больше ценных ресурсов. Это консеквенциалистский аргумент. Неконсеквенциалистским аргументом в пользу такой политики будет утверждение, что человек, открывший ресурс, будь то кусок земли или лекарство от рака, обладает естественным *правом* собственности на него. В качестве другого примера мы можем рассмотреть речь (XXXI) Диона Хризостома против практики Родееана использовать старые бронзовые статуи для почитания благотворителей города, что, как он утверждал, одновременно нарушало права тех, в честь кого эти статуи были первоначально воздвигнуты, и отвращало новых благотворителей, которые понимали, что статуи, воздвигнутые в их честь, вскоре снова будут использованы для кого-нибудь другого. Консеквенциалистские аргументы могут (как кажется) оправдывать жесткие меры против террористов, даже если предпринятые шаги нарушают неконсеквенциалистские ценности, связанные с правами человека и гражданскими свободами⁵.

⁵ Оговорка в скобках (*как кажется*) отражает возможное действие психологии тирании (глава II). Классическая дилемма сдерживания заключается в том, что *ненависть*, которую она вызывает, может в конце концов заслонить собой *страх*, который она призвана внушить.

Особым случаем неконсеквенциалистской мотивации является принцип, который будет фигурировать у меня под разными именами — бытовое кантианство, категорический императив или магическое мышление, — «*Делай то, что для общего блага следовало бы делать всем*». В определенном смысле этот принцип связан с последствиями, так как агент делает то, что даст наилучший результат, если бы все остальные делали то же самое. Однако это последствия не *его* действия, а гипотетической серии действий агента и других людей. В некотором конкретном случае действие согласно этому принципу может иметь катастрофические последствия, если остальные ему не последуют. На международной арене примером может служить одностороннее разоружение.

Другим случаем является следующий принцип еврейской этики. Предположим, враг у ворот и он заявляет: «Отдайте мне одного из вас, он будет убит, и мы пощадим всех остальных. Если откажетесь, мы убьем вас всех». Талмуд предписывает евреям погибнуть, но не выдавать жертву во имя спасения остальных. Однако если в тех же обстоятельствах враг говорит: «Отдайте мне Петра», то это требование приемлемо. Это не запрет на то, чтобы пожертвовать кем-то ради спасения остальных, а запрет на выбор того, кем пожертвовать. Роман «Выбор Софи» представляет ту же самую дилемму.

Социальные нормы (глава XXII) представляют еще один специфический случай неконсеквенциалистского поведения с важным неожиданным поворотом. Они предписывают людям, что надо делать (например, отомстить за обиду или воздерживаться от того, чтобы есть детеныша, сваренного в молоке его матери), не потому что это даст какой-то желаемый результат, но потому, что это действие само по себе обязательно⁶. Хотя эти действия совершаются *не ради* какого-то результата, их можно рассматривать как действия, призванные *предотвратить* какой-то другой результат, а именно обвинения в том, что они были допущены. Но следом мы можем спросить: не выдвигаются ли эти обвинения по сходным консеквенциалистским причинам? Более того, если люди уязвлены действиями других людей, они отвечают тем же самым даже при однора-

⁶ Относительно законов кашрута (кошерности), исторической иллюстрацией которых является, в частности, запрет на то, чтобы есть детеныша, сваренного в молоке матери, какое-то время считалось, что они были оправданы соображениями гигиены. Сегодня, насколько мне известно, их рекомендуют на том основании, что человеку полезно делать нечто трудное и одновременно бессмысленное. Как представляется, эта идея воплощает заблуждение побочных эффектов, которым я уделяю внимание позднее: поведение, которое оправдывается *только* его действием на формирование характера, не возмещает даже этого действия. Хотя я и не знаю, сколько человек из тех, кто соблюдает это правило, делают это именно по этой причине.

зовом взаимодействии в условиях полной анонимности, которая может быть достигнута при проведении эксперимента. Так как интеракция односторонняя, испытуемые ничего не получают от столкновений в будущем, а так как она анонимная, они не боятся обвинений со стороны третьих лиц. Я еще вернусь к этим экспериментам в следующих главах.

Даже для ярого приверженца неконсеквенциализма последствия могут иметь значение, если они важны. Рассмотрим принцип, который многие считают безусловным, — запрет на пытку детей. Представим, что в сценарии тикающей бомбы необходимым и достаточным условием для предотвращения ядерного взрыва в центре Манхэттена является пытка маленького ребенка террористки у нее на глазах. Если можно было бы сделать этот сценарий достоверным, многие неконсеквенциалисты согласились бы с пытками. Другие сказали бы, что поскольку условия этого сценария не достижимы на практике, абсолютный запрет остается в силе. Еще кто-то сохранил бы запрет, даже если такой сценарий возможен. Моя задача здесь не ратовать за один из этих выводов, а сделать эмпирическое наблюдение, согласно которому в реальных жизненных ситуациях ставки редко бывают такими высокими, чтобы можно было заставить неконсеквенциалиста рассматривать последствия своего поведения. Возможно, что если бы на карту было поставлено нечто большее, он отказался бы от своих принципов, но так как до этого дело никогда не доходит, мы не можем сказать наверняка, имеем мы дело с чрезвычайной стоворчивостью или же с полным отказом идти на компромисс.

Эти четыре подхода к мотивации позволяют уловить некоторые из вышеупомянутых феноменов. Инстинктивные факторы, страсти и принцип удовольствия явно имеют много общего. Последний применяется к более широкому ряду случаев, поскольку включает не только стремление избежать боли, но и влечение к удовольствию. Когда студенты затягивают выполнение домашней работы, это происходит не потому, что есть что-то еще, чем они страстно желают заняться. Очень часто они просто идут по пути наименьшего сопротивления.

Суперэго, разум и неконсеквенциалистские мотивации тоже имеют много общего. Хотя не все формы морали являются жесткими и неумолимыми, некоторые действительно таковы. Общеизвестным примером является теория морали Канта. (На самом деле его философия морали происходит из частных правил, которые навязывал себе для контроля над собственными импульсами, например максимум не выкуривать больше одной трубки после завтрака.) В то же время в индивидах, не испытывающих отвращения к неопределенности, мораль может преодолеть упорность. Действительно, терпимое отношение к неопределенности часто считается отличительной чертой здорового эго. Между рациональ-

ностью, личным интересом, эго и консеквенциализмом существуют более тонкие отношения. Абсурдно было бы утверждать, что отличительной чертой здорового эго является рациональное следование личному интересу.

ПОТРЕБНОСТИ И ЖЕЛАНИЯ

Часто мы понимаем мотивации как *потребность установить некое положение дел*. Они также могут принимать форму *желания добиться какого-то положения дел*. Это различие между потребностями и желаниями важно, если мы обратимся к мотивационному компоненту эмоции (глава VII). Эмоции действительно часто могут сопровождаться либо потребностью что-то сделать, либо желанием, чтобы ситуация сложилась определенным образом. В гневе или ярости (*wrath*) потребность А в том, чтобы отомстить Б, не может быть удовлетворена тем, что В сделал с Б то, что с ним хотел сделать А, или тем, что с Б произошел несчастный случай. Важен не результат, не страдания Б, а то, что причиной этих страданий стали действия А. В садизме также важно именно заставить другого страдать, а не только сам факт его страданий. В случае *ненависти*, наоборот, важно то, чтобы ненавистный человек или группа людей исчезли с лица земли, и неважно, произойдет это благодаря моим или чьим-то другим действиям. В случае *злости* (*malice*) значение также имеет страдание другого, а не то, чтобы именно я был причиной этого. И действительно, злоумышленник может даже отступить прежде, чем предпринять активные действия по нанесению вреда другому человеку, не просто из страха разоблачения, но потому, что это было бы несовместимо с его представлением о себе. Еще четче это демонстрирует *зависть*. Многие люди, которые хотели бы, чтобы их недруг потерял все свое имущество, и которые пальцем не пошевелили бы, чтобы это предотвратить, никогда не предпримут активных действий для его уничтожения, даже если им этого ничего не стоит и они при этом ничем не рискуют⁷. Человек, который не станет поджигать соседний дом, возможно, не станет вызывать пожарных, если увидит, что тот загорелся.

Принятие желаемого за действительное (глава VII) основывается на желаниях, а не на потребностях. В некоторых случаях агент может отказаться от тяжелой работы по приведению мира в соответствие с его желаниями и вместо этого пойти по более легкому пути принятия присущих этому миру установок. Если я желаю получить повышение, но неохотно прила-

⁷ У некоторых завистливых людей наверняка нет таких угрызений совести. Они могут жить в обществе, в котором зависть не считается постыдной, или могут быть просто лишены стыда.

гаю к этому усилия, я могу довольствоваться незначительными знаками, чтобы убедить себя в том, что неминуемо его получу. В других случаях мы лишены возможности воздействовать на мир. Я могу быть не в состоянии заставить другого человека ответить мне взаимностью в любви или сделать так, чтобы мой больной ребенок выздоровел. В таких случаях я могу либо предаваться фантазиям, либо смотреть фактам в лицо. Дальнейшее различие может быть проведено между случаями, в которых фантазии не имеют последствий и в которых они становятся предпосылкой к действию. Я могу обманывать себя, полагая, что моя знакомая питает ко мне тайную страсть, и не делать никаких авансов либо потому, что меня принуждает к этому мораль (или эгоизм), либо потому, что самообольщению придаются исключительно ради него самого, его потребительской ценности⁸. Человек также может открыто поделиться с объектом питаемыми на его счет подозрениями, как это случилось с секретаршей Джона Мейнарда Кейнса, которая сказала ему, что видит его пылкую страсть к ней, но ничем не может помочь. После этого ее жизнь была разрушена.

СОСТОЯНИЯ, ЯВЛЯЮЩИЕСЯ ПО СУЩЕСТВУ ПОВОЧНЫМИ ПРОДУКТАМИ

Фактор, затрудняющий различие желания-потребности, состоит в том, что в некоторых случаях я могу получить X, сделав A, но только если я сделаю A ради того, чтобы получить X. Если я много работаю для того, чтобы объяснить нейрофизиологическую основу эмоций, и добиваюсь успеха, я могу заработать хорошую репутацию. Если я уйду в работу ради политических целей, то под конец могу обнаружить, что у меня сформировался характер. Если я хорошо играю на фортепьяно, то могу произвести впечатление на окружающих. Эти косвенные выгоды паразитируют на основной цели деятельности. Если у меня буквальная мотивация, такая как желание заработать репутацию, у меня меньше шансов действительно ее заработать. Если кто-то присоединяется к политическому движению *только* ради развития своего сознания или формирования характера, он обречен на неудачу или же может добиться успеха только случайно. Как заметил Пруст, музыкант «может иногда предать [свое истинное призвание] ради славы, но ища славы таким образом, он только отдаляется от нее, а находит, только от нее отвернувшись».

Самосознание выражается в исполнении. Как он также писал, хотя

⁸ Мы можем отметить для себя, чтобы в дальнейшем (глава XXIII) развить, что мышление, противоположное принятию желаемого за действительное, не может иметь потребительской ценности. Поэтому раз оно все-таки возникает, то с большей вероятностью становится побуждающим мотивом к действию.

«самый лучший способ сделать так, чтобы вас искали — сделаться недоступным», он никому не дал бы подобного совета, поскольку «этот метод достижения социального успеха работает, только если его не применяют сознательно».

Музыкальная слава или социальное преуспевание попадают в категорию *состояний, которые по сути являются побочными продуктами*, — состояний, которых нельзя добиться действиями, мотивированными исключительно желанием их достичь. Эти состояния могут *возникнуть*, но их нельзя *вызвать намеренно* простым решением. Они включают желание забыть, желание верить, желание желать (например, желание преодолеть сексуальную импотенцию), желание заснуть, желание смеяться (невозможно щекотать себя самого) и желание преодолеть заикание. Попытки реализовать эти желания скорее всего окажутся малоэффективными и могут ухудшить *поддержание*. Общим местом у моралистов и романистов стало представление о том, что намеренный гедонизм обречен на провал⁹ и что ничто так не запечатлевает в памяти какой-либо опыт, как попытка его забыть. Хотя мы можем *желать*, чтобы эти состояния реализовались, нам следует избегать *намерения* их реализовать.

Многие люди озабочены *спасением* (в загробной жизни) и *искуплением* (зла, которое они совершили). Они также могут полагать, что в состоянии достичь этих целей своими поступками. Мученическая смерть в битве с неверными может стать пропуском в рай (по крайней мере некоторые так считают). Борьба с нацистами после сотрудничества с ними на предшествующей стадии может искупить совершенное зло. Однако если эти действия совершаются *ради достижения* спасения или искупления, они могут закончиться неудачей. В католической теологии намерение купить место в раю ценой добровольного мученичества будет примером греха симонии. Некоторые специалисты по исламу подвергают сходной критике террористов-самоубийц, движимых верой в то, что таким способом они получают привилегированное место в раю. Монтень пишет: «Когда спартанцы, эти отличные судьи в делах добродетели, стали решать, в свою очередь, кому из них принадлежит честь свершения в этот день наиболее выдающегося деяния, они пришли к выводу, что храбрее всех сражался Аристодем; и все же они не дали ему этой почетной награды, потому что его доблесть воспламенялась желанием смыть пятно,

⁹ В заключительном романе цикла «В поисках утраченного времени» Пруст, размышляя, возможно, о собственной жизни, писал, что тщетные поиски счастья могут, тем не менее, привести к такому пониманию человеческой жизни, которое само по себе «может принести радость». Поиски состояний, являющихся побочными продуктами, может вызвать их косвенным образом, как например, когда детское требование засмеяться может вызвать у человека смех самой нелепостью этого требования.

которое лежало на нем со времени Фермопил»¹⁰. Французский газетный магнат Жан Пруво (Jean Prouvost), сотрудничавший во время оккупации с нацистами, попытался искупить свою вину, выписав для членов Сопротивления чек на большую сумму в тот момент, когда стало ясно, что немцы проиграют войну. После освобождения ему было присуждено *non-lieu* (решение суда, по которому дело приостанавливается или прекращается, не дойдя до суда), чего спартанцы, по всей видимости, не сделали бы¹¹.

ТОЛКАТЬ ИЛИ ТЯНУТЬ

Почему люди переезжают из одной страны в другую? Почему ученые переходят из одного университета в другой? Часто эти проблемы характеризуют в словах *тянуть* (pull) и *толкать* (push). Человек может эмигрировать, потому что ситуация дома непереносима или ситуация за границей представляется более предпочтительной; по крайней мере, таков распространенный подход к этому вопросу. Но во многих ситуациях он может вводить в заблуждение. Как правило, люди уезжают, *сравнив* ситуацию дома и за границей и сочтя, что различие настолько велико, что оправдывает переезд даже с учетом затрат на него¹². И все же имеет смысл различать мотивы побуждения (толчка — «push motives») и мотивы стремления (рывка — «pull motives»): первые находятся в инстинктивной части континуума, а вторые — ближе к его рациональной части. Люди, охваченные страхом, часто бегут от опасности, но не в спасительном направлении. Единственная их мысль — бежать; они не задумываются, что могут попасть из огня да в полымя. В зависимости от вида вещества наркомана может мотивировать или стремление к эйфории (кокаин), или побуждение, вызванное состоянием депрессии (героин). Суицидальное поведение тоже связано скорее с побуждением, чем со стремлением. Это попытка спастись от отчаяния, а не порыв к чему-то.

Действие социальных норм (глава XXII) также может рассматриваться в категориях побуждения и стремления. Многие индивиды испытывают сильную потребность выделиться при помощи различных форм

¹⁰ Монтень М. Опыты. Кн. I. Гл.: О Катоне-младшем. — Примеч. пер.

¹¹ Причина, по которой его освободили, возможно, в том, что движение Сопротивления нуждалось в деньгах и позднее сочло себя обязанным соблюдать молчаливое обещание иммунитета, которое предполагало принятие чека.

¹² Эта формулировка предполагает, что расходы на переезд входят в оценку целесообразности переезда на равных с выгодами от него. Однако затраты на переезд могут накладывать *ограничения* на принятие решения. Если стоимость самого дешевого трансатлантического рейса превышает максимальную сумму, которую может отложить или занять бедный итальянский крестьянин, он останется дома, какой бы привлекательной ни была жизнь в США (глава IX).

социального поощрения, будь то *слава* (желание быть лучшим) или *почести* (победа в состязании или в поединке). Другие индивиды больше озабочены тем, чтобы избежать позора, связанного с нарушением социальных норм. В одних обществах есть социальная норма «Не высовывайся!». Отличиться — значит отклониться в сторону, а отклонения — объект универсального осуждения: «За кого он себя принимает?» Относительная сила этих двух мотиваций варьирует в разных обществах. Классические Афины могут служить иллюстрацией стремления отличиться¹³. В современных малых городах зачастую царит враждебность по отношению к желающим выделиться. Если рискнуть и сделать обобщение, в целом толчок, задаваемый стыдом, кажется более важной мотивацией, чем рывок к славе, что не исключает силы последнего.

КОНФЛИКТ МОТИВАЦИЙ

Существование *конкурирующих* мотиваций — это общее место.

Мне так нужна эта книга, что я готов украсть ее из библиотеки, но при этом я хочу вести себя морально.

Перед лицом хулигана я одновременно испуган и разгневан, мне хочется убежать и в то же время ударить его.

Я хочу, чтобы все дети получали государственное образование, но вместе с тем хочу, чтобы мой ребенок ходил в частную школу и получил самое лучшее образование.

Мне нужен кандидат, выступающий в поддержку абортов, но одновременно поддерживающий снижение налогов.

Я хочу курить и при этом сохранить здоровье.

Если мне делают выгодное, но нечестное предложение, мне хочется одновременно отклонить его по второй причине и принять — по первой.

Я хочу жертвовать на благотворительность, но в то же время соблюдать собственные интересы.

Я испытываю соблазн изменить жене, но также хочу сохранить брак.

Как разрешается конфликт этих мотиваций? Общий ответ должен быть таким. Когда речь идет о ситуации, в которой «победитель забирает все», так что никакой (физический) компромисс невозможен, побеждает более сильная мотивация¹⁴. Если забота о ребенке для меня важнее, чем

¹³ Эсхил писал свои трагедии для представления на драматических состязаниях. Когда его победил юный Софокл, он был так опечален, что уехал из Афин на Сицилию.

¹⁴ Можно (по крайней мере в принципе) использовать вероятностный ком-

озабоченность образованием детей вообще, я пошлю его в частную школу. Если моя неприязнь к абортам сильнее, чем озабоченность налогами, а кандидат, предлагающий обе опции, отсутствует, я буду голосовать за того, кто против абортот и за повышение налогов. Если кто-то предложит мне три доллара из общего фонда десять долларов, намереваясь оставить все остальное себе, я приму это предложение. Если мне предложат всего два доллара, я отвергну это предложение¹⁵. Когда компромисс возможен, более сильная мотивация оказывает большее воздействие, чем слабая. Курильщик может снизить потребление сигарет с тридцати до десяти в день. Демонстрируя широту своего альтруизма, я могу тратить 5% моего дохода на благотворительность¹⁶.

Это не совсем неверный, но слишком простой ответ, так как идея сильной мотивации сложнее, чем предполагают эти подобранные на скорую руку примеры. Мотивация может быть обязана своей силой чисто психическому потенциалу; вот почему, например, инстинктивные мотивы зачастую гораздо сильнее того, что Мэдисон называл слабым голосом разума. Но сильной может оказаться мотивация, которую агент активно одобряет, в силу того, что она высоко оценивается обществом. По сути культура имеет свою нормативную иерархию мотиваций. При прочих равных условиях человек скорее совершит данное действие, руководствуясь мотивом А, чем мотивом Б, поскольку А выше стоит в иерархии. Это *метамотивации*, желания, запускаемые желаниями определенного рода¹⁷. Даже будучи более слабыми с точки зрения инстинктивности, они в конце концов могут возобладать над другими мотивациями.

Эгоистический интерес, особенно страсть, часто демонстрирует определенное *преклонение* перед разумом¹⁸. Сенека говорил: «Разум желает,

промисс, устроив лотерею опций, пользуясь силой оснований в пользу той или иной опции как их весом. Некоторые институты на этом основании выделяют небольшие ресурсы, но я не встречал ни одного случая индивидуального мотивационного конфликта, который разрешался бы таким образом.

¹⁵ Как ранее отмечалось, отвергая предложение, я также оставляю ни с чем предлагающего.

¹⁶ В главе VI я обсуждаю более странный феномен «Проигравший забирает все», наблюдаемый при недостатке воли.

¹⁷ Идея метамотиваций не имеет отношения к идее метапредпочтений. Примером последней может стать человек, у которого две разные системы предпочтений: с одной стороны, он предпочитает еду диете, с другой — диету еде, а метапредпочтение выбирает последнее. Если следовать догадке Лабрюера «Люди тщеславны и больше всего ненавидят, когда их считают таковыми», то метамотивация может предпочесть диету из соображений здоровья ее соблюдению из тщеславных побуждений.

¹⁸ Как мы увидим в главе XII, агенты часто проявляют чрезмерное уважение к рациональности.

чтобы решение, которое он принимает, было справедливым, гнев желает, чтобы решение, которое он принимает, казалось справедливым». Поскольку существует огромное множество внешне достоверных концепций разума, справедливости и честности, вполне возможно представить решение, принятое в гневе, как разумное. Суды над коллаборационистами в странах, оккупированных нацистской Германией во время Второй мировой войны, зачастую были ведомы жаждой мести. Но из-за преклонения перед разумом в сочетании с желанием отмежеваться от незаконной практики оккупационных режимов новые лидеры представляли жесткие меры как основанные на справедливости, а не на эмоциях. Человек может иметь первоочередной заинтересованность в том, чтобы не жертвовать на благотворительность, и во вторую очередь хотеть, чтобы его не считали побуждаемым исключительно личным интересом. Из преклонения перед разумом он может принять философию благотворительности (глава II), которая оправдывает небольшие пожертвования. Если другие жертвуют много, он займет утилитаристскую позицию, делая небольшие взносы, а если другие жертвуют мало, он будет действовать так же, руководствуясь соображениями честности.

В этом случае разум не играет независимой роли. Он лишь постфактум оправдывает действия, которые совершались на совсем иных основаниях. Конфликт не разрешается, а загоняется под ковер. В других случаях поиск оправданий, отсылающих к разуму, может изменить поведение. Если я применяю к благотворительности подход, основанный на честности, когда другие делают небольшие взносы, то мне придется следовать за ними, если они начинают жертвовать больше. *Та же самая потребность в самоуважении*, заставившая меня оправдывать эгоистическое поведение соображениями беспристрастности, мешает мне изменить концепцию беспристрастности теперь, когда она больше не действует в мою пользу. Мы можем вообразить, что в «Короле Лире» Бургундия и Франция влюбились в Корделию из-за ее перспектив, но только Бургундия была так мало озабочена своим имиджем, чтобы смочь отбросить эмоции, когда они стали противоречить ее интересам. Это тот случай, когда интерес отступает перед страстью, а не перед рассудком, принимая по внимание, что страсть (скорее эта конкретная страсть) стоит выше, чем эгоизм в нормативной иерархии. Другие чувства, как, скажем, зависть, вполне могут расположиться на более низкой ступени в сравнении с эгоизмом. Часто при этом мы можем наблюдать попытки совершать такие основанные на зависти поступки, которые могут быть убедительно представлены в качестве эгоистических. Действия, которые нельзя выставить в этом свете, совершаться не будут.

Согласно теории когнитивного диссонанса, когда одна мотивация *немного* сильнее другой, она пытается привлечь к себе союзников так, чтобы

мотивы на одной стороне приобрели решающее значение. Ум неосознанно ищет дополнительные аргументы в пользу предварительного вывода, к которому он пришел сознательно¹⁹. В таких случаях сила мотивации не может приниматься как данность, но должна рассматриваться скорее как продукт самого процесса принятия решения. Когда я покупаю автомобиль, я придаю ценность неравнозначным параметрам (скорости, цене, комфорту, внешнему виду) каждой из альтернатив и прихожу к общему заключению, сравнивая результаты оценки каждой ценности. Я, например, могу дать бренду А оценку 50, а бренду В — оценку 48. Поскольку сравниваемые величины оказались близки, я могу бессознательно модифицировать значения таким образом, чтобы А стал очевидным победителем, например, поставив 60 против 45. Прежде чем сделать покупку, я сталкиваюсь с брендом В, который по старой шкале оценок получил бы у меня 55, но по новой получает 50. Если бы я рассматривал альтернативы в порядке В-А-В, я выбрал бы В. Но так как они попались мне в порядке А-В-В, я выбрал А. Такая *зависимость от траектории* подрывает простую идею, что конфликт мотиваций разрешается в соответствии с их силой.

Согласно тому, что я назвал упрощенным взглядом, решение, красть ли книгу из библиотеки, может быть представлено следующим образом. На одной чаше весов выгода от использования книги, на другой — цена, которую придется заплатить в форме чувства вины. Что я в итоге сделаю, будет зависеть от того, превысит цена выгоду или наоборот. Но результат может измениться, если, предположим, кто-то предложит мне пилюлю от вины, снимающую любое болезненное чувство, вызванное кражей. Если бы вина входила в мои решения только как психическая затрата, было бы рационально проглотить пилюлю, подобно тому как целесообразно принять таблетку от похмелья перед планируемым возлиянием. Я утверждаю, что *большинство людей будут испытывать одинаковое чувство вины как из-за кражи книги, так и из-за принятия таблетки*²⁰. Я не отрицаю, что

¹⁹ Возможно обратное явление, когда ум сознательно ищет обоснование для решения, которое было принято бессознательно. Простым примером может послужить поведение, индуцированное гипнозом. Или возьмем надуманное объяснение, которое влюбленные изобретают, чтобы оправдать слишком частые телефонные звонки предмету своей страсти. Эту форму самоинтерпретации, или ложной самоинтерпретации, не следует путать с трансмутацией, которая возникает перед действием, а не после его совершения.

²⁰ В главе XI я утверждаю, что человек с краткосрочным горизонтом по похожим причинам откажется принять пилюлю, которая заставила бы его придавать больше значения последствиям его действий. Общий принцип, который иллюстрируют эти две пилюли, заключается в том, что рациональный человек не станет в два этапа проделывать то, что не хочет делать в один. Он наверняка может захотеть сделать в два этапа то, что не может сделать в один.

может быть достигнут компромисс между моралью и эгоистическим интересом, только он не может быть представлен в упрощенном виде.

Вот более сложный случай. Я хотел бы не хотеть есть торт. Я хочу есть торт, потому что люблю его. Я хотел бы его не любить, потому что, будучи до некоторой степени тщеславным человеком, полагаю, что важно сохранять стройность фигуры. При этом я хотел бы быть менее тщеславным. *Но не активизируется ли это желание только в момент возникновения желания съесть торт?* В конфликте между моей тягой к тарту, желанием сохранить стройность и желанием не быть таким тщеславным первое и последнее могут объединиться и вытеснить второе. Они могут добиться успеха, застав меня врасплох, но если я *понимаю*, что мое сопротивление тщеславию вызвано тягой к тарту, я смогу им противостоять. В другой раз мое желание краткосрочного удовлетворения и долгосрочное стремление к спонтанности могут объединиться против среднесрочного желания самоконтроля. Когда при выборе между двумя опциями действует более двух мотивов, идея силы мотивации может оставаться неопределенной до тех пор, пока мы не узнаем, какой союз был заключен.

Французский моралист XVII века Лабрюер резюмировал две формы мотивационного конфликта: «Для страсти нет ничего легче, чем победить разум, но величайшая ее победа — завоевать эгоистический интерес». Мы видели, что даже когда страсть побеждает разум, она может захотеть сделать его своим союзником. Хотя апостол Павел сказал: «Добра, которого хочу, не делаю, а зло, которого не хочу, делаю», — более распространенной реакцией было бы стремление убедить себя в справедливости сделанного под воздействием страсти. Когда страсть завоевывает интерес, она может добиться этого двумя способами. Из-за типичной для этой эмоции *настойчивости* (глава VIII) агент может не располагать временем на раздумья о том, в чем его интерес. И наоборот, сила эмоции может быть столь велика, что агент будет *намеренно* действовать вопреки своему интересу. Такое поведение может стать проявлением слабости воли (глава VI).

Библиографические примечания

Теория инстинктивных мотиваций представлена в Дж. Лёвенштейн «Без контроля: инстинктивные влияния на поведение» (Loewenstein G. Out of control: Visceral influences on behavior // *Organizational Behavior and Human Decision Processes*. 1996. No. 65. P. 272–92). Оценка прироста автомобильных аварий после 11 сентября 2001 года взята из Дж. Джигерензер «Чудовищный риск, 11 сентября и автомобильные аварии со смертельным исходом» (Gigerenzer G. Dread risk, September 11, and fatal traffic accidents // *Psychological Science*. 2004. No. 15. P. 286–287. Отсутствие роста таких аварий в Испании отмечает А. Лопес-Руссо «Из-

бегая смертельного риска избежать чудовищного риска: последствия 11 марта в Испании» (Lopez-Rousseau A. Avoiding the death risk of avoiding a dread risk: The aftermath of March 11 in Spain // *Psychological Science*. 2005. No. 16. P. 426–428). Трехотомия «интерес — разум — страсть» анализируется А. Хиршманом в работе «*Страсти и интересы*» (Hirschman A. *The Passions and the Interests*. Princeton, NJ: Princeton University Press, 1977); М. Уайтом в «*Философия, федералист и конституция*» (White M. *Philosophy, The Federalist, and the Constitution*. Oxford University Press, 1987) и в моей книге «*Алхимия разума*» (Alchemies of the Mind. Cambridge University Press, 1999). «Пикоекономика» Джорджа Эйнсли (Ainslie G. *Picoeconomics*. Cambridge University Press, 1992) предлагает механизм, которого не хватает фрейдовской теории. Классическое исследование побуждения и стремления есть у Д. Гамбетта в «*Они сами прыгнули или их подтолкнули?*» (Gambetta D. *Did They Jump or Were They Pushed?* Cambridge University Press, 1983). Аргументы Диона Хризостома взяты мной из П. Вейн «*Греко-римская империя*» (Veyne P. *L'empire gréco-romain*. Paris: Seuil, 2005. P. 217). Приведенный мной принцип еврейской этики разбирается в Д. Доб «*Сотрудничество с тиранией в раввинистическом законе*» (Daube D. *Collaboration with Tyranny in Rabbinic Law*. Oxford University Press, 1965) и в Д. Доб «*Смирение или сопротивление*» (Daube D. *Appeasement or Resistance*. Berkeley: University of California Press, 1987). Идею состояний, являющихся побочными продуктами, я разрабатываю в главе 2 книги «*Зеленый виноград*» (Sour Grapes. Cambridge University Press, 1983) и применяю ее в статье «*Искупление греха*» (Redemption for wrongdoing // *Journal of Conflict Resolution*. 2006. No. 50. P. 324–38), а к вопросу спасения — в статье «*Мотивации и убеждения в самоубийственных террористических актах*» (Motivations and beliefs in suicide missions / Gambetta D. (ed.). *Making Sense of Suicide Missions*. Oxford University Press, 2005). См. также Л. Росс и Р. Нисбетт «*Человек и ситуация*» (Ross L., Nisbett R. *The Person and the Situation*. Philadelphia: Temple University Press, 1991. P. 230–232). Преклонение перед разумом во время судов над коллаборационистами я обсуждаю в главе 8 книги «*Подведение итогов*» (Closing the Books. Cambridge University Press, 2004). Свидетельства изменения важности различных элементов процесса выбора можно найти в А. Браунштейн «*Тенденциозный процесс принятия предварительных решений*» (Brownstein A. *Biased predecision processing* // *Psychological Bulletin*. 2003. No. 129. P. 545–68) и в Дж. Брем «*Изменения оценки привлекательности альтернатив после принятия решения*» (Brehm J. *Postdecision changes in the desirability of alternatives* // *Journal of Abnormal and Social Psychology*. 1956. No. 52. P. 384–389).

V. Эгоизм и альтруизм

МОТИВАЦИЯ И ПОВЕДЕНИЕ

Сравнение эгоистической и альтруистической мотиваций может выглядеть обманчиво простым. Для начала давайте проясним, что под *альтруистической* мотивацией мы понимаем желание способствовать благополучию (welfare) других даже в ущерб своему собственному, а под *альтруистическим действием*, — такое, для которого альтруистическая мотивация является достаточным основанием. Если я вижу, как вы подаете нищему на улице, я называю это актом альтруизма, поскольку такой поступок *может быть* проявлением альтруистической мотивации вне зависимости от того, так ли это.

В качестве более сложного примера рассмотрим результат экспериментов по «альтруистическому наказанию», к которым мы еще будем возвращаться. В этих исследованиях участнику А предоставляется возможность наказать участника Б за отказ от сотрудничества, заплатив за это некоторую сумму. Между участниками нет непосредственного контакта, они никогда больше не встретятся. Несмотря на это многие участники пользуются опцией наказания, принуждая Б быть сговорчивым при будущем взаимодействии с третьей стороной, В. Наказание *может быть* вызвано альтруистической мотивацией, если А предвосхищает и мотивирован выгодой, которую наказание Б принесет В. В действительности он мотивирован скорее жадой мести.

Существуют примеры такого поведения вне лабораторных условий. Во Франции XVIII века крестьяне обычно удовлетворяли просьбы бродяг и нищих, давая им пищу и кров. Если кто-то отказывался, его деревья могли быть вырублены, скот покалечен, а дом сожжен. Такие акты разрушения не приносили выгоды нищим и подвергали их риску быть пойманными с поличным. Хотя нет оснований полагать, что мотивацией нищих действительно было желание заставить крестьян принимать других нищих в будущем, такой мотивации было бы достаточно, чтобы объяснить эти поступки. Крестьянские бунты в доиндустриальной Англии с их непосредственными целями были обычно малоуспешными, а их лидеры жестоко наказывались. Но в долгосрочной перспективе в силу неудобств, которые они причиняли, эти бунты имели успех, заставив имущие классы вести себя более сдержанно, чего в противном случае те не стали бы делать.

Определение альтруистических мотиваций в категориях жертвоприношения собственным благополучием, а не в категориях материальных благ, позволяет исключить случаи, подобные следующему. Если я плачу 100¹ тысяч долларов за пребывание ребенка в колледже, причина, возможно, в том, что благосостояние моего сына так тесно переплетено с моим собственным, что жертва идет на пользу нам обоим¹. Такая мотивация хотя и обращена на другого, не является альтруистической². Проявлением истинного альтруизма был бы факт определения ребенка в государственную школу, тогда как я могу позволить себе частную и считаю, что в ней ребенку было бы лучше. В этом случае я пожертвовал бы благополучием не только своего единственного ребенка, но и собственным. Точно так же донорство для банка крови (а не для своего близкого родственника) с большей вероятностью вызвано альтруистическими мотивами. Впрочем, на практике иногда невозможно понять, является мотивация альтруистической или только направленной на других (*other-regarding*).

Дальнейшие осложнения возникают в связи с тем, что одни и те же люди *любят отдавать*, потому что от этого они лучше себя чувствуют («эффект теплого свечения» — «the warm glow effect»). Если это «теплое свечение» раскрывает *причину*, по которой они помогают другим, мы вряд ли захотим назвать их альтруистами. Они отдают, потому что в конечном счете от этого им становится лучше. Это не значит, что альтруист не испытывает этого «теплого свечения», но только оно не включено в (бессознательные) мотивы, по которым он помогает другим. И снова нужно отметить, что это различие почти невозможно провести на практике. Я подробнее вернусь к этому в главе XV.

С какими бы проблемами ни сталкивалась идентификация альтруистических мотиваций, примеры такого *поведения* встречаются в изобилии. Фонд Карнеги регулярно награждает медалями тех, кто подвергался большому риску ради спасения жизни других. Многие люди сдают кровь совершенно безвозмездно³. В Норвегии органы для пересадки в боль-

¹ Возможно, я платил бы за образование даже в том случае, если бы его высокая стоимость подорвала мое благосостояние. В этом смысле внесение более низкой платы объясняется другими мотивами, вытесняющими альтруистические.

² Когда Маркс писал, что коммунистическое общество не будет основано на альтруизме, а будет находиться *по ту сторону* эгоизма и альтруизма, он, возможно, имел в виду, что любые *материальные* жертвы, на которые люди шли друг для друга, в аналогичных обстоятельствах больше не повлекут жертвоприношения *благосостоянием*.

³ В действительности считалось, что безвозмездная сдача крови может отсечь людей с инфекционными заболеваниями, занимающихся донорством ради денег.

шинстве случаев предоставляют родственники реципиента. Удаление почки влечет медицинские риски, но никакого материального вознаграждения за это нет⁴. Многие индивиды, особенно женщины, ухаживают за престарелыми родителями помимо того, что ходят на работу и заботятся о своих семьях. Во многих странах более половины взрослого населения регулярно жертвует деньги на благотворительность. После цунами 2004 года в развитых странах наблюдался пик количества и размера пожертвований. Во время войны многие люди скрывали свои физические недостатки, чтобы им позволили воевать. Многие солдаты добровольно отправлялись на выполнение опасных (и даже самоубийственных) заданий. Когда люди голосуют на общенациональных выборах и тем самым вносят свой вклад в демократию, они несут некоторые расходы и не получают практически никаких личных выгод. Этот перечень можно было бы продолжать.

Причина того, почему мы не можем вывести альтруистические мотивации из альтруистического поведения, заключается в том, что другие мотивации могут *имитировать альтруизм*. В категориях, предложенных в главе IV, мы можем рассматривать альтруизм как своего рода *разум*, *симулируемый интересом* или *страстью*. (Слова *подражать* или *имитировать* могут, но не должны вызывать ассоциаций с сознательными усилиями, направленными на то, чтобы скрыть от других истинные мотивы.) Многие люди, не слишком озабоченные тем, чтобы быть бескорыстными, стремятся к тому, чтобы их считали таковыми. Таким образом, Юм был не прав, когда утверждал: «Любовь к славе в вознаграждение за добродетельные дела является *верным* доказательством любви к добродетели» (курсив мой. — Ю. Э.). Монтень же утверждал: «И чем больше шума поднимают вокруг того или иного хорошего дела, тем меньшего оно стоит в моих глазах, так как во мне рождается подозрение, что оно совершено скорее ради того, чтобы вокруг него поднялся шум, чем из-за того, что оно хорошее: выставленное напоказ, оно уже наполовину оплачено». В пределе единственными добродетельными поступками являются те, что никогда не становятся достоянием гласности. Бабушка прустовского рассказчика с ее ангельской добротой настолько впитала этот принцип, что приписывала всем своим хорошим поступкам эгоистические мотивы. В той степени, в какой добродетель склонна держаться в тени, за ней может стоять нечто большее, чем то, что привлекает взгляд. По другим причинам за ней наверняка может быть и нечто меньшее.

⁴ В большинстве стран торговля почками живых людей считается незаконной. В данном случае закон мотивирован желанием защитить как опустившихся индивидов от самих себя, так и реципиентов от получения органов для пересадки низкого качества.

СТРЕМЛЕНИЕ К ОДОБРЕНИЮ И СТЫДЛИВОСТЬ

Монтень признавал, что добродетель встречается редко, когда проводил различие между настоящей и фальшивой мотивационной «монетой» — действием во имя блага и действием ради того, что скажут другие. Поскольку первая мотивация встречается редко, политики предпочитают пользоваться второй.

...Это заблуждение человеческого ума имеет заслуги перед обществом... Это оно побуждает людей быть верными своему долгу... Так пусть же это заблуждение укореняется все глубже и глубже; и пусть его насаждают в нас, насколько это возможно. Поскольку люди в силу несовершенства своей природы не могут довольствоваться доброкачественной монетой, пусть между ними обращается и фальшивая. Это средство применялось решительно всеми законодателями, и нет ни одного государственного устройства, свободного от примеси какой-нибудь напыщенной чепухи или лжи, необходимых для того, чтобы налагать узду на народ и держать его в подчинении⁵.

Наполеон откликнулся на эту идею, когда сказал, защищая создание в 1802 году ордена Почетного легиона: «Вот такие побрякушки ведут людей вперед» (ветераны из его республиканской армии горячо протестовали против этого нововведения). *Стремление к одобрению* (approbation), желание, чтобы другие думали о вас хорошо, — это фальшивка, которая может заменить настоящую «монету» альтруизма и морали. В то же время *стыдливость* (shamefulness) — желание, чтобы другие не думали о вас плохо, — может служить в качестве фальшивки. Социальные нормы могут удерживать людей от поступков, которые они совершили бы в противном случае. Однако чтобы другие думали о вас хорошо, недостаточно простого соблюдения норм. Одобрением пользуется то, что является *излишним* (supererogatory), то есть идет дальше нормативного. То, что в одних обществах обязательно, может быть излишним в других. В Норвегии и в США существует (мягкая) социальная норма, согласно которой родственник должен отдать почку для трансплантации, если это требуется (и если есть совместимость)⁶, тогда как во Франции такое проявление альтруизма будет считаться чрезмерным. В некоторых социальных кругах благотворительные пожертвования считаются обязательными.

Эти мотивации можно проиллюстрировать примерами из политики XVIII века. На первом Национальном собрании Франции (1789–1791) де-

⁵ Монтень М. Опыты. Кн. II. Гл. 16: О славе. — Примеч. пер.

⁶ В США врачи часто помогают пациенту сопротивляться подобному давлению, говоря им уже на ранней стадии, что, если потребуется, они готовы предоставить медицинское обоснование отказа от донорства.

путаты несколько раз поступались своими интересами, начиная от отказа от феодальных привилегий до признания себя непригодными для избрания в законодательный орган. Хотя их мотивы были сложными, важным компонентом было желание продемонстрировать бескорыстие. По словам биографа одного из них, они «упивались бескорыстием». Примерно в то же самое время в Соединенных Штатах Джордж Вашингтон постоянно заявлял о том, что боится, что другие люди подумают, что им движет личный интерес. (В то же время он понимал, что слишком большая озабоченность собственной добродетельностью может выглядеть недобродетельно.) В качестве другой пары примеров рассмотрим две концепции чести. Согласно одной честь должна *приобретаться* через доблестные поступки. Согласно другой честь принимается за базовую линию, но может быть *потеряна* из-за постыдных поступков.

Могут ли стремление к одобрению или стыдливость имитировать альтруизм, зависит от существенного критерия, которым пользуются при оценке поведения. Некоторые общества могут придавать большую ценность, а значит, стимулировать проявление добродетелей, которые не стремятся выдать себя за проявления альтруизма. Жажда славы может повлечь все виды социально разрушительного поведения. Наполеоновские «побрякушки» были рассчитаны на то, чтобы побудить солдат жертвовать жизнью во славу Франции, а не для того, чтобы способствовать благосостоянию французов. Некоторые люди выбирают путь самоотречения, потому что общество превозносит религиозных виртуозов, но монахи и отшельники больше думают о служении Богу, а не о своих собратьях. Насколько я знаю, сообщества, в которых высоко ценится образование и знания, не придают альтруизму большей ценности, чем другие сообщества. Культ красоты в современных западных обществах стимулирует эгоистическое поведение, которое представляется не совместимым с заботой о других людях. В обществах, в которых распространена так называемая аморальная семейственность (в качестве примера можно привести юг Италии), есть социальные нормы, запрещающие помогать иностранцам, попавшим в беду, или соблюдать закон. Таким образом, трудно сказать, что имитирует альтруизм — жажда одобрения или стремление избежать вины.

ТРАНСМУТАЦИИ

В предыдущей главе мы упоминали о том, что первоначальная мотивация агента может *трансмутировать* из интереса в разум. За этим процессом стоит механизм *самолюбия* (self-love), стремление к уважению и самоуважению. Если стремление к одобрению и стыдливость, возникающие из потребности в уважении, влияют только на внешнее поведе-

ние, стремление к самоуважению может воздействовать на внутренние мотивации. Многие не желают выглядеть эгоистичными в глазах окружающих. Даже действуя в целях удовлетворения собственных интересов, они стремятся придать своим поступкам оттенок бескорыстия. За несколько десятилетий до Войны Севера и Юга (1861–1865) рабство перестало быть вопросом чистой корысти и стало *делом*, за которое боролись на принципиальной основе. Это очень распространенный тип политической идеологии. Маркс заметил: «Не следует только впадать в то ограниченное представление, будто мелкая буржуазия принципиально стремится осуществить свои эгоистические классовые интересы. Она верит, напротив, что *специальные условия* ее освобождения суть в то же время те *общие условия*, при которых только и может быть спасено современное общество и устранена классовая борьба».

Хотя трудно сказать, является такой оттенок бескорыстия искренним или лицемерным; было бы ошибкой всегда предполагать второе. В действительности у людей в их усилиях оправдать поведение перед самими собой есть две степени свободы. С одной стороны, существует множество *внешне правдоподобных каузальных теорий*, отстаивающих тезис о том, что то, что выгодно для тебя самого, выгодно и остальным. Теории «просачивающегося богатства» («trickle-down theories») предполагают, что бедные тоже получают выгоду от снижения налогов на богатых. С другой стороны, существует множество *внешне правдоподобных нормативных концепций* справедливости, честности или общего блага, что человеку просто не повезло или ему не хватает компетенции, если он не может найти концепцию (согласно какой-нибудь внешне правдоподобной каузальной теории), которая совпадала бы с его личным интересом. В главе IV я отмечал, что для оправдания небольших взносов на благотворительность люди могут выбирать в качестве основы благотворительности честность или полезность⁷. В таких случаях люди инстинктивно тяготеют к соединению каузальной теории и нормативной концепции, которое может оправдать поведение, определяемое эгоистическим интересом. Хотя нам неизвестно, как, мы знаем, что это действительно происходит.

Имитация альтруизма необязательно проистекает из лицемерия или трансмутации. Эгоистические расчеты могут порождать альтруистическое поведение при помощи совершенно открытых механизмов, таких как выбор *под завесой неведения* (behind the veil of ignorance) или вза-

⁷ В этом случае каузальная теория, оправдывающая небольшие пожертвования, если другие дают много, заключается в снижении их предельной полезности для получателя. Используемая для оправдания небольших пожертвований теория, основанная на честности, если другие дают мало, не нуждается в каузальной предпосылке.

имость. В этой главе я приводил пример первого механизма, цитируя аргумент Джорджа Мейсона, что долгосрочные интересы семей должны стимулировать озабоченность благоденствием всех классов общества⁸. Подобные аргументы применимы и к жизни индивида. В обществах с низким уровнем безработицы, переживающих период быстрых структурных изменений, большинство может голосовать за партию, предлагающую большие льготы для безработных, потому что считает увольнение вероятным. То, что внешне кажется солидарностью, может на самом деле оказаться формой подстраховки.

ВЗАИМНЫЙ ОБМЕН

Взаимность (reciprocity) может быть простым отношением, состоящим из двух элементов, как в случае, когда каждая сторона отношений должна выбрать — сотрудничать или не сотрудничать. Одни фермер приступит к уборке урожая в августе, другой — в сентябре; они в состоянии помочь друг другу. Если фермер, чей урожай созревает раньше, попросит у другого помощи, а сам откажется отплатить тем же в сентябре, он вряд ли впредь сможет рассчитывать на поддержку. Стабильные отношения взаимопомощи с большей вероятностью возникнут в случае, когда, не будучи основанными на товариществе, они благоприятствуют его возникновению. Во время Первой мировой войны между некоторыми немецкими и британскими полками существовала молчаливая договоренность о временных перемириях по принципу «Живи и жить давай другим», обстреливая противника с меньшей интенсивностью, чем можно было бы⁹. В этом случае дружественное отношение к противной сто-

⁸ Похожий аргумент предложил губернатор Моррис во время дебатов о представительстве штатов в сенате. «Преданность и значение штатов всегда было проклятием этой страны. Мы, возможно, не можем уничтожить эту змею, но можем вырвать ее зубы. Господь желал, чтобы наши идеи расширялись до масштабов истинных потребностей человека, вместо того чтобы ограничиваться узкими рамками местечковости. В конце концов роль эгоизма в такой политике не так уж велика. Кто может сказать, будет он сам или тем более его дети в следующем году жителями того или иного штата».

В другом случае губернатор Моррис оказался адресатом подобного аргумента. В ответ на то, что Элбридж Джерри поддержал предложение губернатора Морриса ограничить представительство будущих западных штатов, Роберт Шерман ответил: «Мы заботимся о наших потомках, о наших детях и внуках, которые могут оказаться жителями как старых, так и новых западных штатов. Только на основании этого мы не должны заниматься дискриминацией, которую предполагает данный шаг».

⁹ Хотя высшее командование пыталось пресечь подобную практику, ее было трудно отследить.

роне также складывается со временем, но как *результат* сотрудничества, а не как его причина.

Декарт описывал более сложную форму многосторонней или косвенной взаимности.

Основания, заставляющие меня считать, что люди, делающие все исключительно в своих личных интересах, должны не меньше остальных трудиться в пользу другого человека и стараться, насколько это в их силах, всем доставлять удовольствие, если только они хотят сохранить благоразумие, состоят в том, что мы видим обычно, как лица, считающиеся всегда готовыми доставить другим удовольствие, получают взамен множество добрых услуг даже от тех, которых они ни к чему не обязывали, причем они не получали бы этих услуг, если бы люди думали, что они настроены по-иному и что тяготы, несомые ими ради того, чтобы доставлять другим удовольствие, вовсе не столь велики, сколь преимущества, извлекаемые ими из дружбы с теми, кто их знает. Ведь от нас ожидают лишь тех услуг, которые мы способны оказать с удобством для себя, и от других мы ожидаем меньшего; однако часто случается, что немного стоящая кому-то услуга приносит нам немалую пользу и может даже иметь для нас жизненно важное значение. Правда, иногда бывает, что, оказывая благодеяние, мы утрачиваем свою скорбь и, наоборот, что-то выигрываем от причиненного зла; но это никак не меняет правила благоразумия, относящегося к самым частым случаям.

При прямой взаимности *А* помогает *Б* тогда, и только тогда, когда *Б* помог *А*. При косвенной — *А* помогает *Б*, если *Б* помог *В*. Как мы увидим в последующих главах, это разделение применимо и к негативной взаимности: *А* может ударить *Б*, если *Б* ударил *А*, но также если *Б* ударил *В*. Существование косвенной взаимности показывает, что люди могут вести себя как альтруисты для того, чтобы завоевать *репутацию* людей с альтруистическими мотивами. Тогда другим придется решать, отражает их поведение истинный альтруизм или всего лишь стратегическое желание иметь соответствующую репутацию (глава XX). В этом случае репутация будет оцениваться инструменталистски, а не эссенциалистски. Если стремление снискать одобрение заставляет агента искать уважения ради него самого, то репутации добиваются ради материальных выгод, которые она может принести.

Люди могут отвечать друг другу взаимностью в ситуациях одностороннего взаимодействия, которые не предоставляют возможности для последующего вознаграждения. Если *А* совершает альтруистический поступок по отношению к *Б*, последний может ответить первому тем же, даже если оба знают, что дальнейших взаимодействий между ними не будет. Фермер, собирающий урожай в августе, может помочь тому, кто собирает урожай в сентябре, даже зная, что он собирается уехать

к началу следующего сезона. Конечно же, можно найти своекорыстные мотивы такого взаимобмена. Возможно, фермер, собирающий урожай раньше, боится, что второй фермер каким-то образом его накажет, если он не ответит взаимностью, или же его могут подвергнуть ostracismu третьи лица, от которых он зависит. В условиях эксперимента, однако, такие эффекты можно исключить. В экспериментальных играх, которые будут обсуждаться позднее (главы XV и XX), подопытные взаимодействуют анонимно через компьютерные терминалы, тем самым исключаются эффекты непосредственного взаимодействия, такие как стыд или неловкость. Кроме того, эти игры часто задуманы так, чтобы один человек взаимодействовал с конкретным партнером только один раз.

Даже в таких жестких условиях наблюдается феномен взаимности. В игре «Доверительное управление» один игрок, «инвестор», имеет возможность перевести от 0 до 10 единиц своего вклада из 10 денежных единиц другому игроку, «доверительному собственнику». Затем любая переведенная сумма утраивается; таким образом, если инвестор посылает 10, доверительный собственник получает 30. Последний может решить, какую сумму от 0 до 30 (то есть втрое больше того, что перевел инвестор) перевести назад. В одном эксперименте инвесторы переводили в среднем две трети своего вклада и получали обратно в среднем чуть бóльшие переводы. Чем больше был форвардный перевод, тем больше был обратный. Такие результаты согласуются с целым рядом предположений о мотивациях, кроме гипотезы о том, что оба агента мотивированы материальным интересом и осознают это применительно друг к другу. Следуя этой гипотезе, инвестор, ожидая нулевую отдачу, делает нулевой форвардный перевод. Такого результата не наблюдается, следовательно, работают социальные предпочтения, или мотивации, направленные на других. Возможно, отсутствует альтруизм или честность, поскольку в некоторых экспериментах доверительные собственники отдавали обратно столько же, сколько получали, но даже эта сумма превосходила ту, что могла быть продиктована своекорыстием.

МОРАЛЬНЫЕ, СОЦИАЛЬНЫЕ И КВАЗИМОРАЛЬНЫЕ НОРМЫ

В следующих главах я вернусь к следствиям этого и других подобных экспериментов. Здесь я только хочу разграничить три вида мотиваций, направленных на других. *Моральные нормы* включают правило помогать другим людям в беде, правило равенства долей и норму бытового кантианства («делать то, что было наилучшим, если бы все это делали»). *Социальные нормы* (глава XXI) включают правила этикета, правила ме- сти и нормы денежного обращения. То, что я называю *квазиморальными*

нормами, включает правило взаимности («помогай тем, кто помог тебе, и причиняй вред тем, кто причинил вред тебе») и правило условного сотрудничества («сотрудничай, если другие тоже сотрудничают, но не наоборот»). Социальные и квазиморальные нормы являются условными в том смысле, что их применение обусловлено присутствием или поведением других людей. Социальные нормы, как я утверждаю, работают, когда другие могут наблюдать за действиями агента, а квазиморальные нормы — когда сам агент наблюдает за действиями других людей¹⁰. Моральные нормы являются безусловными. Конечно, то, что они предписывают, может зависеть от поведения других людей. Если моя философия благотворительности основана на пользе, то мера сделанного мною добра (и следовательно, отданного) зависит от вклада других. Однако сама норма никак не отсылает к другим дарителям, но только к получателям.

Два случая индивидуальных реакций на нехватку воды может проиллюстрировать различие между социальными и квазиморальными нормами. Когда мэром Боготы был изобретательный Антанас Мокус, при сокращении потребления воды люди следовали квазиморальной норме. Хотя индивидуальный мониторинг был невозможен, суммарное потребление воды в городе показывали по телевизору, чтобы люди могли знать, выполняют другие эту норму или нет. Как оказалось, так поступали многие люди, чтобы происходило условное взаимодействие. Люди говорили себе: «Так как другие урезают свое потребление, будет честно, если я поступлю так же». Когда же нехватка воды наблюдалась в Калифорнии, при ограничении ее потребления в действие вступили, как представляется, социальные нормы. За наружным потреблением воды (например, для полива лужайки) могли следить соседи или даже муниципальные службы. Внутреннее потребление могли отслеживать гости, которые выражали неудовольствие, если унитаз был чистый¹¹. В дей-

¹⁰ Они могут усиливать друг друга, когда агент может наблюдать за действиями самих наблюдателей. Если я вижу, что вы мусорите, я, возможно, не буду возражать, если вы увидите, как я делаю то же самое. Если же я вижу, как вы аккуратно засовываете в карман обертку от мороженого, чувство справедливости и страх неодобрения могут заставить меня соблюдать норму (см. главу XXII).

¹¹ Экономия воды вызывает озабоченность и в обычное время. В Нью-Йорке она достигается благодаря законам, фиксирующим максимальный объем воды в бачке. В Европе это по большей части достигается тем, что унитаз оснащен двумя кнопками для слива, выпускающими разное количество воды в зависимости от цели. Последняя система представляет интерес, поскольку не основывается ни на возможностях, ни на стимулах (глава IX), но только на доброй воле человека, которую невозможно проконтролировать.

ствительности в Боге тоже имело место отслеживание индивидуального поведения, поскольку дети зачастую «давали прикурить» родителям, если те не экономили воду¹².

Квазиморальные нормы могут обладать большой побудительной силой к альтруистическому поведению. *Имитируют* они альтруизм или *являются* истинными альтруистическими мотивациями? Причина того, что я называю их квазиморальными, а не моральными, заключается кроме прочего в том, что я склоняюсь ко второму ответу. Норма о взаимности позволяет не помогать попавшим в беду людям, если ранее они не помогли нам. Типичная моральная норма — помогать людям, попавшим в беду безо всяких условий, даже если предшествующая история помощи отсутствует. Правило условного сотрудничества позволяет потреблять обычное количество воды, если никто больше не сократил потребление, тогда как утилитарный подход и бытовое кантианство предписывают его одностороннее сокращение. Можно сказать, что моральные нормы носят *упреждающий* (проактивный) характер, тогда как квазиморальные нормы выполняют функцию *реагирования* (реактивную). Еще один способ выразить ту же мысль: представляется, что чувство несправедливости обладает большей мотивационной силой, чем чувство справедливости. Как мы увидим в дальнейшем (глава XX), предложения, которые респонденты в экспериментах обычно отклоняют как нечестные, в результате чего ни они, ни те, кто делал предложение, ничего не получают, как правило, одного и того же объема, что и предложения, которые делаются, если предлагающий не испытывает страха быть отвергнутым.

Представляется, что мы можем выявить работу истинно альтруистических мотивов, если соблюдаются два условия. Во-первых, действие, приносящее благо другим, должно быть проактивным, а не реактивным. Во-вторых, оно анонимно, то есть личность благодетеля неизвестна ни бенефициару, ни третьим лицам¹³. Мы можем представить себе человека, анонимно посылающего деньги в благотворительный фонд «Оксфам» или бросающего деньги в ящик для сбора пожертвований в пустой

¹² Результаты экспериментов говорят о том же механизме. Во время кампаний по энергосбережению в душевых были развешаны знаки, просившие студентов выключать душ, когда они намыливаются, и включать его только для того, чтобы смыть пену. Эффект от этих знаков был минимальный. Однако когда один или два товарища по эксперименту стали выполнять правило, его соблюдение другими студентами резко увеличилось. Хотя их товарищи ничего не говорили остальным, их поведение могло служить молчаливым укором нарушителям.

¹³ В экспериментах личность подопытного неизвестна экспериментатору. В делах благотворительности она скрыта от официальных лиц благотворительных организаций.

церкви. Второй пример не настолько чистый, как хотелось бы, так как человека может мотивировать вера в то, за ним наблюдает Господь, который потом его вознаградит. Вера может быть нелогичной (пример побочного заблуждения) и в то же время очень распространенной. Первый пример может показаться более однозначным. Однако даже акты самого чистого альтруизма, такие как анонимные пожертвования незнакомым людям, могут таить в себе сомнительные мотивы. Согласно Канту,

на самом деле совершенно невозможно из опыта привести с полной достоверностью хотя бы один случай, где максима вообще-то сообразного с долгом поступка покоилась на исключительно моральных основаниях и на представлении о своем долге. Правда, иногда может случиться, что при самом жестоком испытании самих себя мы не находим ничего, что помимо морального основания долга могло бы оказаться достаточно сильным, чтобы побудить к тому или иному хорошему поступку и столь большой *самодотверженности*; однако отсюда никак нельзя с уверенностью заключить, что действительно никакое тайное побуждение себялюбия — только под обманным видом той идеи — не было настоящей, определяющей причиной воли; мы же вместо нее охотно льстим себя ложно присвоенными, более благородными побудительными мотивами, а на самом деле даже самым тщательным исследованием никогда не можем полностью раскрыть тайные мотивы, тогда как речь идет о моральной ценности, то суть дела не в поступках, которые мы видим, а во внутренних принципах их, которые мы не видим.

Кант говорит здесь о том, что даже в момент, когда мы не предстаем перед публикой *внешней*, мы никогда не можем быть уверены в том, что не играем перед *публикой внутренней*. Акт сокрытия своих добрых дел, который Монтень находил столь добродетельным, не может быть скрыт от самого себя. Как говорил Ларошфуко, «гордость всегда возмещает свои убытки и ничего не теряет, даже когда отказывается от тщеславия». Также он заметил: «Чиста и свободна от влияния других страстей только та любовь, которая таится в глубине нашего сердца и неведома нам самим». В лучшем случае, говорил Пруст, мы можем узнать о наших истинных мотивах от других: «нам ведомы только чужие страсти, и то, что мы знаем о собственных, мы смогли узнать только от них. На нас они действуют только косвенно, посредством воображения, которое заменяет наши первоначальные мотивы альтернативными, более приемлемыми».

ВМЕНЕНИЕ МОТИВАЦИЙ

Помимо собственных мотиваций агента объяснения его поведения должны апеллировать к его представлениям о мотивациях других людей. При формировании этих представлений он сталкивается с той же герменевтической дилеммой, что историк или социолог. Поскольку он

не может принимать заявления о мотивациях за чистую монету, он может использовать триангуляции общего типа, которые я обсуждал в главе III. Кроме того, он может воспользоваться техникой, применяемой, при непосредственном взаимодействии. Окружающие могут идентифицировать лжеца-дилетанта по языку его тела (или по его отсутствию), поскольку, сосредоточившись на том, что он должен сказать, он упускает из виду жесты, которыми обычно сопровождается спонтанная речь. Помимо этого, чтобы проверить заявленные мотивы, можно устроить для агента ловушку. Если историки не имеют возможности подстроить ловушку тем, кого они изучают, а социологам мешают это сделать этические соображения, то работодатель, супруг или родитель может быть не столь ограничен в средствах.

Вменение мотивов другим людям часто окрашено злобой. Если нужно выбрать между тем, была у альтруистического поступка соответствующая мотивация или он был основан на своекорыстии, мы часто выбираем последнее, даже если для этого нет серьезных оснований. Хотя подобное недоверие имеет смысл с точки зрения благоразумия (глава XXVI), во многих случаях подобное оправдание отсутствует. Сплетни, например, часто мотивированы тем, что французские моралисты вслед за блаженным Августином называли зловерностью или слабостью человеческой природы¹⁴. Согласно Ларошфуко, «если бы у нас не было недостатков, мы бы не испытывали такого удовольствия, видя недостатки в других». Еще он отмечал, что наше желание находить недостатки в других так сильно, что порой оно помогает их находить. «Наши враги в своем мнении о нас ближе к истине, чем мы сами». Но даже если наши враги ближе к истине, только с другого конца, они все равно заблуждаются, хотя и в меньшей степени. По шкале от 0 до 10, даже если у меня 6, я буду полагать, что у меня 9, враги же будут приписывать мне 4.

Для анализа такого отношения, часто называемого герменевтикой подозрения, я могу только процитировать Иеремию Бенета (в переводе с его неуклюжего французского).

Какую бы позицию ни занял король [Людовик XVI], чем бы он ни жертвовал, ему никогда не удастся заставить замолчать этих клеветников: они — паразиты, которые неизменно питаются дурным настроем и тщеславием здорового политического организма. Первым и главным источником этой несправедливости является тщеславие. Человеку ко всему хочется найти искусный подход... и он предпочтет

¹⁴ Я не согласен с теми, кто объясняет сплетни ролью, которую они играют для насаждения социальных норм. Верно, что сплетня может действовать как усилитель на неформальные санкции, которые поддерживают социальные нормы, но я полагаю, что их истоки лежат гораздо глубже.

самые хитроумные домыслы позорному предположению, что у общественного деятеля могут быть похвальные мотивы. Если Вашингтон отошел от политики, то только затем, чтобы через анархию проложить дорогу к деспотизму. Если Некер, вместо того чтобы как все принять плату за свои услуги, платит из своих средств за то, чтобы ему позволили их оказывать, это лишь изощренный способ утолить свою алчность. Если Людовик XVI отрекается от законодательной власти в пользу народа, это может быть только результатом хитроумного плана, рассчитанного на то, чтобы в благоприятный момент вернуть себе все обратно и даже получить больше.

Ирония в том, что последнее из этих показных обвинений, процитированных в тексте (написанном в 1789 году), возможно, оправдалось к осени 1790 года. Один из ближайших советников короля, Сен-При, писал, что к этому времени он перестал сопротивляться покушениям на законодательную власть, потому что «убедил себя в том, что Собрание дискредитирует самого себя своими ошибками». Теории заговора могут быть точны, потому что заговоры существуют. Однако склонность находить их связана не столько с опытом, сколько с болезненным нежеланием признавать, что общественные деятели могут действовать из добрых побуждений.

Библиографические примечания

Многое в этой главе позаимствовано из моей статьи «Альтруистические мотивации и альтруистическое поведение» (Altruistic motivations and altruistic behavior) в сборнике «Учебник по экономике дарения, взаимному обмену и альтруизму» (Kolm S. C., Ythier J. M. (eds). *Handbook on the Economics of Giving, Reciprocity and Altruism*. Amsterdam: Elsevier, 2006). Другие главы этой книги, в частности вводный очерк Колма, дают богатые эмпирические данные и теоретический анализ. Пример с французскими нищими взят из книги Ж. Лефевра «Великий страх 1789 года» (Lefebvre G. *La grande peur*. Paris: Armand Colin, 1988. P. 40), а пример с английскими крестьянскими бунтами — из Э. П. Томпсон «Моральная экономика английской толпы в XVIII столетии» (Thompson E. P. *The moral economy of the English crowd in the 18th century* // *Past and Present*. 1971. No. 80. P. 76–136). Анализ «теплого свечения» альтруизма представлен в статье Дж. Андреони «Нечистый альтруизм и пожертвования на общественное благо: теория дарения с эффектом теплого свечения» (Andreoni J. *Impure altruism and donations to public goods: A theory of warm-glow giving* // *Economic Journal*. 1990. No. 100. P. 464–77). Об отношении к донорству почеч см. Х. Лоренцен и Ф. Патерсон «Донорство живых: являются ли французы и норвежцы альтруистами?» (Lorenzen H., Paterson F. *Donations from the living: Are the French and*

Norwegians altruistic? // Elster J., Herpin N. (eds). *The Ethics of Medical Choice*. London: Pinter, 1994). Идею и слово *approbateness* (стремление к одобрению) я взял из А.О. Лавджой «*Размышления о человеческой природе*» (Lovejoy A. O. *Reflections on Human Nature*. Baltimore: Johns Hopkins Press, 1961). Роль бескорыстия во Французской революции обсуждается в Б. М. Шапиро «*Самопожертвование, своекорыстие или самозащита? Конституционное собрание и „самоотверженное распоряжение“ мая 1791*» (Shapiro B. M. Self-sacrifice, self-interest, or self-defense? The constituent assembly and the «selfdenying ordinance» of May 1791 // *French Historical Studies*. 2002. No. 25. P. 625–656). Касательно американской параллели см.: Дж. Вуд «*Корысть и бескорыстие при создании конституции*» (Wood G. Interest and disinterestedness in the making of the constitution // Beeman R., Botein S., Carter E. II (eds). *Beyond Confederation: Origins of the Constitution and American National Identity*. Chapel Hill: University of North Carolina Press, 1987). Касательно механизмов трансмутации см. мою книгу «*Алхимия разума*» (Alchemies of the Mind. Cambridge University Press, 1999. Ch. 5). Пример взаимного обмена времен Второй мировой войны я взял из книги Р. Аксельрод «*Эволюция сотрудничества*» (Axelrod R. *The Evolution of Cooperation*. New York: Basic Books, 1984). Касательно игры «Доверительное управление» см.: С. Камерер «*Поведенческая теория игр*» (Camerer C. *Behavioral Game Theory*. New York: Russell Sage, 2004. Ch. 2.7). Относительно выявления лжи см.: П. Экман «*Психология лжи*» (Экман П. *Психология лжи*. СПб.: Питер, 2007). Отрывок из Бентама взят из его труда «*Права, представительство и реформа*» (Rights, Representation, and Reform. Oxford University Press, 2002. P. 17–18).

VI. Близорукость и прозорливость

ЗА РАМКАМИ ПОСТЕПЕННОГО ВОСХОЖДЕНИЯ

Фрейдовский принцип удовольствия (глава IV) — стремление к немедленному удовлетворению желаний. Одним из его проявлений является склонность к убеждениям, которые хочется считать правдой, а не к тем, которые подтверждаются действительностью. Принятие желаемого за действительное в этом случае позволяет чувствовать себя комфортно здесь и сейчас, даже если в дальнейшем это приведет к неудаче. Еще одно его проявление возникает при выборе между двумя действиями, порождающими разнонаправленные потоки временной выгоды. Принцип удовольствия диктует выбор потока, предлагающего максимальную выгоду на первом этапе независимо от того, какую форму он примет на последующих.

Говоря более обобщенно, тот, кто принимает решения, будь то земляной червь или фирма, занят *постепенным восхождением* (gradient climbing). В каждый отдельный момент времени он изучает *ближайшие* возможности, чтобы увидеть, не принесет ли какая-то из них *большие непосредственные* дивиденды, чем нынешний статус-кво. Ограниченность ближайшими возможностями — это форма пространственной близорукости (spatial myopia): с глаз долой, из сердца вон. Ограниченность немедленными выгодами — форма темпоральной близорукости, принцип удовольствия. Земляной червь изучает окрестности, чтобы посмотреть, нет ли поблизости более влажного места, чем то, в котором он находится; обнаружив его, он переползает туда. Фирма изучает «пространство» рутинных процедур, близких к тому, чем она в настоящий момент занимается, чтобы найти нечто, что обеспечит основу для лучших краткосрочных показателей, и если находит, внедряет данные процедуры. Через некоторое время фирма или земляной червь может остановиться в месте, превосходящем (в краткосрочной перспективе) все остальные близлежащие позиции. Они достигли *локального максимума*.

Люди поступают умнее. Интенциональность, то есть способность представлять отсутствующее, дает нам возможность выйти за рамки принципа удовольствия и принять в расчет отдаленные по времени последствия

нынешнего выбора. Планирование позволяет нам делать выбор, результат которого будет лучшим в сравнении с тем, что могут дать сиюминутные и сиюсекундные решения. В некоторых случаях дальновидные поступки могут совершаться для лучшего удовлетворения текущих потребностей, подобно тому как алкоголик воздерживается от посещения соседнего ресторана, чтобы купить в отдаленном магазине целую бутылку по той же цене. В других случаях действия направлены на удовлетворение будущих нужд, как в случае когда я откладываю на старость. Если прозорливость первого типа наблюдается и у животных, то поведение второго типа, как принято считать, выходит за рамки их способностей. Некоторые недавние данные, однако, позволяют предположить, что приматы, возможно, способны к планированию на основе не актуальных, а ожидаемых потребностей. Но как бы то ни было, очевидно, что действия на основе спроецированных потребностей являются более сложной операцией.

Позвольте привести четыре примера действий, основывающихся на удаленных во времени последствиях. Первые три примера также обсуждаются в последующих главах.

Отступить, чтобы дальше прыгнуть (*reculer pour mieux sauter*). Это французское выражение, приблизительным аналогом которого является «Шаг назад, два шага вперед», иллюстрируется фундаментальным фактом экономической жизни: чтобы достичь более высокого уровня потребления в будущем, нужно меньше потреблять в настоящем. Агент занимает положение, уступающее статус-кво, потому что таково условие реализации более привлекательной альтернативы впоследствии. Излишне говорить, что это имеет смысл, если (1) более низкое положение позволяет ему выжить и (2) выигрыш от более высокого положения настолько велик, что оправдывает потери, связанные с переходом в более низкий статус.

Ожидание. Многие вина, хотя они уже неплохи, когда их разливают по бутылкам, со временем становятся еще лучше. Чтобы извлечь выгоду из этого факта, агент должен отказаться от опции (выпить вино сразу), которая превосходит статус-кво, поскольку отказ является условием достижения лучшего результата в будущем. В другой стороны, отложенное потребление не всегда имеет смысл, если, предположим, агент не уверен, что проживет долго, чтобы насладиться выдержанным вином. В качестве более веского примера можно привести выбор супруга. Вместо того чтобы делать или принимать предложение, как только подвернется приемлемый кандидат, можно подождать, пока не появится кто-то более подходящий. Риск, богато отраженный в мировой литературе, заключается в том, что никого лучше может и не появиться.

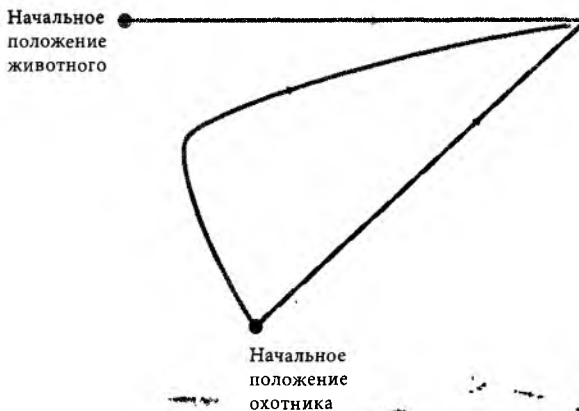


РИС. VI.1

Стрельба с опережением. Чтобы попасть в движущуюся мишень, нужно целиться не туда, где она находится, а туда, где она будет в момент соударения. Аналогичным образом, преследуя движущуюся мишень, нужно целиться по прямой туда, где мишень окажется, а не следовать по кривой, которая возникает при прицеливании в ее текущее положение.

На рис. VI.1 охотник, даже если он движется чуть медленнее, чем животное, может его нагнать, пойдя по прямой по направлению к точке, в которой оно окажется в некий заранее просчитанный момент. Если же он будет двигаться в направлении текущей позиции животного, следуя по кривой на графике, он никогда его не догонит. Как мы увидим (глава XVI), естественный отбор в меняющейся среде тоже может рассматриваться в этой перспективе.

Прямой путь не всегда самый быстрый. При попытке достичь стационарной цели прямая не всегда является самым эффективным маршрутом. На рис. VI.2 спасатель может импульсивно броситься к тонущему пловцу, прежде чем тот достигнет берега, а затем проплыть оставшееся расстояние. Если он остановился (но не надолго!), чтобы подумать, он сообразит, что может бежать быстрее, чем плыть, и что он быстрее доберется до пловца, пойдя по обходному пути, который хотя и окажется длиннее в целом, предполагает проведение меньшего времени в воде.

Так мы ведем себя, когда едем по платной автомагистрали, а не по дороге, которая на карте выглядит короче. В экономике планирование магистрального поведения часто является оптимальным.



РИС. VI.2

ДИСКОНТИРОВАНИЕ ВО ВРЕМЕНИ

Способность к долгосрочному планированию не предполагает, что обязательно будет востребована. Чтобы ожидаемые долгосрочные последствия имели значение для настоящего поведения, агенты должны быть *мотивированы* их учитывать. На языке психологов, они должны быть готовы *отложить удовлетворение желания*. На языке экономистов, они не должны подвергаться чрезмерному *дисконтированию во времени* (time discounting)¹. Необходимы и когнитивный, и мотивационный элементы. Если будущие результаты окутаны неизвестностью, они не могут мотивировать поведение в настоящем. Если они предполагают риск, то их мотивационная сила тоже уменьшается. Способность будущих ре-

¹ В этой книге фраза «высокая ставка дисконтирования во времени» (a high rate of time discounting) означает, что будущие вознаграждения имеют небольшую текущую ценность. Фраза «высокий коэффициент дисконта» (a high discount factor) означает значительную ценность в настоящем. Чтобы проиллюстрировать и обосновать эту внешне странную терминологию, предположим, что агенту все равно, получит он 3 единицы вознаграждения завтра или 2 единицы сегодня. Будущее вознаграждение дисконтируется (сокращается) на одну треть. Коэффициент дисконта (число, на которое мы должны умножить вознаграждение, чтобы получить его ценность в настоящем) — две трети.

зультатов формировать поведение в настоящем зависит как от *времени*, так и от *вероятности* их достижения. Механизмы, через которые они влияют на выбор, — это дисконтирование во времени и отношение к риску.

Как видно из этого выражения, дисконтирование во времени (или близорукость) — это склонность придавать вознаграждению в отдаленном будущем меньшее значение, чем вознаграждению в ближайшей перспективе и в настоящем². Если надо выбрать между 100 долларами сегодня и 110 долларами через год, большинство людей предпочтут первое. Однако у такого предпочтения может быть несколько источников.

Некоторые могут предпочесть скорейшее вознаграждение, так как смогут инвестировать эти средства и получить через год больше, чем 110 долларов.

Другие могут взять 100 долларов сейчас, потому что нуждаются в них, чтобы выжить. Получение большей суммы в будущем для них бессмысленно, потому что к тому времени они уже могут умереть. Предположим, что у меня есть выбор: поймать рыбу в реке сейчас руками или сделать сеть и наловить гораздо больше рыбы потом. Поскольку я не могу ловить рыбу, пока плету сеть, цена изготовления сети может быть так высока, что я не смогу себе ее позволить.

Еще кто-то может выбрать меньшее вознаграждение потому, что болен болезнью, при которой с 10%-й вероятностью через год он умрет. Планируя на будущее, нам приходится учитывать тот факт, что мы знаем, что умрем, но не знаем, когда.

Если ожидаемая в будущем сумма предполагает 50%-ю вероятность получения 130 долларов и 90 долларов, то из нежелания рисковать можно предпочесть верные 100 долларов сегодня.

Наконец, некоторые люди могут предпочесть скорейшее вознаграждение просто потому, что оно более быстрое. Это *чистое дисконтирование во времени*. Подобно тому как большой дом с дальнего расстояния кажется меньше, чем маленький дом вблизи, большая сумма в будущем может субъективно казаться меньшей, чем маленькая сумма в настоящем. В дальнейшем я буду рассматривать только этот случай.

Иррационально ли чистое дисконтирование во времени? Предположим, молодой человек делает очень большой дисконт с будущих вознаграждений. Вместо того чтобы отправиться в колледж, для чего придется временно пожертвовать доходом, чтобы получать более вы-

² Некоторые индивиды (например, патологические скупцы) могут придавать большее значение будущей выгоде, нежели настоящей. Для них время потребления не наступает никогда.

сокий заработок позднее, он сразу после школы поступает на мало-престижную работу с минимальными возможностями для карьерного продвижения. Поскольку он пренебрегает долгосрочным воздействием курения и еды с высоким содержанием холестерина, у него небольшой предположительный срок жизни³. Если он не уважает закон по моральным основаниям, соображения благоразумия не удержат его от нарушения этого закона. Другими словами, вполне вероятно, что его жизнь будет короткой и несчастной. Если это не иррациональное поведение, то что это?

По моему мнению, чистое дисконтирование во времени само по себе не является иррациональным. Оно может стать причиной того, что жизнь агента будет хуже, чем если бы он больше заботился о своем будущем, но это касается и эгоистических мотиваций. Тот, кто думает только о себе, может в результате прожить печальную и обедненную жизнь, но мы не должны из-за этого говорить, что эгоизм иррационален. Я обсуждаю эти вопросы в главе XI. В ней я сосредоточусь на том, как следует подходить к концептуализации дисконтирования во времени. Имеется несколько подходов с радикально расходящимися следствиями.

При моделировании дисконтирования во времени специалисты по теории принятия решений традиционно предполагают, что люди дисконтируют будущую прибыль экспоненциально (*exponentially*). Одна единица прибыли t периодов в будущем имеет настоящую ценность k_t , где $k < 1$ является коэффициентом дисконта на этот период. Экспоненциальное дисконтирование привлекательно, с нормативной точки зрения, тем, что позволяет осуществлять *последовательное планирование* (*consistent planning*). Если один поток вознаграждений имеет в данный момент времени большую ценность, чем другой, он будет иметь большую настоящую ценность во всех временных точках. Следовательно, у агента никогда не происходит инверсия предпочтений (*preference reversal*), которая обычно (в отсутствие причин для пересмотра решения) рассматривается как знак иррациональности.

Однако эмпирически понятие последовательного планирования не имеет особого смысла. Случайное наблюдение показывает, а систематическое подтверждает, что у большинства из нас часто возникает инверсия предпочтений. Очень часто мы не в состоянии последовать намерениям экономить, делать зарядку по утрам, заниматься на фортепьяно,

³ Пятьдесят лет назад многие люди игнорировали подобные последствия, то есть не осознавали их. Хотя сегодня это менее вероятно, они по-прежнему могут игнорировать их, то есть придавать им меньше значения при принятии решений. Нередко они могут оказаться в состоянии мотивированного игнорирования (форма принятия желаемого за действительное) последствий.

не опаздывать на встречи и так далее. Я могу позвонить дантисту 1 марта и записаться к нему на 1 апреля только затем, чтобы потом отменить запись 30 марта, сказав (солгав), что должен идти на похороны. Чтобы объяснить эти разновидности повседневной иррациональности (а также многие другие явления), мы можем заменить предположение об экспоненциальном дисконтировании *гиперболическим дисконтированием* (hyperbolic discounting).

Предположим, что дисконтированная настоящая ценность 1 единицы прибыли t периодов в будущем равна $1/(1+k_t)$ (в приведенном ниже примере я принимаю $k=1$, но в более общем случае k может быть любым положительным числом: чем оно больше, тем меньше агента заботит будущее). Более того, предположим, что агент при $t=0$ сталкивается с выбором: вознаграждение 10 при $t=5$ и вознаграждение 30 при $t=10$. При $t=0$ настоящая ценность первого составляет 1,67, а у последнего она равна 2,73. У агента, который максимизирует настоящую ценность, сформируется намерение выбрать отложенное вознаграждение. При $t=1$ настоящая ценность более быстрой награды равняется 2, более поздней — 3. При $t=2$ эти значения составляют соответственно 2,5 и 3,3; при $t=3$ они составляют 3,3 и 3,75; и при $t=4$ они равняются 5 и 4,29. В какой-то момент между $t=3$ и $t=4$ скорейшее вознаграждение перестает быть крайней опцией и становится более предпочтительной *только в результате того, что идет время*. Действительно, легко увидеть, что переключение происходит при $t=3,5$; именно тогда я звоню своему дантисту, чтобы отменить встречу.

Еще лучше эта модель видна на графике. На рис. VI.3 агент может выбрать или небольшое вознаграждение B при t_1 или дождаться t_2 и полу-

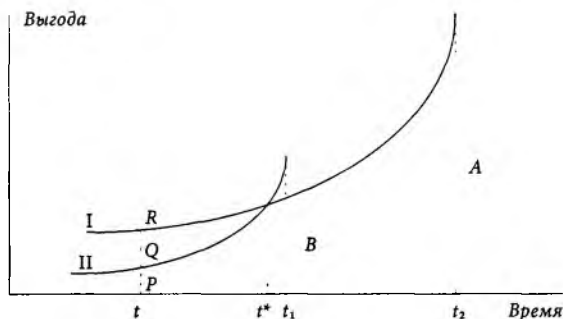


РИС. VI.3

чить большее вознаграждение А. Гиперболические кривые I и II представляют то, как оцениваются значения этих вознаграждений в разные предшествующие моменты. По сути они являются *кривыми безразличия* (глава IX), которые представляют компромиссы между временем получения вознаграждения и размером этого вознаграждения. К примеру, в момент t агенту все равно, получить вознаграждение PQ немедленно или получить небольшое вознаграждение на t_1 , а также ему все равно, получить PR немедленно или получить большую награду в момент t_2 . Поскольку во момент t настоящая ценность у А больше, чем у Б, она сформирует намерение выбрать А. Однако поскольку гиперболические кривые пересекаются в t^* , в этот момент возникает инверсия предпочтений и агент выбирает Б вместо этого⁴.

ПАРИ ПАСКАЛЯ

Мы можем использовать паскалевское пари, чтобы проиллюстрировать отношения между экспоненциальным и гиперболическим дисконтированием во времени. Паскаль хотел убедить своих друзей, свободомыслящих игроков, в том, что они должны поставить на Бога, поскольку даже самая малая вероятность вечного блаженства компенсирует величайшие земные удовольствия. В аргументе Паскаля заключено много

⁴ Есть альтернативный, слегка иной способ представления гиперболического дисконтирования. Он основан на интуитивной идее, что люди проводят радикальное различие между настоящим и всеми другими временами, придавая благополучию в настоящий момент большее значение, чем благополучию в последующие периоды. К тому же они различают эти последующие периоды. В примере с тремя периодами при написании u_i для благополучия, испытываемого в период i , настоящая ценность или дисконтированная сумма прибыли будет $u_i + b(d_1u_2 + d_2u_3)$. Здесь задействованы два множителя дисконта. В сравнении с настоящим все последующие прибыли, независимо от того, когда они будут получены, дисконтируются с множителем b . Кроме того, все будущие прибыли дисконтируются по экспоненте с множителями d . Настоящий момент инстинктивно наделяется значением, выделяясь на фоне всех остальных моментов, тогда как последующие периоды постепенно теряют мотивационную силу благодаря своего рода оптической иллюзии. Эту модель, называемую квазигиперболическим дисконтированием (*quasi-hyperbolic discounting*), объединяет с собственно гиперболическим дисконтированием то, что она тоже может вызывать инверсию предпочтений. Они различаются тем, что настоящая ценность бесконечного потока равных вознаграждений (как в паскалевском пари) имеет конечную сумму. Есть данные нейрофизиологии, свидетельствующие о том, что квазигиперболическое дисконтирование хотя и было введено как полезное приближение к гиперболическому дисконтированию, на самом деле является более точной репрезентацией.

сложностей, некоторые из которых мы рассмотрим в следующей главе. Здесь я хочу привлечь внимание к вопросу, который Паскаль не упоминает: имеет настоящая (дисконтированная) ценность вечного блаженства конечную или бесконечную ценность? Если она конечна, игрок может предпочесть получить свои удовольствия на земле, вместо того чтобы ждать загробной жизни.

Предположим для простоты, что каждый период загробной жизни дает 1 единицу прибыли; что человек, по его предположениям, умрет через n число лет от настоящего момента; и наконец, что он дисконтирует будущее благосостояние по экспоненте с множителем k ($0 < k < 1$). Если Бог дарует человеку спасение по вере его, настоящая ценность блаженства в первый год после его смерти составит k^n единиц прибыли, во второй год k^{n+1} и так далее. Если следовать элементарной алгебре, эта бесконечная сумма ($k^n + k^{n+1} + k^{n+2} + \dots$) складывается с конечной суммой $k^{n/1-k}$. Можно представить по крайней мере, что эта сумма может быть меньше настоящей ценности n лет гедонистической жизни на земле. Наоборот, если агент подвержен гиперболическому дисконтированию, бесконечная сумма $1/(n+1) + 1/(n+2) + 1/(n+3) + \dots$ увеличивается, выходя за пределы любой данной конечной ценности, предполагая, что если мы сравним настоящие ценности, любые земные удовольствия будут в конечном счете заслонены блаженством спасения. Даже если последнее будет умножено на сколь угодно малую вероятность того, что Бог существует, результат по-прежнему будет перерастать пределы любого конечного числа.

Предположим, однако, что собеседник Паскаля имеет возможность регулярно играть в азартные игры. Предварительно рассмотрев ситуацию, он предпочитает пойти к мессе, а не поиграть, потому что первая заставит его поверить и обещает ему вечное блаженство. Однако, по логике гиперболического дисконтирования, имманентная возможность играть в азартные игры вызывает инверсию предпочтений. У него появится намерение сыграть еще один раз, а затем начать ходить к мессе. Вслед за блаженным Августином он скажет: «Дай мне целомудрие и воздержание, но не сейчас». Однако на следующей неделе он рассудит таким же образом. Так сама структура дисконтирования во времени, обещающая более высокую настоящую ценность вечного блаженства, помешает игроку предпринять шаги для его достижения.

СЛАБОСТЬ ВОЛИ

Как показывает этот пример, гиперболическое дисконтирование может прояснить классическую проблему слабости воли. Слабовольный (или *акратический*) человек характеризуется следующим образом:

1. Он имеет основание сделать *X*.
2. Он имеет основание сделать *Y*.
3. Согласно его собственному суждению, у него есть более веское основание сделать *X*, чем *Y*.
4. Он делает *Y*.

Часто именно эмоции считаются способными предопределять действия агента, противоречащие его здравому суждению. Когда Медея в пьесе Еврипида собирается убить своих детей, она говорит: «Я ведаю, какое зло я совершаю. Но ярость моя сильнее последующих раздумий». По версии Овидия, она говорит: «Неведомая сила влечет меня вниз вопреки воле. Понукаемая то в одну, то в другую сторону... я вижу добро, но следуя по пути зла».

Эти признания, подобно четырем высказываниям, использованным для характеристики слабоволия, являются двусмысленными и неопределенными, поскольку в них отсутствует упоминание того, *когда*, как предполагается, они должны быть истинными. Давайте дадим *строгое определение слабоволия* следующим образом:

1. Человек имеет основание сделать *X*.
2. Человек имеет основание сделать *Y*.
3. Человек делает *Y*, полагая в момент совершения действия, что основания для совершения *X* более веские, чем основания для совершения *Y*.

Представим человека, который принял решение бросить курить и идет на вечеринку, где ему предложат сигарету. Он примет предложение, хотя знает, что ему не следует этого делать. Человек, сидящий на диете, может угоститься десертом, сознавая, что это лучшая идея. Хотя такая концепция слабости воли вполне возможна, она наталкивается на две эмпирические проблемы. Проще предположить, что здоровое рассуждение изменилось за доли секунды до совершения действия, чем считать, что действие и здоровое суждение сосуществует в одно и то же время. Кроме того, насколько я знаю, никто не установил причинно-следственный механизм, при помощи которого желание сделать *Y* приобретает большую побудительную силу, чем желание сделать *X*.

Чтобы обойти эти проблемы, мы можем дать широкое определение слабоволию, которое позволит развести во времени суждение агента о том, что ему нужно сделать *X*, и выбор *Y*:

1. Человек имеет основание сделать *X*.
2. Человек имеет основание сделать *Y*.

3. По спокойному и здравому рассуждению, основания для совершения X является более веским, чем основание для совершения Y .

4. Человек делает Y .

Сократ отрицал, что слабоволие в строгом смысле возможно. Аристотель тоже близко подошел к подобному утверждению. Он допускал слабоволие в более широком смысле, приводя пример человека, чья способность к суждению в момент совершения поступка находится под действием алкоголя. Предположим, я иду на вечеринку в офисе, слишком много пью, оскорбляю своего начальника и пристаю с домогательствами к его жене. В тот момент эти действия кажутся совершенно естественными. Но если бы кто-нибудь предположил до этого, что я способен на такое, я отверг бы подобные мысли как не совместимые с моим спокойным, здравым рассуждением. Если бы меня убедили в том, что мое здоровое суждение может раствориться в алкоголе, я остался бы дома. После произошедшего я могу горько сожалеть о своем поведении.

Пример, представленный на рис. VI.4, иллюстрирует случай *временного обращения предпочтений* (temporary preference reversal), а не слабоволия в строгом смысле. Есть как минимум три механизма, которые производят подобные изменения. Один — это *близость по времени* (temporal proximity), как это объяснялось при обсуждении гиперболического дисконтирования. Другой механизм — *пространственная близость* (spatial proximity), которую можно проиллюстрировать феноменом зависимости от сигнала (cue dependence). Этот механизм объясняет, например, многие случаи рецидива у наркоманов. Даже после многих лет воздержания поданный окружением сигнал, ассоциирующийся с употреблением наркотиков, может вызвать рецидив. Иногда для этого достаточно увидеть по телевизору наркоманские принадлежности. Решимость сесть на диету может быть подорвана видом провозимой мимо тележки с десертами. В этих случаях агент делает выбор того, что он предпочитает с учетом всех возможностей *в момент непосредственного выбора*. Наконец, *страсти* способны спровоцировать временное изменение предпочтений в силу своей быстротечности (глава VIII). Они также могут

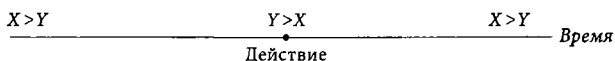


РИС. VI.4

вызвать инверсию предпочтений, заставив агента меньше задумываться об отдаленном будущем⁵.

Мы можем расширить эту идею, приняв в расчет временные (и мотивированные) изменения в убеждениях агента. В этой гораздо более широкой концепции слабование может быть также порождено самообманом (или принятием желаемого за действительное). Заранее решив, что на вечеринке он выпьет всего два бокала, чтобы не подвергать себя риску возвращаясь домой на машине, человек, побуждаемый желанием выпить третий бокал, может сказать себе, несмотря на веские контраргументы, что это никак не отразится на его водительских навыках⁶. Его предпочтения (безопасное вождение) остается неизменным, но изменились его представления об условиях, при которых вождение является безопасным. В то же время он может испытать временное изменение предпочтений, если решит, что хорошо провести время на вечеринке настолько важно, что это компенсирует (адекватно воспринимаемые) риски пьяного вождения.

Библиографические примечания

Касательно свидетельств способности приматов к планированию будущих (а не испытываемых в настоящий момент) потребностей см.: Н. Малкахи и Дж. Колл «Обезьяны сохраняют орудия для использования в будущем» (*Mulcahy N., Call J. Apes save tools for future use // Science. 2006. No. 312. P. 1038–1040*). Два источника по дисконтированию во времени и другим аспектам межвременного выбора — Дж. Левенстайн и Дж. Эльстер «Выбор во времени» (*Loewenstein G., Elster J. (eds). Choice over Time. New York: Russell Sage Foundation, 1992*) и Дж. Ле-

⁵ В действительности за вызванную гиперболическим временным предпочтением инверсию предпочтений может быть принято эмоционально мотивированное изменение множителя дисконтирования. Предположим, у агента есть выбор из двух опций, А и Б, которые дают вознаграждение соответственно (2, 5, 6) и (5, 4, 1) в течение трех последующих периодов. При ставке дисконта 0,8 для одного периода (а для двух периодов 0,64) настоящая ценность двух опций (из расчета на один период) соответственно равна 9,84 и 8,84. При ставке дисконта на один период 0,6 (а для двух периодов 0,36) ценности составят 7,16 и 7,96. Неудивительно, что агент перестает выбирать опцию с лучшими долгосрочными последствиями, когда из-за эмоций начинает меньше задумываться о будущем.

⁶ Наоборот, если он больше озабочен тем, что его остановит полиция, чем тем, что попадет в аварию, ему будет сложнее заставить себя поверить в то, что третий бокал не поспособствует увеличению алкоголя в крови. Как я утверждаю в следующей главе, реальность налагает (иногда) свои ограничения даже на принятие желаемого за действительное.

венштайн, Д. Рид и Р. Баумейстер *«Время и решение»* (Loewenstein G., Read D., Baumeister R. (eds). *Time and Decision*. New York: Russell Sage Foundation, 2003). Паскалевское пари я подробнее обсуждаю в работе *«Паскаль и теория принятия решений»* (Pascal and decision theory // Hammond N. (ed.). *The Cambridge Companion to Pascal*, Cambridge University Press, 2004). Нейрофизиологические данные о квазигиперболическом дисконтировании во времени приводятся в С. Макклур и др. *«Раздельные нейросистемы оценивают непосредственное и отложенное временное вознаграждение»* (McClure S. et al. *Separate neural systems evaluate immediate and delayed monetary rewards* // *Science*. 2004. No. 306. P. 503–507). Современное обсуждение слабоволия восходит к работе Д. Дэвидсона *«Как слабование возможно?»* (Davidson D. *How is weakness of the will possible?* // *Essays on Action and Events*. Oxford University Press, 1980). Я комментирую его идеи в статье *«Дэвидсон о слабоволии и самообмане»* (Davidson on weakness of will and self-deception // Hahn L. (ed.). *The Philosophy of Donald Davidson*. Chicago: Open Court, 1999). Мотивированное формирование убеждений обсуждается в книге Д. Пирса *«Мотивированная иррациональность»* (Pears D. *Motivated Irrationality*. Oxford University Press, 1984). Я более подробно обсуждаю связь между слабованием и инверсией предпочтений в статье *«Слабость воли и обращение предпочтений»* (Weakness of will and preference reversal // Elster J. et al. (eds). *Understanding Choice, Explaining Behavior: Essays in Honour of Ole-Jørgen Skog*. Oslo Academic Press, 2006).

VII. Вера

ЧТО ЗНАЧИТ ВЕРИТЬ ВО ЧТО-ТО?

Чтобы понимать роль верования (belief) в порождении действия, мы должны понимать его природу, причины и последствия. Как я отмечал во вступительных замечаниях ко второй части, не всегда ясно, что значит верить в то, что есть загробная жизнь. Многие великие религиозные деятели писали о постоянной борьбе с сомнением. Была ли их вера в те моменты, когда они действительно верили, такой же простой и безусловной, как вера того, кто никогда не знал сомнений? Действительно ли последователи коммунизма, верившие в то, что партия всегда права, верили в это?¹ Как мы можем различить прирожденного пессимиста, склонного верить в худшее, и благоразумного руководителя, который лишь действует так, как если бы был верен наихудший сценарий?

Кроме того, в повседневном языке понятие «верование» подразумевает нечто меньшее, чем полную поддержку. Я *верю*, что завтра пойдет дождь, но я также знаю, что могу ошибаться. Я не просто верю в то, что женат, я это *знаю*. При философском анализе знание обычно определяется как оправданное истинное верование, которое находится в особых отношениях как с миром (оно истинное), так и с корпусом информации, имеющимся у агента (оно оправданно). И все же ни одна из этих сторон знания не передает субъективной определенности, которая часто лежит за фразой «Я знаю» в повседневной речи. Эта определенность — не просто предел, то есть 97, 98, 99, 99,9%-я вероятность и так далее. Она качественно отличается от чего-то, в чем нет определенности².

¹ После падения коммунизма одна женщина из бывшей Восточной Германии сказала во время публичного выступления, что ее поколение было с детства воспитано так, чтобы соответствовать, следовать генеральной линии. Длительная шизофрения опустошила их как людей. Поэтому, по словам этой женщины, сейчас она не смогла бы вдруг начать говорить открыто или говорить то, что думает. Она даже точно не знает, что она на самом деле думает.

² В действительности все обстоит намного сложнее. Когда люди имеют дело с высокой числовой вероятностью, такой как 99,9%, могут сработать два разных механизма (в том смысле, в каком это слово используется в главе II): различие между неопределенностью и высокой вероятностью либо игнорируют, либо преувеличивают (подобным образом очень низкая вероятность

Эффект определенности продемонстрирован в следующем эксперименте. Одну группу испытуемых попросили выразить свое предпочтение одному из предложенных вариантов (цифры в скобках указывают пропорцию участников, которые предпочли данный вариант):

- 50%-я вероятность выиграть трехнедельную поездку по Англии, Франции и Италии (22%).
- Гарантированный недельный тур по Англии (78%).

Другой группе были предложены следующие варианты:

- 5%-й шанс выиграть трехнедельную поездку по Англии, Франции и Италии (67%).
- 10%-й шанс выиграть шестидневный тур по Англии (33%).

Члены первой группы склоняются к тому, чтобы предпочесть вариант «только Англия», потому что он *гарантирован*. Как только его ценность падает из-за такой же вероятности, как у его альтернативы, последняя начинает выглядеть более привлекательно. Солдаты, которых спрашивают, пойдут ли они добровольцами на очень опасное задание, могут иметь гораздо меньше сомнений, чем те, кого просят добровольно вызваться выполнять самоубийственное задание. Конечно, первые тоже могут пасть жертвой принятия желаемого за действительное («Со мной это не случится»), которое для вторых, однако, не действует.

ЧЕТЫРЕ КОГНИТИВНЫХ ОТНОШЕНИЯ

Даже если оставить в стороне эти проблемы, идея верования (убеждения) сохранит неопределенность. Мы можем различить четыре когнитивных отношения к миру в порядке убывания их силы. Первое — модус *уверенности* (определенности); второе — модус *риска*, в котором агенты определяют вероятность — будь она основана на частотности события в прошлом или на субъективном мнении — каждого из возможных взаимоисключающих и вместе с тем исчерпывающих исходов. Третье отношение — модус *неуверенности* (неопределенности), в котором люди знают все возможные взаимоисключающие и исчерпывающие исходы, но считают, что они не в состоянии определить их (решающие) вероятности³.

тоже либо игнорируется, либо преувеличивается). Однако это осложнение не возникает в примере с эффектом определенности, приведенном в тексте.

³ Они могут иметь возможность приписывать ординальные вероятности, то есть говорить, что один исход более вероятен, чем другой, хотя и не имея возможности сказать, почему он более вероятен.

Наконец, модус *незнания*, в котором и диапазон возможных исходов, и вероятность их возникновения неопределенны или известны не полностью. Говоря памятными словами бывшего министра обороны Дональда Рамсфелда, мы сталкиваемся не только с известными и неизвестными величинами, но и с «неизвестными неизвестными»⁴.

Я сосредоточусь на уверенности и риске не из-за того, что они всегда являются собственно когнитивными отношениями, но потому, что они самые распространенные. Даже когда у людей нет оснований для формирования *какого бы то ни было* отношения к вопросу, они все равно чувствуют непреодолимое желание сформировать собственное мнение, причем не определенное мнение (как в случае принятия желаемого за действительное), но *то или иное мнение*. Эта склонность до некоторой степени предопределена культурными факторами. Альберт Хиршман сказал, что большинство латиноамериканских культур «придают большую ценность *изначально твердым мнениям* фактически *по всем вопросам*». В таких обществах расписаться в незнании равносильно тому, чтобы признать поражение. Но на самом деле это явление распространено повсеместно. Монтень говорил: «Многие, или, говоря смелее, все злоупотребления в этом мире возникают оттого, что нас приучают бояться открыто признавать свое невежество и что мы якобы должны принимать все, что не в состоянии опровергнуть». Нетерпимость по отношению к неуверенности и незнанию происходит не от гордыни, а из универсального человеческого желания находить во всем логику и смысл. Разум страшится пустоты.

Особой разновидностью стремления искать смысл в окружающей действительности является присваивание статуса *инстанции* (agency) событиям, которые с той же или даже большей достоверностью могли произойти совершенно случайно. При старом режиме во Франции народ никогда не мог смириться с тем, что обязан своей нищетой только природе. При дороговизне зерна существовало всеобщее мнение, что спекулянты вздувают цены, даже если реальной причиной был неурожай. Иногда нехватку зерна объясняли желанием элиты умиротворить народ с голоду в рамках классовой борьбы. Согласно статье 22 Хартии ХАМАС, евреи «стоят за Французской революцией, коммунистической революцией и большинством революций, о которых мы слышали и слышим, здесь и там. С помощью денег они организовывали тайные общества, «Ротари клуб», «Лайонс клуб» и другие в разных частях мира в целях подрыва общества и во благо сионистских интересов. С помощью денег

⁴ Более сложные случаи тоже возможны. Я могу иметь возможность приписывать вероятности некоторым исходам, будучи даже не в состоянии определить остальные.

они могли контролировать империалистические страны и подстрекать их к колонизации новых стран, чтобы дать им возможность эксплуатировать их ресурсы и распространять в них коррупцию». Такой параноидальный или склонный к теории заговора тип сознания в целом не восприимчив к опровержениям, ибо его носители воспринимают недостаток фактов или даже их опровержение в качестве *подтверждения*, которое становится еще одним доказательством дьявольского хитроумия заговорщиков.

Эти порождающие ошибки механизмы тем или иным образом полагаются на *мотивацию*. Но ошибка может возникнуть и из-за *незнания*. Это утверждение кажется очевидным, но оно не лишено проницательности. Дарвин заметил, например, что невежество чаще, чем знание, порождает уверенность. Невежество вкупе с уверенностью — хороший рецепт для ошибки. И наоборот, по мере расширения круга света растет и окружающая его зона темноты, что внушает большую сдержанность. Эксперименты показывают, что некомпетентность не только становится причиной низких когнитивных показателей, но и вызывает неспособность признать нехватку компетенции. Некомпетентные люди — вдвойне инвалиды.

Но есть и другая, возможно, еще более опасная вещь — не столько невежество, сколько *малое* знание. Монтень писал: «Есть детское невежество, которое предшествует знанию, и другое — ученое невежество, которое за ним следует». Паскаль говорил то же самое, только гораздо пространнее. С получением большего количества информации уверенность человека сначала растет, но в конечном счете падает. Я предполагаю, что гипотеза Монтеня — Паскаля более достоверна, чем гипотеза Дарвина. Это не наш выросший *уровень* знаний, а просто наш *выросший* уровень заставляет разум обгонять самого себя и производить больше уверенности, чем позволяют факты.

СУБЪЕКТИВНАЯ ОЦЕНКА ВЕРОЯТНОСТИ

Вероятностные суждения могут исходить из наблюдения за объективной частотностью или же быть чисто субъективными оценками⁵. Когда агент может черпать данные из значительного числа наблюдений за сходными ситуациями, частотный метод может дать хорошие результаты. Если я планирую в следующем месяце устроить пикник на день

⁵ При более глубоком анализе первый (объективный) метод сводится ко второму (субъективному). Чтобы быть полезными, объективные данные нуждаются в субъективной интерпретации. Но для многих практических целей подобное разделение является ясным и полезным.

рождения и должен сформировать мнение о том, какая будет погода, лучшее, что я могу сделать, — это посмотреть статистику за этот день в прошлые годы. Но если я хочу понять, какая погода будет завтра, лучшим индикатором может послужить сегодняшняя погода. Однако это не единственный индикатор. Статистика прошлых лет может сказать мне, является солнечная погода в этот день редким или нормальным явлением. Если это редкое явление, то сегодняшняя солнечная погода теряет часть своей значимости в качестве указателя. Я могу посмотреть на барометр, чтобы узнать, поднимается или падает атмосферное давление, или взглянуть на вечернее небо, полет воробьев и так далее.

Интегрировать всю эту информацию в общее вероятностное суждение о завтрашней погоде — непростая задача. Большинству из нас это не слишком удастся. Зачастую проблема не в недостатке информации, а в ее избытке, в отсутствии формальной процедуры, позволившей сформировать на ее основе интегральное мнение. Однако некоторые люди умеют лучше других обобщать в виде комплексной оценки обширную и разрозненную информацию с разной степенью релевантности. Они обладают неувловимым, но фундаментальным качеством *суждения* (judgment). Оно бывает у успешных генералов, бизнесменов, политиков; потому они и успешны. Им должен обладать глава центрального банка, но для экономистов оно необязательно⁶. Максимум, на что способны все остальные, — это признать, что у нас его нет, и научиться не доверять интуиции. Например, я могу понять, что зачастую не доверяю людям по причинам, которые при близком знакомстве оказываются нерелевантными («Он похож на того хулигана, которого я знал в пятом классе»). То есть я могу научиться не доверять собственному недоверию⁷.

Мы, однако, склонны думать, что способностью к суждению обладают не только успешные генералы, политики и бизнесмены, но и специально обученные эксперты. В сложных случаях диагноза или прогноза, когда надо выявить психически больного индивида или предсказать,

⁶ Говоря о Алене Гринспэне (Нью-Йорк Таймс. 2005. 28 окт.), Пол Кругман отметил, что хотя тот склонен не доверять формальным моделям, у него есть «способность угадывать по отрывочным и порой противоречивым данным, куда подует экономический ветер».

⁷ Знать, что кто-то склонен к предвзятости, — это одно, другое — уметь корректировать эту предвзятость. Исследования показывают, что намеренные попытки непредвзятых суждений не обладают особой ценностью, поскольку можно легко попасть в ловушку недостаточной, ненужной или чрезмерной коррекции. Можно научиться не доверять собственному суждению, но труднее его усовершенствовать. Если бы такое было возможно, то этого бы и не потребовалось.

совершит ли новое преступление тот, кто подал прошение о досрочном освобождении, мы доверяем экспертам. Благодаря своему опыту эксперты восприимчивы к знакам, которые люди, не имеющие специальной подготовки, могут упустить или чье значение они могут не понять. Более того, когда показания расходятся, эксперты могут воспользоваться своим опытом, чтобы решить, на что именно в данном случае следует полагаться. Так мы по крайней мере представляем себе экспертов. Поскольку большинство из нас считают себя экспертами в той или иной области (хотя бы в предсказании поведения нашего босса, супруга или ребенка), мы сделали большие вложения в создание образа эксперта с превосходными когнитивными навыками.

К сожалению, это *абсолютно ложный образ*. Во многих исследованиях диагностические и прогностические способности экспертов сравнивали с результатами простейшей механической формулы, основанной на нескольких переменных. В сущности все сводится к сравнению объективных (частотных) и субъективных методов. Значения, присваиваемые той или иной переменной, выводятся при помощи статистических методов, позволяющих предсказать наиболее правдоподобный результат. Почти без исключений формула работает ничуть не хуже, а чаще даже лучше, чем эксперт⁸. В исследовании прогрессирующей дисфункции мозга, основанном на тестировании интеллектуальных способностей, формула, полученная на основании ряда историй болезни, а затем перенесенная на новый образец, правильно идентифицировала 83% новых случаев заболевания. Группы опытных и неопытных клиницистов идентифицировали соответственно 63 и 58%. Более того, эксперты часто расходились во мнениях. В другом исследовании психиатры с большим опытом работы, посмотрев одну и ту же беседу с пациентом, не могли прийти к единому мнению относительно диагноза, мотиваций или чувств пациента. Некоторые психотерапевты используют для постановки диагноза сомнительные чернильные пятна. Но похоже, что пациенты для них не меньшая загадка, чем эти чернильные пятна для самих пациентов.

НЕКОТОРЫЕ ОШИБКИ СТАТИСТИЧЕСКИХ ВЫВОДОВ

Эксперты ошибаются не меньше, чем простые люди, потому что игнорируют очевидные и не столь очевидные принципы статистического анализа. В одном исследовании участникам дали описание молодого человека с длинными волосами, любителя читать стихи, после чего

⁸ Это превосходство сохранится, даже если мы присвоим всем переменным равные значения.

их спросили, кем он, по их мнению, является — скрипачом или водителем грузовика. Большинство высказались скорее за скрипача, то есть проигнорировали *ключевой параметр* двух групп, а именно абсолютное количество индивидов в каждой из них. Водителей грузовиков в стране настолько больше, чем скрипачей (и водители грузовиков такие разные), что в действительности больше вероятность найти молодого любителя поэзии именно среди водителей грузовиков.

Еще один источник ошибок при формировании убеждений — *отклонение выборки* (selection bias). Пациенты в центрах диализа на удивление неохотно встают в очередь на пересадку почки. Одна из причин такой реакции в том, что все попадающиеся им пациенты с пересаженными почками — это те, кому операция не помогла, из-за чего им снова пришлось вернуться к диализу. На такого рода отклонение указывал Монтень, когда ссылаясь на Диагора, к которому «...показав в храме многочисленные дарственные приношения с изображением людей, спасшихся при кораблекрушении, обратились с вопросом: Ну вот, ты, который считаешь, что богам глубоко безразличны людские дела, что ты скажешь о стольких людях, спасенных милосердием? Пусть так, — ответил Диагор. — Но ведь тут нет изображений утонувших, а их несравненно больше». Точно так же психиатр, утверждающий, что ни один педофил не остановится по собственной воле, игнорирует тот факт, что он просто не сталкивался с теми, кто это сделал.

Израильские ВВС совершили менее очевидную ошибку при оценке относительной эффективности поощрения и наказания при подготовке пилотов. Заметив, что показатели летчиков улучшаются, когда их наказывают за плохие показатели, а не когда поощряют за хорошие, они пришли к выводу, что наказание более эффективно. При этом они не приняли во внимание феномен, известный как *возврат в среднее состояние* (regression to the mean). В любой серии событий, полностью или частично определяемых случаем, есть тенденция к тому, чтобы за крайним значением в один момент следовало менее крайнее в следующий момент. У высоких отцов рождаются дети более низкого роста, а у пилотов за плохими показателями следуют не такие плохие, независимо от поощрений и наказаний. Когда атлеты очень хорошо выступают в течение одного сезона и хуже — во время следующего, поклонники и тренеры часто говорят, что их испортил успех, тогда как на самом деле мы, возможно, наблюдаем всего лишь возврат в среднее состояние.

Ошибка игрока (gambler's fallacy) и ее (безымянная) противоположность дают еще один пример. Приобретение страховок от землетрясения резко повышается после землетрясения, а потом стабильно падает по мере того, как память о происшествии стирается. Как и игроки, совершающие ошибку, когда верят, что красное выпадет снова, если

оно выпадало несколько раз подряд, покупатели формируют свои убеждения, пользуясь *эвристикой наличности* (availability heuristic). Их суждение о вероятности события определяется легкостью, с которой его можно вспомнить, а недавние события вспоминаются с большей готовностью, чем более ранние события. Исчезновение эмоций с течением времени (глава VIII) также может играть свою роль. И наоборот, люди, живущие в областях, где часто бывают наводнения, считают, что вероятность наводнения в год $(n+1)$ меньше, если оно уже произошло в год n . Как игроки, совершающие ошибку, полагая, что если красное выпадало несколько раз подряд, вероятность того, что оно выпадет снова, мала, они формируют свои убеждения, полагаясь на *эвристику представленности* (representativeness heuristic). Они считают (или действуют так, будто считают), что короткая последовательность событий может характеризовать более длинную, частью которой является.

Люди зачастую не могут уловить связь между случайными процессами и распределением исходов. Во время Второй мировой войны многие лондонцы были убеждены, что немцы систематически бомбили только определенные зоны города, потому что бомбы падали группами (кластерами). Они не понимали базовый статистический принцип, согласно которому кластерность генерируют случайные процессы и что бомбы, падающие кучно, в большей степени указывает на умышленный выбор цели. Факт, неизменно поражающий тех, кто с ним раньше не сталкивался, — в группе, насчитывающей по крайней мере 23 человека, вероятность совпадения двух дней рождения (день и месяц) составляет более 50 %.

МАГИЧЕСКОЕ МЫШЛЕНИЕ

Далее мы рассмотрим различные формы *магического мышления* (magical thinking), то есть склонности верить в возможность оказания причинного воздействия на результаты, которые человеку неподвластны. Например, люди склонны возлагать большие надежды на монету, которую еще не бросали, чем на ту, которую уже бросали, однако не разглашали результат броска. У Пруста друг рассказчика Робер де Сен-Лу был подвержен «своего рода суеверной вере в то, что верность его любовницы может зависеть от его собственной верности ей». Кроме того, люди могут не улавливать различия между *диагностическим* и *причинным значениями*. В одном эксперименте участники, которым сказали, что то, как долго они смогут продержаться руки в холодной воде, служит наилучшим индикатором их продолжительности жизни, продержали руки в холодной воде дольше, чем те, кто не получил этой (ложной) ин-

формации⁹. Точно так же, используя собственное поведение для предсказания поведения других, люди могут выбрать в «дилемме заключенного» кооперативную стратегию, как будто они каким-то образом могут¹ заставить других сотрудничать. В одном эксперименте, когда сотрудничающих участников попросили предсказать выбор их партнера по интеракциям, а также того, кто не является их партнером и связан с другим человеком, они были склонны ожидать (и в большей степени были уверены в своих предсказаниях) сотрудничество со стороны скорее своего партнера по интеракциям, чем того, кто им не был¹⁰.

Пример такого рода магического мышления дает кальвинизм (глава III). Учитывая его веру в предопределение, может показаться, что у кальвиниста нет причин воздерживаться от всевозможных земных наслаждений, которые предположительно не могут оказать влияния на его судьбу после смерти. Макс Вебер утверждал, что кальвинизм заставлял своих последователей вести аскетический образ жизни не для того, чтобы заслужить спасение, но чтобы иметь субъективную уверенность в пребывании среди избранных. Мы можем истолковать его высказывание как указание на то, что кальвинисты путали причинное и диагностическое значения их поведения. Это хорошо видно в письме, которое распространяли английские баптисты в 1770 году: «Любая душа, проходящая к Христу... должна поощряться... Пришедшая душа не должна страшиться, что она не является избранной, ибо только избранные души захотят прийти». Если Бог избрал меня, он также *заставит меня стремиться* к определенному образу поведения.

Эти ошибки (и многие другие, которые подтверждаются многочисленными свидетельствами) являются по большей части «холодными», или немотивированными, ошибками, в некотором отношении сходными с оптическими иллюзиями. Другие, «горячие», ошибки возникают, потому что убеждения агента *мотивированы*, то есть подвержены влиянию желаний. Как мы увидим в главе XI, каузальное влияние желаний на убеждения по своей природе не является иррациональным. Желание

⁹ Разделение между причиной и симптомом не всегда очевидно. Уже в 1959 году великий статистик Р. А. Фишер, предполагая, что в основе предрасположенности к курению и раку лежит одна генетическая черта, утверждал, что курение может диагностировать рак, а не вызвать его. (Верно и то, что в то время ему платили табачные компании.) Или взять открытие, рассматривавшееся в главе II: чем дольше человек сидит без работы, тем меньше у него шансов найти ее в указанный промежуток времени. Длительность безработицы позволяет диагностировать способность получить работу или же может оказать каузальное влияние (через деморализацию и т. д.) на шансы ее найти.

¹⁰ Это расхождение позволяет нам исключить возможность приписывания кооперативного поведения партнеру по интеракции в результате эффекта ложного консенсуса (глава XXIII).

может дать основание для того, чтобы вкладывать определенное количество ресурсов в получение информации. Полученная информация может послужить основанием для того, чтобы воспринять определенные убеждения. Хотя желание не дает оснований для того, чтобы придерживаться убеждений, оно входит в рациональный комплекс их формирования. Клин между первоначальным желанием и финальным убеждением вбивает тот факт, что исход поиска информации по определению неизвестен в тот момент, когда принимается решение о его начале.

МОТИВИРОВАННОЕ ФОРМИРОВАНИЕ УБЕЖДЕНИЙ

Влияние желаний на убеждения, о котором я только что говорил, непротиворечивым образом согласуется с рациональностью. Более смелую идею предложил Паскаль в своем пари. Паскаль, как я показал в предыдущей главе, утверждал, что агент, верящий в то, что вероятность существования Бога не равна нулю, сколь бы мала она ни была, по чисто инструментальным причинам максимизации ожидаемого выигрыша должен попытаться приобрести твердую (в модусе уверенности) веру в то, что Бог существует, поскольку если это так, то она обеспечит ему вечное блаженство. У этого аргумента следующие предпосылки: (1) твердая вера наверняка ведет к спасению и (2) инструментальное происхождение этой веры не лишает ее эффективности в деле спасения. Хотя обе предпосылки могут показаться сомнительными с теологической точки зрения, нас это не должно смущать. Вопрос в том, является ли это решение верить рациональным проектом. В некотором смысле нет: я не могу принять волевое решение верить, подобно тому как я могу решить поднять руку. Можно, однако, использовать обходную стратегию. По утверждению Паскаля, действуя так, *как будто* он верит, человек в конце концов действительно приобретает веру. Однако механизм этого превращения сложно представить¹¹.

Есть и другие случаи, в которых у человека может возникнуть желание поверить в то, что он полагает ложным, ввиду благотворных последствий такой веры. Если я хочу сократить потребление алкоголя, но считаю, что риск стать алкоголиком не дает мне достаточной мотивации, у меня может возникнуть желание поверить в то, что риск больше, чем мне представляется. В общем и целом, однако, надежной технологии, обеспечивающей такую веру, не существует. Если только в этом процессе не задействован *самостирающийся компонент*, при помощи которого из сознания устраняется происхождение веры из желания верить, такое желание скорее всего лишь им и останется.

¹¹ По причинам, которые будут объяснены ниже, редукция диссонанса не является достоверным механизмом.

В непротиворечивом случае желание агента приводит к сбору определенного рода информации, которая в свою очередь вызовет то или иное убеждение; в противоречивом — желание вызывает специфическое поведение, порождающее в свою очередь специфическое убеждение, которое хочет приобрести агент. Обе стратегии являются обходными. Теперь я хочу обратиться к убеждениям, *напрямую* формируемым мотивацией. Это происходит одним из двух способов, соответствующих основным чертам мотиваций, — возбуждению и удовлетворению. Подобно тому как мы говорим, что камень разбивает лед благодаря своему весу, а не цвету, мы можем сказать, что мотивация влияет на убеждения не своим содержанием, а благодаря уровню возбуждения, который ее сопровождает. Умеренное физиологическое возбуждение может улучшить качество формирования убеждений через фокусирование внимания и стимулирование воображения. «Когда человек знает, что его через две недели повесят, — говорил доктор Джонсон, — он может прекрасно сосредоточиться». Однако за пределами определенного уровня возбуждения умственные способности ухудшаются. В состояниях крайнего истощения, стресса, страха или желания, связанного с аддикцией, трудно правильно думать, потому что возбуждение мешает держать в уме предыдущие ходы рассуждений. Предположительно умственная концентрация притупляется за день до повешения. В школьных тестах очень сильная мотивация получить высокий результат может привести к низкому результату так же, как у стрелка, который очень хочет попасть в цель, может дрогнуть рука, и он промахнется¹². В следующей главе я постараюсь доказать, что давление многочисленных эмоций может привести к тому, что формирование рациональных убеждений у агента пойдет в обход нормального механизма. Таким образом, убеждения могут быть *сформированы мотивацией* и при этом не быть *мотивированными*, потому что у агента нет желания верить в то, что они истинные. Возбуждение *затуманивает* разум, но не *склоняет* его в пользу какого-то определенного убеждения.

РАЦИОНАЛИЗАЦИЯ

Убеждения, вызванные содержанием, бывают двух основных разновидностей. Как я отмечал ранее, агент может быть мотивирован придерживаться *того или иного* убеждения по данному вопросу из-за потребности в завершенности или нежелания признавать свое невежество. И наоборот, он может быть мотивирован придерживаться *определенно-*

¹² И вам не поможет то, что вы сами будете пытаться расслабиться или вам прикажут расслабиться, потому что состояние расслабленности, подобно состоянию сна, является побочным эффектом.

го убеждения, например верить в то, что жена ему верна¹³ Самые важные механизмы, порождающие эту разновидность, — рационализация, принятие желаемого за действительное и самообман. Различие между первым и двумя последними механизмами лежит в отношении к поведению. При рационализации сначала возникает поведение, а убеждение следует за ним. (Это не значит, что после того как убеждения возникли, они не могут порождать поведение.) При принятии желаемого за действительное и самообмане мы наблюдаем обратную последовательность.

В качестве примера рационализации рассмотрим стандартный эксперимент с когнитивным диссонансом. Две группы испытуемых получили задание написать сочинение, предлагающее аргументы для позиций в споре о запрете абортов, которые они *не* поддерживают. Члены одной группы получают значительную сумму денег за участие, тогда как в другой группе участников просят сделать это только ради эксперимента. Написав сочинение, члены второй группы (но не члены первой) проявляют более благосклонное отношение к позиции, которую они защищали. Объяснение, возможно, заключается в том, что все участники хотят иметь *основание* для тех или иных действий. Члены первой группы могут просто сослаться на деньги как на основание¹⁴. Члены второй могут указать на (адаптированные) убеждения как основание своих утверждений¹⁵.

Французская пословица гласит: «Тот, кто обидел, никогда не простит». Если я несправедливо нанес вред другому человеку, то, возможно, не смогу признаться себе в том, что это моя вина. Вместо этого я буду искать недостатки в другом человеке, которые могли бы оправдать или по крайней мере извинить мое поведение. Насильники скажут: «Она была в провоцирующей одежде», — отговорка, иногда принимаемая судом. Участвующие в антисемитском насилии расскажут такую историю: поскольку евреи добились успеха аморальными или незаконными средствами, они заслуживают наказания. То, что именно поведение формирует убеждение, а не наоборот, очевидно из поразительной гибкости рационализации. Как отмечалось в главе III, риторика антисемитизма включает характеристику евреев как хищников и неполноценных людей и утверждения об их всемогуществе в истории.

¹³ Или, как в случае с Отелло, неверна. Такие контрмотивированные убеждения вкратце описывались в главе II и будут обсуждаться в главе XXIII.

¹⁴ Наказание за *отказ* участвовать тоже может быть достаточным основанием. Это может объяснить, почему при коммунистических режимах граждане имеют систему двойной бухгалтерии таким образом, что их внешний энтузиазм никогда не подрывает внутреннее отторжение.

¹⁵ В пари Паскаля заключен мощный стимул к тому, чтобы действовать так, как будто человек верит, — вечное блаженство, так что верующему не нужно искать других объяснений своего поведения.

ПРИНЯТИЕ ЖЕЛАЕМОГО ЗА ДЕЙСТВИТЕЛЬНОГО

Позвольте обратиться к принятию желаемого за действительное (*wishful thinking*) и к самообману. Эти два до конца непонятых феномена объединяет то, что желание, чтобы *P* стало реально, рождает веру в то, что *P* действительно реально. При принятии желаемого за действительное это простой одношаговый процесс; желание — мать мысли. Факты не столько отрицаются, сколько игнорируются. Сформированное в результате этого убеждение может оказаться тем же самым, которое может быть подтверждено фактами, если бы с ними сверялись¹⁶. Самообман в том виде, как его обычно представляют, включает четыре шага: рассматриваются свидетельства; во-вторых, формируется убеждение; в-третьих, это убеждение вытесняется или отвергается, потому что оно несовместимо с нашим желанием; и наконец, наше желание формирует на его месте другое, более приемлемое убеждение. Самообман — парадоксальное явление, существование или даже сама возможность которого ставилась под вопрос, поэтому позвольте начать с более простого случая — с принятия желаемого за действительное.

Прежде чем предложить описание механизма, который порождает принятие желаемого за действительное, позвольте отметить, что, в отличие от самообмана, существование этого явления невозможно отрицать. Можно оспаривать, что оно возникает в ситуации высокого риска или что оно влияет на сложные формы взаимодействия, например поведение на биржах или во время выборов, но не факт существования яв-

¹⁶ Игнорирование этого пункта может стать источником иррациональных верований. Поскольку, как правило, легко выявить работу мотивированного образования убеждений у других, мы склонны не доверять полученным таким образом выводам, не удосуживаясь проверить, не подтверждаются ли они фактами. До 1990 года я вместе со своими друзьями считал, что на шкале от 0 до 10 баллов (худшее) коммунизм должен получить где-то 7–8 баллов. После последних разоблачений я полагаю, что адекватная оценка — 10. Причиной моего ошибочного восприятия фактов была не идеалистическая вера в то, что коммунизм — это достойный идеал, который был предан самими коммунистами. В этом случае я был бы просто жертвой принятия желаемого за действительное или самообмана. Я был скорее сбит с толку излишней эмоциональностью тех, кто все это время утверждал, что коммунизм — это 10. Мое игнорирование их утверждений не было целиком иррациональным. Как правило, имеет смысл отбрасывать заявления откровенно истерического характера: истерик может быть прав, хотя и на ложных основаниях. Поскольку я считал (и продолжаю считать), что многие из этих яростных антикоммунистов говорили бы то же самое, независимо от реальных данных, я не мог поверить, что то, что они говорят, действительно соответствует фактам. Я совершил ошибку, посчитав их часами, отстающими на час, а не сломанными часами, которые дважды в день показывают правильное время.

ления, свидетельствами которого полнится мировая литература. Более того, многие убеждения, образовавшиеся в результате принятия желаемого за действительное, становятся предпосылкой к *действию* и, следовательно, представляют собой нечто большее, чем просто квазиубеждения. Многие курильщики, которым удалось убедить себя, что курение не опасно в целом и, в частности, для них, бросили или попытались бы бросить, будь их убеждения более рациональными¹⁷. Слишком самоуверенные индивиды, которые в своем самообольщении считают себя более способными, чем в действительности, могут совершать поступки, которых они в противном случае избегали бы. Люди, которые убедили себя в том, что они так же успешны, как и все остальные, могут потерять желание совершенствоваться дальше.

Распространен следующий механизм. Сначала человек мотивирован поверить в то, что он *успешен*. Затем он находит в своей жизни области, в которых дела у него действительно идут хорошо. Третьим шагом он преувеличивает важность этих областей, чтобы доказать себе, что успешен в целом. Наконец, он ослабляет свои усилия по достижению успеха в других областях жизни.

Чтобы двигаться по жизни, необходимо иметь адекватные убеждения. В то же время убеждения могут быть внутренне приятными или неприятными, то есть вызывать положительные или отрицательные эмоции. Если мне скажут, что у меня рак, я могу обратиться к врачам, но вера заставит меня ужасно себя чувствовать. На языке Фрейда, те, кто руководствуется принципом реальности, стремятся к адекватным убеждениям, тогда как те, кем движет принцип удовольствия, будут стремиться исключительно к приятным. Это разделение применимо только к убеждениям в их строгом смысле, но не к квазиубеждениям. Люди, у которых возникает нереалистичная вера в то, что они получают большой денежный приз за свои достижения, но которые, тем не менее, воздерживаются от трат до его получения, пребывают во власти безвредной версии принципа удовольствия. При более опасной его разновидности убеждения заставили бы их по-настоящему влезть в долги.

Формирование убеждений тоже может иметь свою цену. Если убеждения формируются на основе действительных выгод, то это может происходить ценой отказа от инструментальных выгод. Эта цена зависит как от сравнения исхода, который будет получен, если мотивированное убеждение было ложным, с тем, который имел бы место, будь оно истинным, так и от вероятности ошибочного убеждения. Я буду называть это *составляющей исхода* (outcome component) и *составляющей вероятности* (prob-

¹⁷ Как и в случае с алкоголем, для того чтобы бросить курить, потребовалась бы иррациональная вера в то, что курение *более* опасно, чем на самом деле.

ability component) в цене. И наоборот, за убеждения, принятые для достижения инструментальных выгод, возможно, придется расплачиваться отказом от действительных выгод. Я буду обсуждать случаи, в которых формирование рациональных убеждений будет иметь краткосрочные действительные издержки и долгосрочные инструментальные выгоды, тогда как мотивированная вера имеет краткосрочные действительные выгоды и долгосрочные инструментальные издержки. Тем самым я сравню объяснения формирования мотивированной веры экономистами и психологами, чтобы доказать, что мы должны использовать и те и другие.

Экономисты сосредоточили свое внимание на *издержках*. Некоторые, например, утверждали, что у рабочих формируется мотивированная вера в технику безопасности в зависимости от того, превышает ли выгода от этой веры *издержки* ее поддержания. Если психологические выгоды от подавления страха, вызываемого каким-то видом деятельности, превышают потери, связанные с возросшей вероятностью несчастного случая, рабочий будет верить, что этот вид деятельности безопасен. Как следствие агент будет бессознательно изучать факты, чтобы понять, может ли он позволить себе веру, поддерживающую такую мотивацию. Далее предполагается, что в формировании веры нет никаких ограничений, рабочий может верить во все что угодно, независимо от имеющейся у него информации. Я думаю, это неправильная модель, потому что в ней не только нет места ограничениям, но и неверно фиксируются издержки. Выгоду от мотивированной веры получают немедленно, а возможные издержки возникают позднее. Чтобы этот аргумент был принят, нам придется предположить, что человек способен достигать такого межвременного баланса бессознательно. Как я утверждал в начале этой главы, свидетельств в пользу этого нет.

Более правдоподобная идея, по моему мнению, заключается в том, что принятие желаемого за действительное порождается *знанием об итоговом коэффициенте издержек*. Когда ставки невысоки, агент может сформировать веру без учета вероятностной составляющей издержек, то есть без учета данных. Нет компромиссов, есть только двухшаговый процесс. Во-первых, агент рассматривает ставки. Если ставки низкие, то он принимает более приятное убеждение; если высокие, то изучает факты и при необходимости собирает дополнительные сведения. Понятие низких ставок будет, естественно, варьировать от человека к человеку. Можно будет только сказать, что для данного индивида при прочих равных условиях принятие желаемого за действительное более вероятно, когда ставки низкие.

Психологи фокусируются на *ограничениях*. У агента, начавшего курить, может сложиться убеждение, что курение неопасно, по крайней мере для него. При этом он, однако, может оказаться ограничен своими

прежними представлениями о вреде курения. Первый раз провалившись на экзамене, человек может себе сказать, что ему не повезло, однако когда это происходит четыре раза подряд, такая отговорка сработает с меньшей вероятностью, если он провалится и в шестой раз. Или возьмем пример с дорогими билетами на бродвейское шоу, который я приводил в главе I. Если я заплатил за билет 75 долларов, а шоу оказалось так себе, мои воспоминания о заплаченных деньгах еще слишком свежи, чтобы я мог легко себя обмануть и занижить цену. Учитывая неосознаваемую и многомерную природу эстетического восприятия, мне проще завысить свою оценку шоу. Точно так же, хотя есть свидетельство того, что более вероятные события считаются более желательными и что более желательные события воспринимаются как более вероятные, последний эффект ограничен в большей степени, чем первый.

В наглядном эксперименте испытуемые должны были участвовать в игре на знание истории с определенным человеком в качестве либо партнера, либо противника. Увидев, что этот человек показал отличный результат, те, для кого он должен был стать партнером по игре (и кто хотел, чтобы он играл лучше), оценили его знание истории выше, чем те, для кого он должен был стать противником по игре (и кто хотел, чтобы он играл хуже). В то же время участники были ограничены характером информации, которую они получили, поскольку даже те, кто ждал, что он станет его противником, оценивали его выше, чем в среднем. Ограниченность эксперимента в том, что он не предлагал участникам возможность *действовать* согласно своим убеждениям с потенциально дорогостоящими последствиями, которые может иметь недооценка противника. Исходя из того, что мы знаем, это просто квазиубеждения.

В приведенном примере принятие желаемого за действительное ограничено предшествующими фактуальными убеждениями. В других случаях оно может быть ограничено правдоподобными каузальными убеждениями. Принятие желаемого за действительное часто влечет рассказывание историй, при этом идея истории тесно связана с идеей механизма, которую я обсуждал в главе II.

Множество механизмов упрощают подбор той или иной истории, оправдывающей убеждение, которое хочется считать правдой. Я могу отмахнуться от неприятного слуха пословицей «Слухи всегда врут» и радостно принять приятный мне слух пословицей «Слухи редко лгут». Или, например, устраиваясь на место социального работника, я могу узнать, что для этой работы желательна эмоциональная стабильность. Если моя мать бросила работу, чтобы сидеть со мной, я могу подогревать веру в мою эмоциональную стабильность, говоря себе, что постоянное присутствие родителей дома хорошо сказывается на детях. Если же она осталась на работе и отправила меня в детский сад, я, наоборот, буду гово-

рить себе, что на детях хорошо отражается пребывание рядом с другими детьми и что их родители самореализовались за пределами семьи¹⁸. В случае поражения моей любимой футбольной команды я могу поддерживать веру в ее превосходство, если другая команда выиграла благодаря пустячному событию (если оно может быть так истолковано). «Если бы судья не отклонил мяч, боковой игрок получил бы пас в положении, из которого он мог бы забить». Если моя лошадь приходит второй, я могу поддерживать веру в мое умение делать ставки, говоря, что она почти выиграла. В еще более иррациональном примере принятия желаемого за действительное, если я поставил на 32, а вышло 33, я тоже могу сказать, что почти выиграл, хотя эти два номера далеко друг от друга на рулетке¹⁹.

Иногда, однако, под рукой не оказывается подходящей и правдоподобной истории. Предположим, человек ставит деньги на 24. Выпал номер 15, который лежит недалеко от 24 на колесе рулетки; таким образом, вера человека в свой талант игрока нашла подтверждение. Вероятно, он рассматривал бы как подтверждение и другие результаты, такие как 5, 10 и 33, потому что на колесе все они рядом. Кроме того, результаты 22, 23, 25 и 26 он тоже счел бы подтверждениями, потому что они ближе по числовому значению, а 20, 21, 26 и 27 — потому что они рядом на игровом столе. То есть 13 из 37 его результатов могут рассматриваться как подтверждение его таланта к игре. Но это также означает, что для 24 результатов нет подходящей истории. Если выпадет такой результат, даже человек, предрасположенный к принятию желаемого за действительное и сильно мотивированный принять определенное убеждение, будет вынужден посмотреть фактам в лицо.

САМООБМАН

Затронем теперь скользкую тему самообмана. В повседневной жизни люди обманываются по поводу таких вещей, как вес, здоровье, привычка к выпивке, привычка тянуть время или верность собственных жен. В одном типичном сценарии они получают информацию, что что-то не так, и отказываются идти дальше и делать более определенный вывод. Глядя на себя в зеркало, я вижу, что у меня лишний вес, но не могу точно сказать, каков он. Отказываясь вставать на весы, я говорю себе, что это, возможно, всего несколько фунтов, которые я могу сбросить в любое время, когда

¹⁸ На самом деле никаких особых различий в дальнейшем развитии детей из этих двух типов семей не отмечается.

¹⁹ В то же время людей может еще больше расстраивать факт близости их номера к выигравшему. Некоторые национальные лотереи предлагают скромные утешительные призы для тех, кто почти победил.

захочу. Женщина чувствует уплотнение в груди, но отказывается пойти к врачу, чтобы выяснить, злокачественная опухоль или доброкачественная. В таких случаях самообман облегчается недостатком точного знания. Женщина не просто сначала приходит к выводу, что у нее рак, а потом подавляет эту веру. Она скорее подозревает, что у нее *может* быть рак.

Случай, который я буду рассматривать как парадигматический, характеризуется следующими особенностями:

1. Первоначальное подозрение о раке принимает форму веры с низкой вероятностью.
2. Ее сопровождает твердая уверенность в том, что если у нее действительно рак и она ничего не будет с этим делать, исход почти наверняка будет фатальным.
3. Ее также сопровождает твердая уверенность в том, что если у нее рак и она предпримет какие-нибудь меры, исход все равно может оказаться фатальным; и даже если он таковым не будет, лечение будет крайне неприятным.
4. Женщина никогда не задает себе вопрос, компенсирует ли боль, которую может принести лечение (в сравнении с болью от запущенной болезни), дифференциальный риск смерти.
5. Вместо этого она просто не идет к врачу, чтобы выяснить, есть ли у нее рак.

В этом примере ключевые черты (1) и (3). Поскольку начальная вера имеет небольшую вероятность, цена за ее изменение невелика. Женщина может легко переключиться на многочисленные истории о безвредных утолщениях и ненужных страхах. *И все же в отсутствие черты (3) у нее нет мотивации изменять свои убеждения.* Если бы она знала, что существует безболезненное и дешевое лечение, которое ее исцелит, у нее не было бы мотивации избегать врачей. Неизвестна ни одна форма иррациональности, которая поощряла бы тенденцию к блокированию *беззатратного* избавления от маловероятных бедствий, например, с помощью перевода низкой субъективной вероятности в нулевую.

Понимаемый таким образом самообман не включает одновременное присутствие двух взаимоисключающих убеждений, одно из которых сознательное, другое — бессознательное. Когда первоначальная оценка вероятности замещается следующей, она исчезает *насовсем*, а не перемещается в бессознательное. Многие авторы считают это противоречие главной чертой самообмана. Но дело так обстоит далеко не всегда. Обманывающий себя человек не похож (или необязательно похож) на того, кто ненавидит, когда ему на глаза попадают кошки, но понимает, что, чтобы их игнорировать, он должен сначала их заметить. Его можно сравнить с тем, кто видит в темноте тень, которая может быть кошкой,

но может оказаться и чем-то еще. Его отвращение к кошкам может быть удовлетворено тем, что он истолкует эту тень именно так и не станет подходить ближе, чтобы проверить свою правоту.

Я сомневаюсь, что психология или философия может сравниться с литературой в демонстрации самообмана. В романе «По направлению к Свану» Пруст описывает, как Сван размышляет о том «времени, когда люди описывали ему Одетту [его возлюбленную] как содержанку» и каким нелепым казалось это описание применительно к той, которую он знал. Следуя по цепочке ассоциаций, он подумал о своем банкире и напомнил себе, что нужно снять немного денег, чтобы помочь ей справиться с материальными трудностями.

И тут он вдруг задал себе вопрос: а не входит ли все это в понятие содержать женщину <...> и не может ли все-таки быть применимо к Одетте <...> совершенно, на его взгляд, не подходящее к ней слово *содержанка*? Он не смог углубиться в эту мысль: приступ умственной лени, которая была у него врожденной, перемежающейся и роковой, погасил в его сознании свет с такой быстротой, с какой в наше время, когда всюду проведено электрическое освещение, можно выключить электричество во всем доме. Несколько секунд мысль Свана шла ощупью, потом он снял очки, протер стекла, провел рукой по глазам и наконец увидел свет только после того, как наткнулся на совсем другую мысль, а именно — что ему в следующем месяце нужно будет послать Одетте не пять, а шесть или даже семь тысяч франков — чтобы сделать ей сюрприз и чтобы порадовать ее.

Библиографические примечания

Сведения о многих открытиях, о которых здесь сообщается, можно найти в следующих книгах: сборник *«Принятие решений в неопределенности: Правила и предубеждения»* под ред. Д. Канемана, П. Словика и А. Тверски (Канеман Д., Словик П., Тверски А. Принятие решений в неопределенности: Правила и предубеждения. М.: Генезис, 2005); *«Принятие решений»* под ред. Д. Белла, Г. Райфа и А. Тверски (Bell D., Raiffa H., Tversky A. (eds). Decision Making. Cambridge University Press, 1988), *«Суждение и принятие решений»* под ред. Т. Коннолли, Х. Аркса и К. Р. Хаммонда (Connolly T., Arkes H., Hammond K. R. (eds). Judgment and Decision Making. Cambridge University Press, 2000), *«Рациональный выбор, ценности и фреймы»* под ред. Д. Канемана и А. Тверски (Канеман Д., Тверски А. Рациональный выбор ценности и фреймы // Психологический журнал. Т. 24. 2003. № 4. Июль-август), *«Эвристика и предвзятость: психология интуитивного суждения»* под ред. Т. Гиловича, Д. Гриффина и Д. Канемана (Gilovich T., Griffin D., Kahneman D. (eds). Heuristics and Biases: The Psychology of Intuitive Judgment. Cambridge University Press, 2002), *«Новое в поведенческой экономике»* под ред.

К. Камерера, Дж. Лёвенштейна и М. Рабина (Camerer C., Loewenstein G., Rabin M. (eds). *Advances in Behavioral Economics*. New York: Russell Sage, 2004), «Психология экономических решений» под ред. И. Брокаса и Х. Карилло (Brocas I., Carillo J. (eds). *The Psychology of Economic Decisions*. Vols. 1, 2. Oxford University Press, 2003, 2004). Преобладающее стремление приписывать партнерам кооперацию засвидетельствована в статье Л. Мессе и Дж. Сивачека «Предсказания ответов других в играх со смешанными мотивами: самооправдание или ложный консенсус?» (Messé L., Sivacek J. «Predictions of others» responses in a mixed-motive game: Self-justification or false consensus? // *Journal of Personality and Social Psychology*. 1979. No. 37. P. 602–607). Факты двойной некомпетентности невежд приводятся в статье Дж. Крутера и Д. Даннинга «Неквалифицированные и не подозревающие об этом» (Kruger J., Dunning D. Unskilled and unaware of it // *Journal of Personality and Social Psychology*. 1999. No. 77. P. 1121–1134). Тонкий анализ ненадежности суждений некоторых экспертов («ежей») и более надежных суждений остальных («лисиц») приводятся в книге Ф. Тетлока «Экспертные политические суждения» (Tetlock P. *Expert Political Judgment*. Princeton, NJ: Princeton University Press, 2005). Данные по землетрясениям и наводнениям можно найти у П. Словича в «Восприятии риска» (Slovic P. *The Perception of Risk*. Sterling, VA: Earthscan, 2000). Касательно (не) логичности теорий заговора см. статью Б. Кили «О теориях заговора» (Keeley B. Of conspiracy theories // *Journal of Philosophy*. 1999. No. 96. P. 109–126). О теориях голода см. работу С. Каплана «Представления о заговоре как причине голода во Франции 18-го века» (Kaplan S. *The famine plot persuasion in eighteenth-century France* // *Transactions of the American Philosophical Society*. 1982. No. 72.) и Ф. Плу «Из уст в уста: Рождение и распространение слухов во Франции 19-го века» (Transactions of the American Philosophical Society. 1982. No. 72; Ploux F. *De bouche à oreille: Naissance et propagation des rumeurs dans la France du XIXe siècle*. Paris: Aubier, 2003). Исследование мышления, склонного верить в теорию заговора, приводится в книге Р. Хофstadтера «Параноидальный стиль в американской политике» (Hofstadter R. *The Paranoid Style in American Politics*. Cambridge, MA: Harvard University Press, 1964). Его роль на Ближнем Востоке является темой книги Д. Пайпса «Скрытая рука: Ближневосточный страх заговоров» (Pipes D. *The Hidden Hand: Middle East Fears of Conspiracy*. New York: St. Martin's Press, 1998). Мои комментарии, касающиеся того, как два экономиста (Джордж Акерлоф и Мэттью Рабин) и один психолог (Зива Кунда) рассматривают мотивированное формирование убеждений, получили дальнейшее развитие в работе «Цены и ограничения в экономике разума» (Costs and constraints in the economy of the mind // Brocas I., Carillo J. (eds). *The Psychology of Economic Decisions*. Vol. 2. Oxford University Press, 2004). Лучшее изложение проблемы самообмана можно найти в специальном выпуске журнала «Бихевиористские науки и науки о мозге» (*Behavioral and Brain Sciences*. 1997. No. 20, организованном вокруг статьи А. Меле «Настоящий самообман» (A. Mele. *Real self-deception*)).

VIII. Эмоции

РОЛЬ ЭМОЦИЙ

Эмоции входят в человеческую жизнь тремя способами. На пике интенсивности они являются самыми важными *источниками счастья и несчастья*, значительно превосходящими гедонистические удовольствия или физическую боль. Светлая любовь Энн Эллиот в «Доводах рассудка» — необыкновенное счастье. На другом полюсе — эмоция стыда, которая может быть чрезвычайно разрушительной. Вольтер писал: «Быть объектом презрения для тех, с кем живешь рядом, — этого никто не мог и не сможет вынести».

Стыд также иллюстрирует второй путь, которым эмоции на нас влияют, а именно их *влияние на поведение*. В главе IV я упоминал несколько случаев, когда люди кончали с собой из-за того, что не могли справиться с чувством стыда. В этой главе я буду преимущественно рассматривать *тенденции к действию* (action tendencies), которые ассоциируются с эмоциями. То, в какой степени эти побуждения отражаются в поведении, будет интересовать нас в следующих главах.

Наконец, эмоции имеют значение благодаря их влиянию на *другие ментальные состояния*, а именно на убеждения. Когда желание добиться какого-то состояния поддерживается сильной эмоцией, стремление поверить в то, что оно действительно достигнуто, может оказаться непреодолимым. Стендаль в книге «О любви» писал: «С той минуты, как он влюбляется, даже самый мудрый человек больше ничего не видит в *его настоящем свете*... Он больше не допускает элемент случайности и теряет чувство вероятности, судя обо всем по тому влиянию, какое оно оказывает на его счастье, все, что он воображает, становится реальностью». В прустовском цикле «В поисках утраченного времени» эта тема развита на сотнях страниц, с большим количеством вариаций и сюжетных поворотов, чем можно себе вообразить.

ЧТО ТАКОЕ ЭМОЦИИ?

Прежде чем рассмотреть каждый затронутый аспект более подробно, я должен сказать что-нибудь о том, *что такое эмоции и какими они бывают*.

Общепризнанного определения того, что считать эмоцией как некое согласованного списка необходимых и достаточных условий, не существует. Хотя я коснусь большого количества общих черт тех состояний, которые мы в предварительном анализе считаем эмоциями, для каждой из них найдется контрпример. К любой такой черте можно подобрать эмоциональные состояния, в которых она отсутствует. Мы можем полагать, что тенденция к действию имеет ключевое значение для эмоции, но в качестве опровержения можно сослаться на эстетическую эмоцию. Нам может казаться, что существенной чертой эмоции является короткая жизнь, то есть свойство быстро распадаться, но в некоторых случаях безответная романтическая любовь (как у Сирано де Бержерака) или страстное желание отомстить может сохраняться годами. Уместно предположить, что эмоции вызываются убеждениями, но тогда как объяснить, что люди испытывают волнение, читая книги и просматривая фильмы, вымышленность которых очевидна? Можно привести много других якобы универсальных черт, которые в некоторых случаях оказываются отсутствующими¹.

В свете этих вопросов естественным ответом становится отрицание научной продуктивности самой категории «эмоция». Говоря языком философов, эмоции, кажется, не образуют *естественный вид* (natural kind). Несмотря на различия, киты и летучие мыши, будучи млекопитающими, принадлежат к одному и тому же естественному виду. Киты и акулы, вопреки их сходству, к нему не относятся, так же как летучие мыши и птицы. Гнев и любовь объединяет способность затуманивать и искажать сознание, но это сходство не делает их естественным видом. Чтобы показать, как такие аналогии могут сбивать с толку, можно заметить, что прием амфетаминов и романтическая любовь производят много сходных эффектов: обостренное сознание, прилив энергии, снижение потребностей во сне и в еде, чувство эйфории. Насколько я понимаю, никто не возьмет за утверждать, что они принадлежат к одному естественному виду².

В рамках социальных наук эту головоломку можно оставить без разгадки. Мы можем сосредоточиться на эмоциональных состояниях, у которых есть ряд регулярно наблюдаемых черт, и задаться вопросом: как это может помочь нам объяснить поведение или другие ментальные состояния? Тот факт, что в других состояниях, которые интуитивно относят к эмоциям, некоторые из этих черт отсутствуют, интересен с концептуальной точки зрения, но не лишает эту идею ее объяснительной

¹ В целях диагностики можно постулировать, что ментальное состояние является эмоцией, если обладает (скажем) восемью или более из двенадцати черт. Однако для объяснения этого недостаточно.

² Хотя они могут задействовать иногда одни и те же нейронные контуры.

способности там, где эти черты обнаруживаются. Я хотел бы обратить внимание следующие из них:

- *Когнитивные antecedentes.* Эмоции провоцируются (часто вновь приобретенными) убеждениями агента. Эмоции могут иметь и другие каузальные условия (мы легко раздражаемся, когда устали), но их наличие само по себе вызывает появление эмоции не в большей мере, чем скользкая дорога порождает аварию.
- *Физиологическое возбуждение.* Эмоции сопровождаются изменением сердцебиения, электропроводности кожи, температуры тела, кровяного давления, дыхания и многих других параметров.
- *Физиологическое выражение.* Эмоции сопровождаются характерными знаками, такими как поза, тембр голоса, приливы крови к лицу (от смущения), улыбка или обнажение зубов, смех и насупленность, плач и рыдания, бледность или краснота в приступе гнева.
- *Тенденции к действию.* Эмоции сопровождаются тенденциями или побуждениями к выполнению специфических действий. Хотя такие тенденции могут и не вызывать реальное поведение, они нечто большее, чем просто возможность; это формы зарождающегося поведения, а не просто их потенция.
- *Интенциональные объекты.* В отличие от других внутренних ощущений, таких как боль или голод, эмоции всегда испытываются *по отношению* к чему-то. У них могут быть пропозициональные объекты («Меня огорчает, что нечто...») или не-пропозициональные объекты («Меня огорчает нечто»).
- *Валентность (valence).* Это технический термин для характеристики испытываемых нами эмоций по шкале *боль — удовольствие*. Как уже отмечалось, валентность может варьировать от сияющего счастья Энн Эллиот до невыносимого стыда, который испытывали те, кто был публично уличен в потреблении детской порнографии.

Нет ли у эмоций, подобно цветам, присущих им характерных *чувств*? Стыд и вина, например, вызывают разные чувства, различие между которыми не сводится к большей остроте дискомфорта, вызываемого стыдом. Есть свидетельства того, что в мой мозг можно вживить электрод и заставить испытывать печаль, смущение или страх, даже если я буду не в состоянии идентифицировать причину или объект этих чувств. Каким бы важным ни оказался этот аспект для нашего понимания эмоций, он еще недостаточно понятен, чтобы выдвинуть конкретные каузальные гипотезы.

КАКИЕ СУЩЕСТВУЮТ ЭМОЦИИ?

Я перечислю и дам краткое описание примерно двух десятков эмоций, не претендуя на то, что эта классификация превосходит многочисленные другие классификации. Моя цель — дать некоторое понимание эмоций, которые имеют непосредственное или каузальное значение для социальной жизни, а не удовлетворить (законную) обеспокоенность теоретиков эмоций. В частности, мне нечего будет сказать по поводу того, какие эмоции являются базовыми, а какие небазовыми.

Одна из важных групп эмоций — *оценочные* (evaluative emotions). Они дают положительную или отрицательную оценку своего или чужого поведения или характера³. Если эмоция вызывается поведением другого человека, это поведение может быть обращено на агента либо на третьих лиц. Эти различия дают нам десять (или одиннадцать) эмоций:

- *Стыд* порождается негативным представлением о собственном характере.
- *Презрение и ненависть* вызываются негативными представлениями о характере другого человека. Презрение вытекает из представления, что другой человек хуже тебя; ненависть — убеждением в его злонамеренности.
- *Вина* вызывается негативным представлением о своем поведении.
- *Гнев* порождается негативным представлением о действиях других в отношении агента.
- *Картезианское негодование*⁴ вызывается негативным представлением о поступке другого человека в отношении третьих лиц.
- *Высокомерие* порождается положительным представлением о собственном характере.
- *Симпатия* следует из положительного представления о характере другого человека.
- *Гордость* вызывается положительным представлением о своем действии.
- *Благодарность* порождается положительным представлением о действии другого человека в отношении агента.

³ Эмоции, порожденные собственной негативной оценкой, всегда имеют негативную валентность. Вызванные негативной оценкой других более неопределенны в этом отношении.

⁴ Эта эмоция была впервые выявлена Декартом с существенным добавлением с его стороны: когда агент *любит* третью сторону, реакцией будет гнев, а не негодование.

- *Восхищение* вызывается положительным представлением о действии другого человека по отношению к третьей стороне.

Вторая группа эмоций связана с представлениями о том, что кто-то заслуженно или незаслуженно владеет чем-то хорошим или плохим⁵. Объект этих эмоций — не индивидуальное действие или индивидуальный характер, но положение дел. Пойдя по стопам Аристотеля и его «Риторики», мы можем выделить шесть (или семь) случаев:

- *Зависть* вызывается заслуженным благом другого.
- *Аристотелевское негодование* вызывается незаслуженным благом другого⁶. Тесно связанная с ним эмоция *обиды* (resentment) вызывается инверсией иерархии престижа, когда группы или индивиды, ранее занимавшие более низкое положение, добиваются господства.
- *Сопереживание* вызывается заслуженным благом, которым обладает другой человек.
- *Жалость* вызывается незаслуженным злом, причиненным другому.
- *Злоба* вызывается незаслуженным злом, причиненным другому.
- *Злорадство* вызывается заслуженным злом, причиненным другому.

Третья группа позитивных и негативных эмоций связана с мыслями о хороших или плохих вещах, которые случились или случатся с человеком, — *радость* (joy) и *печаль* (grief), с их несколькими разновидностями и родственниками. Как многие неоднократно отмечали, плохие события в прошлом также могут порождать положительные эмоции в настоящем, а хорошие события — негативные эмоции. Так, в основном корпусе пословиц античного мира, «Сентенциях» Публия Сира, мы можем найти как «Воспоминания о прошлых опасностях приятны», так и «Прошрое счастье преумножает настоящее несчастье».

Все до сих пор обсуждавшиеся эмоции вызываются убеждениями, которые существуют (или *возможно* существуют) в модусе уверенности. Но есть также эмоции — *надежда*, *страх*, *любовь* или *ревность*, которые

⁵ Я включаю «незаслуженное» в рамки «заслуженного». Таким образом, когда кто-то выигрывает большой приз в лотерею, я скажу, что это заслуженно, что несколько противоречит распространенным взглядам.

⁶ Хотя термин, которым Аристотель называл эту эмоцию, традиционно переводится как «негодование», очевидно, насколько оно отличается от картезианского негодования.

по сути своей связаны с убеждениями, существующими в модусе вероятности, или потенциальности. Эти эмоции порождаются мыслями о хороших и плохих вещах, которые могут случиться или не случиться в будущем; о хороших или плохих ситуациях, которые могут сложиться или не сложиться в настоящем⁷. В общем и целом эти эмоции требуют того, чтобы соответствующее событие (или ситуация) рассматривалось как более чем представимое, то есть должен быть серьезный шанс или нисходящая каузальная теория (*downhill causal theory*), что он действительно может произойти. Мысль о выигрыше большого приза в лотерею может породить надежду, но не восходящую (*uphill*) мысль о получении значительного дара от неизвестного миллионера. Эти эмоции, как представляется, требуют, чтобы событие (или ситуация) не дотягивало до того, чтобы считаться чем-то, что должно произойти наверняка. Если я *знаю*, что меня скоро казнят, я могу испытывать отчаяние, а не страх. Согласно Стендалю, любовь в равной степени вянет, когда человек уверен в том, что его чувство взаимно, и когда он уверен, что никогда не добьется взаимности. Согласно Ларошфуко, ревность может исчезнуть в тот момент, когда человек *узнает*, что человек, которого он любит, влюблен в кого-то другого.

Некоторые эмоции порождаются *контрфактическими* (*counterfactual*) мыслями о том, что могло случиться или что человек мог сделать. *Разочарование* — это эмоция, возникающая, когда положительное событие, на которое возлагали надежду, не реализуется⁸. *Сожаление* — это эмоция, проистекающая из осознания, что желаемое событие могло бы произойти, если бы мы сделали иной выбор. Позитивные противоположности этих эмоций (вызываемые тем, что не произошло отрицательное событие) иногда называют соответственно *восторгом* (*elation*) или *ликованием* (*rejoicing*) (в повседневном языке они обычно сливаются воедино под именем *облегчения* (*relief*)). Если разочарование и восторг предполагают сравнение разных исходов данного выбора, вызванных разными ситуациями, сожаление и ликование предполагают сравнение разных

⁷ Иногда эмоции могут порождаться неуверенностью относительно прошлого. После падения ГДР многие граждане захотели получить доступ к имевшимся на них досье, чтобы выяснить, шпионил ли за ними кто-нибудь из друзей или родственников. Пикантность в том, что некоторые хотели выяснить, связаны их личные неудачи (отсутствие повышения по службе или разрыв с любимым человеком) с их личными недостатками или же другие люди просто получили инструкции от секретных служб подвергнуть их ostracism.

⁸ Если что-то ускользнуло из рук, то эмоции более интенсивные; так, серебряные медалисты на Олимпийских играх говорят о том, что испытывают меньшее счастье, чем бронзовые призеры.

выборов в рамках одной ситуации. В некоторых случаях негативные события могут приписываться обоим источникам. Если я промок по пути на работу, я могу отнести это на счет случайного метеорологического события или того факта, что я не взял зонт. Хотя я могу предпочесть первый фрейминг, этот пример принятия желаемого за действительное может подвергнуться ограничениям со стороны реальности (глава VII), если, выходя из дома, я слышал прогноз о том, что будет дождь.

ЭМОЦИИ И СЧАСТЬЕ

Роль эмоций в порождении счастья (или несчастья) предполагает идею валового национального счастья (gross national happiness product). Обычные методы измерения экономических процессов, конечно, более объективны. Но, в конце концов, нас не так уж беспокоит объективность в смысле физической измеримости. Причина того, почему мы хотим знать об экономических результатах, заключается в том, что они вносят вклад в *субъективное* чувство благополучия или счастья. Более того, счастье может исходить из источников, которые не поддаются объективному количественному измерению. В 1994 году, когда Норвегия принимала у себя зимние Олимпийские игры, стране пришлось с существенными затратами построить новые спортивные сооружения и здания для размещения участников. Доходная часть может включать деньги, потраченные в стране туристами и посетителями спортивных соревнований, а также выручку, генерируемую возведенными сооружениями в дальнейшем. Экономисты, проводившие эти расчеты, не были уверены в том, что доходы и расходы на Игры могут сравняться. Но я совершенно уверен (но, конечно, не могу этого доказать), что если мы включим сюда эмоциональные выгоды для населения Норвегии, то Игры, возможно, принесли даже большую прибыль. Неожиданно большое число золотых медалей у норвежской сборной создало в стране атмосферу коллективной эйфории, которая была тем сильнее, чем неожиданнее была победа. Объективное число побед в значительной степени было обязано своим резонансом элементу субъективного удивления⁹. Более близкий по времени пример: победа французской футбольной сборной на ЧМ в 1998 году, а также ее поражение в 2002 году породили чувства эйфории и уныния, интенсивность которых во многом была связана с эффектом неожиданности.

⁹ Предположим, что более ранняя вероятность победы составляла p и что удовлетворение от победы пропорционально $1/p$ (так как удивление тем выше, чем ниже p). В этом специальном модусе удивления *ожидаемое* удовлетворение от победы не зависит от вероятности победы. В целом влияние удивления будет более комплексным.

Обычно трудно сравнивать эмоциональный компонент благополучия с другими его составляющими. То, что положительные эмоции на пике интенсивности больше влияют на счастье, чем простое гедонистическое благополучие, ничего не доказывает, пока мы не выясним, как часто случаются эти пики интенсивности. К тому же мы не понимаем, сопровождается ли (и если да, то в какой степени) предрасположенность к эмоциональным подъемам склонностью к эмоциональным спадам. Если ответ положительный, то не является ли жизнь в устойчивом довольстве в целом более счастливой, чем жизнь, в которой эйфория и дисфория сменяют друг друга? Как заметил Монтень, ответ зависит от обстоятельств, которые создает среда: «Если мне скажут, что преимущество иметь притупленную и пониженную чувствительность к боли и страданиям связано с той невыгодой, что сопровождается менее острым и менее ярким восприятием радостей и наслаждений, то это совершенно верно; но, к несчастью, мы так устроены, что нам приходится больше думать о том, как избегать страданий, чем о том, как лучше радоваться»¹⁰. Идеал притупления всех эмоций, который мы находим у многих древних философов, особенно в стоицизме и буддизме, возник в обществах, дававших больше поводов для эмоций с негативной валентностью. Монтень, писавший во времена Религиозных войн, опустошавших Францию, возможно, был в сходной ситуации.

ЭМОЦИЯ И ДЕЙСТВИЕ

Опосредованная связь между эмоцией и действием — это тенденция (готовность) к действию. Последнюю можно также рассматривать как временное предпочтение. Как представляется, с каждой из основных эмоций ассоциируется одна (или несколько) таких тенденций.

Хотя гнев и картезианское негодование вызывают одну и ту же тенденцию к действию, стремление, связанное с гневом, сильнее. Эксперименты показывают, что подопытные готовы нести большие расходы, чтобы причинить вред тем, кто причинил вред им самим, чем тем, кто причинил вред третьим лицам. По окончании Второй мировой войны американцы с большей страстью желали наказать нацистов, которые плохо обращались с американскими военнопленными, а не тех, кто нес ответственность за холокост. Исключение, подтверждающее правило, составляли евреи — члены Администрации Рузвельта.

Эмоции гнева, вины, презрения и стыда тесно связаны с моральными и социальными нормами. Нарушители норм могут испытывать чувство вины или стыда, а те, кто стал свидетелем нарушения, — гнев или

¹⁰ Монтень М. Опыты. Кн. II. Гл. 12. — Примеч. пер.

ТАБЛИЦА VIII.1

| Эмоция | Тенденция к действию |
|------------------------------------|--|
| Гнев или картезианское негодование | Заставить объект эмоций страдать |
| Ненависть | Заставить объект эмоций прекратить свое существование |
| Презрение | Подвергнуть остракизму, избегать |
| Стыд | «Провалиться сквозь землю», убежать, совершить самоубийство |
| Вина | Признаться, загладить вину, нанести себе вред |
| Зависть | Уничтожить предмет зависти или его обладателя |
| Страх | Бежать, сражаться |
| Любовь | Приблизиться и коснуться другого, помочь другому, доставить другому удовольствие |
| Жалость | Утешить или облегчить страдания другого |
| Благодарность | Помочь другому |

презрение. Эти отношения различаются по своей структуре, как показано на рис. VIII.1.

Социальные нормы, которые будут обсуждаться в главе XXII, опосредуются разоблачением перед другими. Вот почему самоубийства, упомянутые в начале этой главы, имели место только тогда, когда постыдные действия были преданы огласке. Как я утверждал в главе V, моральные нормы в этом отношении различаются¹¹.

Некоторые тенденции представляются направленными на восстановление морального равновесия в мире. Причинить вред тем, кто нанес его тебе, и помогать тем, кто содействовал тебе, похоже, является *способом сведения счетов*. Для некоторых случаев это может быть справедливо. Но теория перспектив (глава XII) предполагает, что модель «два oka за одно» может стать лучшим описанием тенденции к действию, свой-

¹¹ Я подозреваю, что нарушение квазиморальных норм вызывает такие же эмоции, что и нарушение моральных норм, хотя не уверен в своей догадке.



РИС. VIII.1

ственной гнева, чем «око за око»¹². Хотя многие акты благодарности могут совершаться лишь ради того, чтобы отдать долг, некоторые могут возникать из подлинного чувства доброй воли по отношению к своему благодетелю. Моральное равновесие обладает большей принудительной силой в случае вины, когда тенденция к действию по заглаживанию вины носит ярко выраженный восстановительный характер. Более того, когда агент не может исправить нанесенный им ущерб, он может восстановить равновесие, нанеся такой же ущерб себе самому. Если я мошенничал с подоходным налогом и выяснил, что налоговая служба не принимает анонимный перевод денег на сумму, которую я ей задолжал, я могу восстановить равновесие, если вместо этого сожгу деньги.

Эмоциональные тенденции к действию не просто порождают желание действовать — они склоняют действовать *рано, а не поздно*. Чтобы ввести эту идею в контекст, позвольте разграничить *нетерпение* (impatience) и *безотлагательность* (urgency). Нетерпение я определяю как предпочтение более быстрого вознаграждения более медленному, то есть неко-

¹² Так, в ответ на убийство 365 ливанских мусульман ливанская женщина сказала, что «в этот момент мне хотелось, чтобы [мусульманская милиция]... зашла в офисы и убила первых семьсот тридцать беззащитных христиан, которые им попадутся». В «Гражданской войне во Франции» Маркс цитирует письмо священника, взятого коммунарами в заложники, в котором он говорит, что «принято решение за каждую казнь, устроенную правительством, казнить двух из многочисленных заложников, которых они держат». Наконец, Сенека заметил: «Непревзойденное зло не является отомщенным».

ТАБЛИЦА VIII.2

| t_1 | t_2 | t_3 | t_4 |
|--|-------|-------|-------|
| Случай 1. Нетерпение | | | |
| А | 3 | | |
| Б | | 5 | |
| Случай 2. Безотлагательность | | | |
| А | | 3 | |
| | Б | 4 | |
| Случай 3. Нетерпение и /или безотлагательность | | | |
| А | | 3 | |
| | Б | | 6 |

торую степень дисконтирования во времени. Как я отмечал в главе VI, эмоции могут заставить агента придавать меньшее значение отдаленным по времени последствиям действия в настоящем. Безотлагательность, еще один эффект эмоций, я определяю как предпочтение более раннего действия более позднему. Различие показано в табл. VIII.2.

В каждом случае агент может совершить одно, и только одно из двух действий, А или Б. В случае 1 эти опции имеются в одно и то же время, в случаях 2 и 3 в последовательные моменты времени. В случае 2 вознаграждения (размеры которых указаны в числах) появляются в одно и то же более позднее время, в случаях 1 и 3 — последовательно в более поздние моменты времени. Предположим, что в неэмоциональном состоянии агент во всех случаях выбирает Б, но в эмоциональном состоянии он выбирает А. В случае 1 выбор А связан с вызванным эмоциями нетерпением. В случае 2 он определяется вызванной эмоциями безотлагательностью. В случае 3 он вызван обоими факторами или их взаимодействием.

Нетерпение, часто обсуждавшийся вопрос, было центральной темой главы VI. Хотя безотлагательность обсуждается гораздо реже, я полагаю, что она может быть очень важной. В частности, безотлагательность эмоций может обуславливать один из тех механизмов, через который они влияют на формирование убеждений. Как мы увидим в главе XI, формирование рациональных убеждений требует оптимального сбора информации. Вместо того чтобы воспользоваться имеющимися под рукой све-

дениями, прежде чем начать действовать, рациональный агент соберет дополнительную информацию, если решение, которое следует принять, довольно важное, а цена ожидания невелика. Безотлагательные эмоции часто возникают в ситуациях, в которых стоимость ожидания *высока*, то есть перед лицом крайней физической опасности. В таких случаях быстрое действие без промедления на сбор дополнительной информации имеет первостепенное значение. Но когда откладывание важного решения может его улучшить, вызванное эмоциями желание немедленного действия может нанести вред. Сенека говорил: «Разум выслушивает обе стороны, потом старается отложить действие, даже свое собственное, чтобы выиграть время для тщательного анализа истины; но гнев поспешен». Пословица «Поспешись жениться, век будешь каяться» указывает как на эмоциональный порыв, так и на неблагоприятные последствия неспособности справиться с ним¹³.

Я говорил, что не все эмоции имеют короткую жизнь. Тем не менее, как правило, эмоции быстро распадаются с течением времени. В некоторых случаях это связано с тем, что вызвавшая их ситуация прекратила свое существование. Когда мне удалось счастливо спастись от угрожавшего мне медведя, страх больше ничем не поддерживается, хотя чаще эмоция приглушается по мере того, как с прошествием времени стираются воспоминания. Гнев, стыд, вина и любовь редко сохраняют ту же интенсивность, какой они отличались при зарождении. После 11 сентября 2001 года, например, количество молодых людей, выразивших интерес к службе в армии, возросло, но во время призыва прироста не последовало. Эти факты соответствуют гипотезе о том, что начальный всплеск интереса был связан с эмоциями, которые остыли в течение нескольких месяцев, которые занимает процесс призыва на военную службу. Роста интереса к службе в армии среди молодых женщин почти не наблюдалось — факт, не имеющий очевидного объяснения.

Люди, как правило, не способны предчувствовать спад своих эмоций. Захваченные сильной эмоцией, они могут ошибочно полагать, что это будет длиться вечно, и даже утратить чувство будущего. Если бы упоминавшиеся мною самоубийцы знали, что стыд (и презрение со стороны наблюдателей) пойдет на спад, они, возможно, не покончили бы с собой. Если бы молодые пары знали, что их любовь не продлится вечно, они

¹³ Урезание гражданских свобод, проведенное несколькими западными правительствами после событий 11 сентября 2001 года, — хороший пример. Иллюстрирует оно необходимость действовать быстро под угрозой неминуемой опасности или же является панической реакцией, которая, сделав эти правительства еще более одиозными в глазах их врагов, скорее увеличила, нежели уменьшила, вероятность дальнейших атак? Сравните с замечаниями о психологии тирании в предыдущих главах.

не стремились бы принимать на себя тяжелые обязательства, в частности заключать брачные договоры, которые потом сложно расторгнуть.

Позвольте мне завершить эту тему указанием на взаимодействие двух явлений, вызванных эмоциями, — инверсию предпочтений и туманное формирование убеждений (*clouded belief formation*). В розовом свете они могут вообще отменять друг друга: при инверсии предпочтений человек может захотеть действовать вопреки своему спокойному и здравому рассуждению, но из-за туманных убеждений он не способен выполнить это намерение. Чаще, как я полагаю, эти явления будут усиливать друг друга. Примером может служить месть. Если я не буду мстить за нанесенную обиду, риск минимален; он увеличится, если я стану мстить, но буду ждать благоприятного момента; и он будет максимальным, если я буду мстить сразу же, не задумываясь о рисках. Похожее наблюдение сделал Монтень: «Философия хочет, чтобы, собираясь ответить за понесенные нами обиды, мы предварительно побороли свой гнев, и не для того, чтобы наша месть была мягче, а напротив, для того, чтобы она была лучше нами обдумана и стала тем чувствительней для обидчика; а этому, как представляется философии, неудержимость наших порывов только препятствует». И все же это игнорирование парадокса, что если мы не чувствуем эмоции, мы можем не захотеть мстить, а если чувствуем, то можем оказаться не в состоянии реализовать нашу месть эффективно.

ЭМОЦИЯ И УБЕЖДЕНИЕ

Эмоция может влиять на формирование убеждений как прямо, так и косвенно. Прямое воздействие производит предвзятые убеждения, не прямое — неполные. Одна из форм предвзятости иллюстрируется теорией *кристаллизации* Стендаля. Происхождение этого термина следующее: «В соляных копиях Зальцбурга, в заброшенные глубины этих копеей кидают ветку дерева, оголившуюся за зиму; два или три месяца спустя ее извлекают оттуда, покрытую блестящими кристаллами; даже самые маленькие веточки, которые не больше лапки синицы, украшены бесчисленным множеством подвижных и ослепительных алмазов; прежнюю ветку невозможно узнать».

Аналогия с любовью очевидна: «С того момента, когда вы начинаете по-настоящему интересоваться женщиной, вы больше не видите ее такой, *какая она есть*, но такой, какой хотите ее видеть. Вы сравниваете лестные иллюзии, создаваемые этим рождающимся интересом, с прекрасными алмазами, скрывающими безжизненную ветку — и которые видны, заметьте, только глазам влюбленного молодого человека».

Во французской пословице, которую я цитировал раньше и процитирую еще раз, мы *легко верим в то, чего боимся*. Это форма предвзя-

тости. В дополнение к тому факту, что мы естественным образом (даже в неэмоциональных состояниях) склонны придавать чрезмерную важность рискам с низкой вероятностью (глава XII), чувство инстинктивного страха тоже может заставить нас поверить в то, что опасности больше, чем на самом деле. Когда мы ночью идем по лесу, малейший звук или движение могут пробудить в нас страх, который заставит нас воспринимать как злоеющие другие звуки и движения, которые мы прежде игнорировали. Страх питает сам себя. Мы также можем рассматривать это как «эффект Отелло». Я останавлиюсь на этом подробнее в главе XXIII.

Безотлагательность эмоции влияет не на само убеждение, а на сбор информации, предшествующий его формированию. В результате получается неполное убеждение, основанное на меньшем количестве информации, чем следовало бы, но непредвзятое, склоняющееся за или против конкретного вывода, который агент хочет считать истинным. На практике, однако, оба механизма идут рука об руку и усиливают друг друга. Сначала агент формирует вызванную эмоциями предвзятость, а потом безотлагательность эмоции мешает ему собрать информацию, которая могла бы скорректировать пристрастность. Как мы видели в предыдущей главе, реальность до некоторой степени ограничивает принятие желаемого за действительное. Противоположный ему феномен также подвержен этим ограничениям. Следовательно, если бы агент собрал больше информации, ему было бы труднее упорствовать в предвзятом убеждении.

Тем не менее я должен повторить, что сбор большого количества информации может оказаться нерациональным. Если вы потратите слишком много времени на то, чтобы решить, похож предмет на дороге на палку или на змею, вы рискуете погибнуть. Следует также заметить, что в неэмоциональных состояниях существует тенденция *упускать из виду* затраты упущенных возможностей и обращать больше внимания на наличные расходы (глава XII). Столкнувшись с возможным риском, агент может не захотеть принимать серьезные меры предосторожности из-за прямых издержек от них, не придавая значения тому факту, что бездействие тоже может обойтись дорого. Таким образом, в некоторых случаях безотлагательность эмоции может обеспечить полезную корректировку этой иррациональной тенденции. В то же время, повторим, безотлагательность и неполные убеждения могут скорее создать проблемы, чем решить их.

ЭМОЦИИ И ТРАНСМУТАЦИЯ

Из-за нормативной иерархии мотиваций (глава IV) люди могут стыдиться своих эмоций. Например, в зависти признаются немногие. Когда Яго говорит: «Пусть этого не будет. А у Кассио жизнь так красива, что я сам себе кажусь уродом», — он необычно, возможно даже, недопусти-

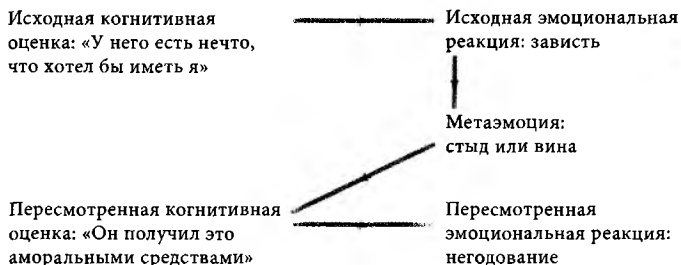


РИС. VIII.2

мо откровенен. Самая распространенная реакция на осознание испытываемой зависти — мысленное пожимание плечами. Человек осознает болезненное чувство неполноценности, у него появляется мимолетная деструктивная интенция, но затем он переступает через это и идет дальше. Иногда, тем не менее, эмоция может оказаться несколько сильной, что ее нельзя игнорировать. В то же время она не может быть признана. Решением конфликта может стать трансмутация зависти в праведное негодование посредством соответствующего переписывания сценария. Я могу рассказать себе историю, согласно которой другой получил предмет моей зависти незаконными или аморальными средствами и, возможно, за мой счет; таким образом, зависть преобразована в аристотелевское негодование или гнев.

В главе VII я обсуждал мотивированное формирование убеждений. Как показано на рис. VIII.2, этот феномен может быть встроен в процесс *мотивированной мотивации*, которую я рассматривал в главе IV. Для того чтобы агент мог принять мотивацию, которой он стыдится, может потребоваться когнитивное переписывание сценария.

Преобразование зависти в праведное негодование может также происходить более прямым образом. Зависть не только постыдна, она причиняет боль. Вера в то, что другой добился успеха там, где мог бы я, приложи я больше усилий, может быть очень неприятна. Чтобы облегчить боль, я могу принять ту же самую историю о недостойных корнях успеха другого человека.

Вина тоже переносится с трудом, особенно если человек горд. Гордые индивиды легко могут сочинить историю, в которой переложат всю вину на других людей: «Тех, кому они причиняют вред, они еще и ненавидят» (Сенека). Эта модель роднит, например, властителей эпохи Возрождения,

древних тиранов и других владетельных особ. Озабоченность своим образом может также вызвать стремление избегать раскаяния. Некоторым индивидам так трудно признать свою ошибку, что они продолжают убыточную деятельность, вместо того чтобы прекратить ее, следуя требованиям рациональности. Это ошибка необратимых затрат, которую также иногда называют ошибкой Вьетнама, ошибкой Конкорда. Я подробнее остановлюсь на этом в главе XII.

КУЛЬТУРА И ЭМОЦИИ

Универсальны ли эмоции? Если нет, есть ли универсальные эмоции? Я скажу твердое «да» в ответ на второй вопрос и осторожное «да» — на первый.

То, что эмоции универсальны, кажется ясным. Есть полтора десятка эмоций — счастье, удивление, страх, печаль, отвращение и гнев, проявление которых на лице человека узнается независимо от культуры. Если кто-то, подобно мне, считает, что социальные нормы существуют во всех обществах, то эмоции, на которые они опираются, должны быть универсальными. Можно было бы представить общество, в котором люди чувствуют гнев, когда обижены, но не испытывают (картезианского) негодования, когда кто-то причиняет вред третьей стороне. Мне трудно поверить в то, что такое общество существует; возможно, я ошибаюсь. Если любовь универсальна (см. дальнейшее обсуждение), то не является ли такой же и ревность?

Говорят, что у японцев есть эмоция *амаэ* (грубо говоря, беспомощность и стремление быть любимым), которая не существует в других обществах. Кроме того, утверждалось, что древняя Греция была культурой стыда, в отличие от современных культур вины, что романтическая любовь — современное изобретение, что чувство скуки (если это вообще эмоция) возникло недавно. Однако нельзя исключать, что эти якобы отсутствующие эмоции могли существовать, но не быть концептуализированными членами этого общества. Эмоция как таковая может быть распознана сторонним наблюдателем, но не быть признана самим обществом. На Таити мужчина, которого оставила женщина, будет проявлять поведенческие симптомы печали, заявляя при этом, что всего лишь устал. На Западе *концепция* романтической любви появилась относительно недавно — во времена трубадуров. До этого была только веселая чувственность или безумие. Но возможно и, по моему мнению, очень вероятно, что *опыт* романтической любви встречался, даже когда у общества не было концепции для него. Люди могут любить друг друга, сами того не замечая, в то время как эмоции будут бросаться в глаза посторонним наблюдателям из собственного или другого общества.

Древние греки демонстрировали набор эмоций, связанных с чувством вины, — гнев, прощение или возмещение ущерба, что указывает на ее наличие, даже если у них отсутствовало для этого название. То, как люди осмысливают эмоции, может зависеть от конкретной культуры, хотя сами эмоции могут от культуры не зависеть.

Следует при этом отметить, что если какая-то эмоция эксплицитно не концептуализирована, она может в меньшей степени проявляться в поведении человека. Ларошфуко писал: «Некоторые люди никогда бы не влюбились, если бы ничего не слыхали о любви». Чувство вины также в большей степени распространено в обществах, в которых людям с раннего возраста внушают необходимость испытывать вину в той или иной ситуации.

Библиографические примечания

Лучшая книга об эмоциях — «Эмоции» Н. Фрижды (Frijda N. The emotions. Cambridge University Press, 1986). Много из этой работы я почерпнул для моей «Алхимии разума», которая носит подзаголовок «Рациональность и эмоции» (Alchemies of the Mind. Cambridge University Press, 1999), в которой читатель может найти дальнейшие отсылки к обсуждаемым здесь концепциям. Идея безотлагательности, отличной от нетерпения, является дополнением к концептуальной рамке этой книги. Серьезные обсуждения роли эмоций в поведении содержатся в «Понимая этническое насилие: страх, ненависть и рессентимент в Восточной Европе двадцатого века» П. Петерсена (Petersen P. Understanding Ethnic Violence: Fear, Hatred and Resentment in Twentieth-Century Eastern Europe. Cambridge University Press, 2002) и в главе 8 моей книги «Подводя итоги: Переходная справедливость в исторической перспективе» (Closing the Books: Transitional Justice in Historical Perspective. Cambridge University Press, 2004).

ДЕЙСТВИЕ

Ч А С Т Ь Т Р Е Т Ь Я

Хотя я часто использую термины «действие», «поведение», «решение» и «выбор» как синонимические, иногда полезно их различать. Самая широкая категория — *поведение* (behavior), понимаемое как любое телесное движение, исходящее изнутри, а не извне (как в случае, когда его уносит лавина) агента. *Действие* (action) — это интенциональное поведение, вызываемое желаниями и убеждениями агента. Таким образом, рефлекторные поступки не являются действиями; эрекция — это не действие (но она может быть спровоцирована каким-то действием, например приемом виагры); засыпание — это не действие (но оно может быть вызвано приемом снотворного). Действию может предшествовать (или не предшествовать) сознательное *решение* (decision). Когда я еду на работу привычным маршрутом, я не принимаю сознательного решения повернуть здесь направо, а там — налево, хотя каждое действие является интенциональным, или ориентированным на достижение цели. Однако когда я ехал на работу в первый раз, всем действиям предшествовало эксплицитное решение. Действительно, им предшествовал эксплицитный *выбор* (choice) из альтернативных маршрутов. Хотя любой выбор есть решение, обратное не является верным. Когда я решаю взять книгу, которую начал читать, у меня в голове нет никакой эксплицитной альтернативы. Я вижу на столе книгу, ее вид напоминает мне о том, что мне нравится ее читать, и я решаю ее взять. Выбор здесь не задействован.

Самая важная черта этого концептуального пейзажа, возможно, заключается в том, что *не все решения ведут к действиям*. Человек может решить чего-то не делать, например не спасать тонущего человека, если действие подвергает некоторому риску его самого. Если человек тонет и в этом не замешаны третьи лица, у меня нет каузальной ответственности за исход этого события. Я могу нести моральную (а в некоторых странах и юридическую) ответственность, но это другой вопрос¹.

¹ В США никому не вменяется в обязанность быть добрым самаритянином, за исключением некоторых специфических случаев. В континентальной Европе не оказание помощи человеку, находящемуся в опасности, может быть сурово наказано, если риск для оказывающего помощь меньше той опасности, которой подвергается человек, в ней нуждающийся. Некоторые американские теоретики-юристы утверждают, что американское законодательство более эффективно, поскольку всеобщая обязанность оказывать помощь может заставить потенциальных спасателей избегать тех мест, в которых может потребоваться спасение людей, из-за угрозы ответственности. Это может быть так или (скорее всего) не так. Одно можно сказать наверняка: американская система была создана и продолжает существовать не из соображений эффективности.

Но предположим, что присутствует третье лицо или, как в случае Китти Дженовезе, много третьих лиц. Если кто-то из них замечает, что я в состоянии помочь тонущему, но не делаю этого, он может справедливо заключить, что ситуация менее серьезная, чем она показалась бы в противном случае, и в результате тоже не оказать помощь. В этом случае мое решение ничего не делать заставит другого человека поступить так же. Таким образом, решения могут иметь каузальную силу, даже когда не порождают никаких действий.

Многое в третьей части вращается вокруг теории рационального выбора. Как я объяснял во Введении, со временем я стал более скептически, чем раньше, относиться к объяснениям при помощи этой теории. Хотя поведение так или иначе является иррациональным, в определенном смысле рациональность имеет первичное значение. Человеческие существа *хотят* быть рациональными. Мы не испытываем гордости за отступления от рациональности. Мы стремимся скорее избежать или исправить их, если только гордыня не мешает нам их признать.

IX. Желания и возможности

4

ДЕЛАЯ ВСЕ ВОЗМОЖНОЕ

Характеризуя поведение, мы иногда говорим: «Он сделал все, что мог». Если разобрать это предложение, то оно будет содержать два элемента: желания и возможности, то есть желания определяют то, что для агента считается «всем»; возможности — это варианты или средства, из которых агент «может» выбирать. Эта характеристика может также служить рудиментарным *объяснением* поведения через рациональный выбор. Если мы спрашиваем: «Зачем он это сделал?», ответ: «Это все, что было в его силах», — может полностью нас удовлетворить. Во многих случаях требуется нечто большее, чтобы дать удовлетворительное объяснение с позиций рационального выбора. В частности, нам придется апеллировать к убеждениям агента, а не только к его желаниям и возможностям. Мы коснемся этих вопросов в главе XI. Здесь я собираюсь проанализировать, как далеко может завести нас простая рамка *желания-возможности*¹. Я также продемонстрирую, что эта рамка порой не такая простая, какой может показаться, поскольку желания и возможности не всегда (как порой считается) независимы друг от друга.

Есть другой, сходный взгляд на эту проблему. Анализируя поведение, мы можем начать со всех гипотетически возможных действий, которые мог предпринять индивид. Действие, которое мы наблюдаем, может рассматриваться как результат двух успешных процедур фильтрации. Первый фильтр учитывает все *ограничения* — физические, экономические, юридические и другие, с которыми сталкивается агент. Действия, согласующиеся со всеми существующими ограничениями, образуют набор допустимых решений (*opportunity set*)². Второй фильтр — это механизм, который определяет, какое именно действие из набора будет реализо-

¹ При этом я делаю скрытое допущение, что набор вариантов, которые, по мнению агента, у него имеются, совпадает с объективным множеством допустимых решений.

² Когда я ссылаюсь на юридические ограничения, я не имею в виду действие законов, делающих одни поступки более затратными, чем другие (это относится ко второму фильтру), но их эффект, делающий одни действия возможными, а другие — нет. Я могу голосовать только в дни проведения выборов и жениться только в специально предназначенных для этого местах.

вано. Я предполагаю здесь, что агент выбирает действие с наилучшими последствиями с точки зрения своих желаний (или предпочтений). В последующих главах мы будем рассматривать механизмы, связанные со вторым фильтром.

Фильтрационный подход подводит к следующему вопросу: что, если ограничения столь велики, что второму фильтру не с чем работать? Может случиться так, что ограничения предложат одно, и только одно действие, которое учитывает все из них? Богатые и бедные в равной мере имеют возможность спать в Париже под мостом, но у бедных может не быть никаких других возможностей³. Для неимущего потребителя ограниченность в средствах и ограничение в потреблении калорий могут совместно определить уникальный набор благ⁴. Можно сказать, что те, кто защищает *структурализм* в социальных науках, утверждают, что ограничения столь велики, что совсем (или почти совсем) не оставляют места выбору⁵. Почему мы должны так считать, остается загадкой. Невозможно, например, утверждать, что богатые и власть имущие не оставляют бедным и угнетенным иного выбора, кроме как работать на них, потому что такое утверждение предполагает, что у первых по крайней мере есть выбор.

В некоторых случаях подход с точки зрения *желаний-возможностей* неполон. У меня может быть возможность совершить действие, которое лучше всего реализует мое желание, например выбрать правильный ответ в тесте или попасть в мишень в соревнованиях по стрельбе, но не *способность* определить это действие. В некоторых случаях неимение этой способности восходит к более ранним этапам, когда у агента не было возможностей либо желания ее приобрести. В других случаях ее отсутствие связано с глубокими психологическими ограничениями. У меня могут быть и желание, и возможность выбрать действие, которое максимально увеличит мое долгосрочное благополучие, но не способность определить прямо на месте, какое это действие. Когда экономисты

³ Если дать более тонкое описание множества допустимых решений, можно предположить, что у бедных есть выбор того, под каким именно мостом спать. Это общее замечание; для любого описания вариантов можно предложить ситуацию, в которой у агента будет только одно осуществимое решение, но для любой ситуации можно найти описание, согласно которому таких вариантов будет более одного.

⁴ Однако, как правило, существует несколько стратегий выживания на прожиточный минимум. В отличие от выбора того, под каким мостом спать, они могут различаться самым небанальным образом. Один из недостатков трудовой теории стоимости Маркса заключается в том, что он не смог понять этот факт.

⁵ Другая, не связанная со структурализмом идея, рассматривалась в главе 1.

и ориентированные на математику политологи пытаются определить, какое поведение является оптимальным в некоторой ситуации, им зачастую нужны многостраничные математические выкладки, чтобы все расписать. Некоторые индивиды никогда не смогут провести такие расчеты, независимо от пройденной подготовки, или же потратят на это больше сил, чем заслуживает задача, или больше времени, чем имеется в их распоряжении.

ОБЪЯСНЕНИЯ ВОЗМОЖНОСТЕЙ

Даже когда поведение является совокупным результатом действия желаний и возможностей, *дисперсия* (variance) с течением времени может быть в значительной степени предопределена возможностями. Потребление алкоголя обычно объясняется как силой желания выпить (в сравнении с другими желаниями), так и тем, что люди могут себе позволить. Когда цены на алкоголь стремительно растут, например, в военное время, потребление спиртного резко падет.

Для объяснения можно было бы использовать кривые безразличия (см. рис. IX.1). Предположим, что потребитель должен разделить свой доход между затратами на алкоголь и на корзину потребительских товаров. Относительные цены и его доход первоначально таковы, что он имеет возможность находиться внутри треугольника OAA' . Предположив, что он распоряжается всем своим доходом, мы можем ограничиться бюджетной линией AA' ⁶. Сила его желания купить алкоголь в сравнении с потребительской корзиной представлена в форме *кривых безразличия* I, I' и I'' . Определение отражает ту идею, что потребителю безразличны комбинации алкоголя и набора благ, которые лежат на любой из данных кривых, при этом он предпочтет любое сочетание, лежащее на верхней кривой, сочетанию, лежащему на нижней⁷. Чтобы выбрать наилучшую из имеющихся у него возможностей, потребитель должен взять на бюджетной линии точку, лежащую на касательной к кривой безразличия, потому что это будет самая высокая среди кривых, включающих комбинацию, которую он может себе позволить. На рис. IX.1 это дает потребление алкоголя OX .

⁶ Иными словами, я не принимаю в расчет то, что человек может работать сверхурочно, заниматься самогонноварением или покупать контрабанду. В полицейских досье эти варианты имеют важное значение.

⁷ Форма кривых соответствует тому факту, что чем больше алкоголя агент потребляет в данный момент, тем больше алкоголя ему нужно компенсировать (и остаться на том же уровне благосостояния) при сокращении набора благ.

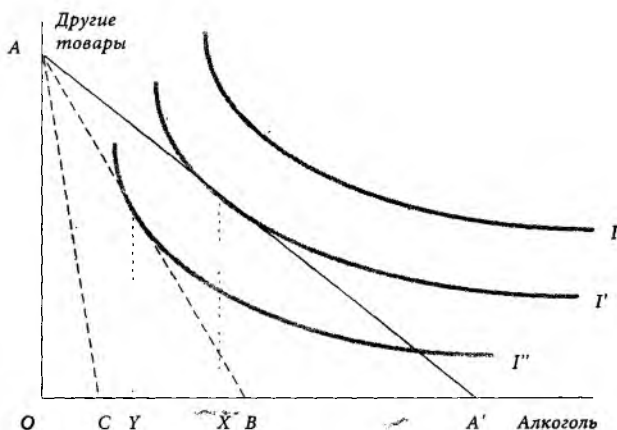


РИС. IX.1

Представим теперь, что цена алкоголя растет, так что потребитель теперь сталкивается с бюджетной линией AB . Так как точка касания сдвигается влево, он будет теперь потреблять на OY . Мы могли бы провести те же рассуждения, если бы дальнейший рост цен переместил бюджетную линию на AC . Но даже ничего не зная о кривых безразличия, мы можем предположить, что в этой ситуации потребитель не станет потреблять больше, чем на OC , как это произошло бы, если бы он тратил весь свой доход на алкоголь. Сам по себе набор возможностей может объяснить многое в дисперсии с течением времени. Вторым фильтром на самом деле может быть чем угодно — оптимизирующим поведением, непреодолимой тягой к алкоголю, привычкой или чем-то еще, но потребитель в любом случае будет жестко ограничен первым фильтром.

Я выбрал этот пример, чтобы обсудить вопрос о якобы непреодолимых желаниях, таких как тяга наркоманов, заядлых курильщиков или алкоголиков к веществам, от которых у них есть зависимость. На что больше похожи наркотики — на инсулин, который диабетики будут покупать по любой цене, или на сахар, потребление которого сокращается по мере роста цен? В подтверждение того, что наркотики больше похожи на сахар, часто приводят факт, что их потребление снижается при увеличении цены. И все-таки, как мы видели, это может быть связано только с неспособностью человека, испытывающего зависимость, выйти за пределы своего бюджета (диабетик может потерять возможность покупать инсулин, если цены на него вырастут). Таким образом, падение потребления

алкоголя в военное время часто вызвано его отсутствием, оставляя вопрос о преодолимых и непреодолимых желаниях открытым. Благодаря другим фактам нам, тем не менее, известно, что потребление алкоголя *зависит* от цены. Даже когда потребители могут поддерживать прежний уровень потребления при более высоких ценах, они этого не делают.

Чтобы показать, что возможности имеют большую объяснительную силу, чем желания, иногда приводят другой аргумент. Экономисты утверждают, что все потребители имеют одинаковые желания и предпочтения, а различаются только возможности. Хотя может показаться, что у людей разные пристрастия в классической музыке, наблюдаемые различия в потреблении вызваны (как утверждается) лишь тем обстоятельством, что у некоторых индивидов больше «капитал на потребление музыки»; следовательно, они получают от нее больше удовольствия, чем остальные. Подготовленное ухо ближе скорее к способности, чем к возможности, это верно, но именно способность зависит от возможности ее приобрести. Считается, что желание приобрести способность у всех одинаково. Последнее утверждение, однако, приоткрывает слабость этого аргумента. Сегодня доступ к классической музыке по беспроводному радио практически ничего не стоит, даже с учетом альтернативных издержек (можно слушать музыку, занимаясь другими вещами). Поскольку все хотят иметь натренированный слух, объяснение может быть найдено в субъективной склонности к его приобретению, которая является производной (помимо прочего) от того, как классическая музыка влияет на *неподготовленное* ухо.

ДРУГИЕ ПРИМЕРЫ ТОГО, ПОЧЕМУ ВОЗМОЖНОСТИ ВАЖНЕЕ ЖЕЛАНИЙ

Возможности имеют фундаментальное значение в еще одном ключевом аспекте: их легче наблюдать не только социологам, но и другим членам общества. В военной стратегии основное правило состоит в том, что нужно строить планы, исходя из (верифицируемых) возможностей противника, а не из его (нераскрытых) намерений⁸. Если у нас есть основание полагать, что противник *может* иметь враждебные намерения, это правило может привести к наихудшему допущению: противник нанесет нам удар, если сможет⁹. Ситуация осложняется тем, что наша вера во враж-

⁸ Переоценка советской или иракской мощи может нам, тем не менее, напомнить, что даже возможности очень трудно верифицировать.

⁹ Даже если у нас нет оснований думать, что другой народ вынашивает враждебные намерения, может иметь место планирование на случай чрезвычайной ситуации. Возможно, где-то глубоко в недрах американского Министерства обороны существует план вторжения в Канаду.

дебные намерения противника может основываться на ощущениях, что он полагает, что у нас есть средства и, возможно, намерение напасть на него. В этой трясине субъективности объективные возможности могут показаться единственным твердым фундаментом для планирования.

Другая причина того, почему возможности представляются важнее желаний, имеет отношение к средствам влияния на поведение. Обычно проще изменить обстоятельства и возможности людей, чем их мысли¹⁰. Это аргумент с точки зрения сравнения затрат и результата, связанный с долларовой эффективностью альтернативной политики, а не довод об относительной объясняющей силе. Даже если у правительства есть теория, которая позволяет давать объяснения и делать прогнозы, она может не обеспечивать приемлемый уровень контроля, если элементы, на которые она воздействует, лишены достаточной важности с точки зрения каузальности.

Предположим, что слабые экономические показатели могут быть связаны с нежеланием предпринимателей рисковать или с сильным влиянием профсоюзов. Правительство может пребывать в полной уверенности, что первостепенную важность имеет настрой менеджеров, и при этом не иметь возможности на него повлиять. Профсоюзы же, как показали годы правления Рейгана и Тэтчер, могут быть сломлены действиями правительства.

В качестве важного примера рассмотрим суицидальное поведение. Чтобы совершить самоубийство, одного желания лишить себя жизни недостаточно, нужны средства это сделать. Высокий процент самоубийств среди врачей, например, может быть отчасти вызван их легким доступом к смертельно опасным лекарствам, которые являются любимым средством покончить с жизнью в этой группе¹¹. Хотя правительство может попытаться ограничить намерения самоубийц, создавая «горячие линии» помощи и убедив прессу меньше сообщать о таких случаях, чтобы не провоцировать их эпидемию, самые эффективные результаты дает затруднение доступа к средствам самоубийства¹². Такая политика может включать установку барьеров и ограждений, затрудняющих прыжки с мостов или с крыш высотных зданий, усиление контроля за некоторыми лекарствами, выдаваемыми по рецепту, ограничение продажи огнестрельного оружия, замену в кухонных плитах смертельно опасного

¹⁰ Кроме того, как будет показано далее, лучший способ изменить мысли — это изменить обстоятельства.

¹¹ Вопреки распространенному мнению доступ полицейских к огнестрельному оружию не делает их более склонными к самоубийству, чем других.

¹² Когда в Британии в 70-е годы резко упал процент самоубийств, первоначально это связывали с созданием благотворительной организацией «Самаритяне» линий помощи, но позднее оно было объяснено переходом со смертельно опасного угольного газа на менее опасный натуральный в домашних плитах.

угарного газа на натуральный и установку на автомобилях каталитических конвертеров, которые уменьшают выхлопы углекислого газа. В будущем мы можем стать свидетелями запрета интернет-сайтов, помогающих самоубийцам. Даже простой переход с бутылочек на блистеры способствовал снижению процента самоубийств путем отравления парацетамолом. Сокращение количества таблеток, которые выписывает врач или готовит фармацевт для отдельного больного, тоже может уменьшить вероятность серьезного отравления. Во Франции, в отличие от Англии, содержимое каждой упаковки парацетамола официально ограничено восьмью граммами. Это считается причиной того, что во Франции поражение печени и смерть от сильного отравления парацетамолом случаются реже, чем в Англии.

Наверняка полный решимости индивид найдет способ покончить с собой. Когда люди лишаются привычных средств сведения счетов с жизнью, следующий за этим спад в количестве самоубийств может быть временным. И все же в некоторых случаях удастся добиться долгосрочного эффекта, как можно было ожидать. Если порыв убить себя скорее мимолетен, чем твердо обоснован, он может пройти к тому времени, как человек найдет подходящее средство¹³. Таким образом, простое *откладывание* (а не блокирование) доступа к средствам будет эффективным для предотвращения импульсивного самоубийства. Введение периода ожидания перед приобретением огнестрельного оружия помогло бы сократить количество как самоубийств, так и убийств¹⁴.

РАЗНОВИДНОСТИ ВЗАИМОДЕЙСТВИЯ ЖЕЛАНИЙ И ВОЗМОЖНОСТЕЙ

Более сложный пример взаимодействия желаний и возможностей можно найти в анализе фракций, сделанном Мэдисоном в «Федералисте». Он утверждает, что в условиях прямой демократии (которая неизбежно будет ограниченной в размерах) или в небольшой представитель-

¹³ Я не думаю, что *ценовой* фактор может ему помешать. В решении покончить с собой могут присутствовать соображения, касающиеся затрат и результатов, например, когда агент сравнивает боль, которую он причинит другим людям, с болью, от которой самоубийство избавит его самого, но затраты на то, чтобы изыскать подходящее средство, для решительно настроенного индивида имеют значение.

¹⁴ На самом деле в большинстве американских штатов, в которых практикуется определенное ожидание перед приобретением оружия, оно введено для того, чтобы дать властям время проверить, нет ли у будущего покупателя уголовного прошлого или психического заболевания, а не для того, чтобы покупатель получил возможность успокоиться.

ской республике *фракции* будут иметь и мотивы, и средства для ослабления друг друга. С одной стороны, «общее увлечение или интерес почти во всех случаях будут владеть большинством»; с другой — малая численность граждан позволит большинству угнетать меньшинство, так как в рамках небольшого государства это легче сделать технически. В большой республике, наоборот, «значительно уменьшится вероятность того, что у большинства возникнет общий повод покушаться на права остальных граждан, а если таковой наличествует, всем, кто его признает, будет труднее объединить свои силы и действовать заодно». При таком переносе рассуждения о желаниях и возможностях с индивидуального уровня на коллективный оно приобретает несколько иную форму. Хотя, по мысли Мэдисона, институциональной модели недостаточно, чтобы изменить мотивацию индивидов, она может повлиять на вероятность появления у большинства общего мотива. Хотя институциональное устройство, строго говоря, не может влиять на возможности членов фракционного большинства (если таковое существует) действовать сообща, оно может ограничить их *способность* сделать это, так как им будет труднее узнать о существовании друг друга. До эпохи соцопросов, должно быть, нередкими были случаи, когда молчаливое большинство было просто не в состоянии опознать себя в качестве такового.

У Мэдисона двунаправленный аргумент: большая республика не только мешает формироваться фракционному большинству, но она затрудняет совместные действия. В «Демократии в Америке» Токвиля мы найдем много аргументов типа «не только, но и». Рассмотрим, например, его рассуждения о влиянии рабства на рабовладельцев. Во-первых, рабство убыточно в сравнении со свободным трудом. «Свободный работник получает жалование, тогда как раб получает воспитание, пищу, лечение и одежду, хозяин тратит свои деньги на содержание раба понемногу, небольшими суммами, едва это замечая. Жалование рабочего выплачивается сразу и, как кажется, только обогащает того, кто его получает, но на самом деле раб обошелся дороже, чем свободный человек, и его труд менее продуктивен»¹⁵. Но «влияние рабства распространяется еще шире, проникая в душу хозяина и придавая особый оттенок его идеям и вкусам». Поскольку труд ассоциируется с рабством, белые с Юга презирают «не только сам труд, но и те предприятия, для которых он является необходимым условием успеха». У них нет ни возможностей, ни желания богатеть: «Рабство... не только мешает белым людям зарабатывать состояние, но даже отвращает их от этого желания». Если Токвиль прав, клас-

¹⁵ Этот аргумент нехарактерно тёмен. Более простой довод заключается в том, что, за исключением некоторых отраслей сельского хозяйства, рабство неприбыльно, потому что не создает для раба стимулов трудиться с отдачей.



РИС. IX.2

сический спор об экономическом застое в рабовладельческих странах является мнимым. Нет нужды спрашивать, что именно дает правильное объяснение — отсутствие желания инвестировать или отсутствие возможностей инвестировать; обе стороны могут оказаться правыми.

Аргументы, приведенные Мэдисоном и Токвилем, имеют одну и ту же структуру: одна и та же третья переменная влияет и на желания, и на возможности, которые совместно влияют на действие (или в некоторых случаях мешают ему совершиться). Если говорить абстрактно, есть четыре возможности (знаки «плюс» и «минус» указывают на положительные и отрицательные каузальные эффекты) (см. рис. IX.2).

Случай А иллюстрирует мэдисоновский анализ прямой демократии или маленьких республик. Случай В иллюстрируется его аргументом в пользу больших республик и анализом воздействия рабства на рабовладельца Токвиля.

Случай С часто наблюдается там, где нехватка ресурсов имеет двойной эффект увеличения стимула для улучшения собственной ситуации и уменьшения возможностей это сделать. Хотя и говорится, что «нужда — мать изобретательности», это верно, пока трудности усиливают мотивацию на внедрение инноваций. Но поскольку инновации часто требуют ресурсов (которые соответственно могут быть названы «отцами» изобретательности), мотивация сама по себе может ни к чему не привести. Инновации часто требуют затратных инвестиций с долгосрочными и неопределенными перспективами окупаемости, а именно

этого не могут позволить себе фирмы, находящиеся на грани банкротства. Процветающие фирмы могут осуществлять инвестиции, но их это часто не интересует. Экономист Джон Хикс сказал: «Самая большая *монопольная выгода — это тихая жизнь».

Сходным образом желание эмигрировать может быть подкреплено нищетой на родине, но та же самая нищета может ограничивать возможности эмиграции из-за стоимости переезда. До начала XIX века те, кто эмигрировал в Соединенные Штаты, могли использовать в качестве залога свое тело. Их будущие работодатели оплачивали их переезд в обмен на период службы по договору сервитута. Сегодня те, кто нелегально занимается переправкой эмигрантов, могут использовать страх перед службой иммиграции и натурализации, чтобы заставить нелегальных иммигрантов выполнять свои обязательства по оплате затрат на переезд с доходов, полученных в принимающей стране. Но когда ирландцы в 1840-х годах бежали от голода, самые бедные остались умирать дома.

Еще один пример случая *C* можно найти в исследовании крестьянских бунтов. Хотя самый большой стимул к восстанию был у самых нищих крестьян, у них могло не быть к этому средств. Участие в коллективных действиях требует возможности на время отойти от производственной деятельности, а именно этого бедные крестьяне себе позволить не могли. Середняки, имевшие небольшие запасы, могли присоединиться к восстанию, но их мотивация была слабее. Маркс утверждал, что цивилизация зародилась в умеренных зонах, потому что именно там желания улучшений совпали с их возможностями. В местах со слишком богатой природой нет желания, а там, где природа слишком бедна, нет возможностей. Как показывает этот пример, существует широкий разброс возможностей, в рамках которых желания и возможности развиваются настолько, чтобы порождать действие, но априори ничего нельзя сказать о том, насколько он широк или узок, даже существует ли он.

Пример случая *D* мы видели в главе II. Верхняя часть рис. II.1 показывает то, как, согласно рассуждениям Токвиля, демократия (при посредничестве религии) сдерживает желание вести себя распушенно, какое поведение становится возможным благодаря таким демократическим институтам, как свобода печати и собраний. Более тривиальное наблюдение Токвиля основывается на соединении *C* и *D* с молодежью как третьей переменной. «В Америке большинство богатей были в молодости бедны, большинство праздных людей трудились в юности, в результате в том возрасте, когда есть тяга к знаниям, на них нет времени, а когда есть время, исчезла тяга»¹⁶.

¹⁶ Несовпадение другого рода представлено во французской поговорке «Если бы молодость знала, если бы старость могла», которая предполагает, что



РИС. IX.3

Желания и возможности также могут влиять друг на друга напрямую. Рассмотрим первый случай E на рис. IX.3. В главе II я коснулся того, как возможности могут влиять на желания: люди могут начать желать то, что они могут получить, или предпочитать то, что у них есть, тому, чего нет. Мы можем снова процитировать определение рабства Токвиля: «Это дар Господа или последнее проклятье, такое состояние души человека, которое внушает ему своего рода порочное влечение к причине его страданий». Данный механизм подсказывает еще одну причину, по которой можно считать возможности более фундаментальными, чем предпочтения. Возможности и желания вместе являются непосредственными причинами действия, но на большем отдалении значение имеют только возможности, поскольку они воздействуют и на желания тоже. Механизм адаптивного формирования предпочтений (форма редукции диссонанса) гарантирует выбор наиболее предпочтительной опции внутри набора возможностей, а не за его пределами.

Можно задаться вопросом: *имеет ли значение* этот механизм для поведения, так как опции, которых нет внутри набора возможностей, не выбираются¹⁷? Предположим, что агент первоначально располагает варианты в таком порядке: A, B, C, D, а потом узнает, что A отсутствует. В результате адаптивного формирования предпочтений он теперь располагает их в последовательности B, A, C, D. Он выберет B, так как

два условия для действия — знание того, что нужно делать, и энергия сделать это — никогда не совпадают. Норвежская поговорка поворачивает дело иным образом: «Средний возраст — это когда ты достаточно стар, чтобы знать, и достаточно молод, чтобы все-таки делать».

¹⁷ Адаптация может иметь значение для благополучия агента, обеспечивая ему спокойствие духа. Этот эффект, однако, неоднозначен, поскольку самоинтоксикация духа, сопровождающая многие примеры «кислого винограда» и «сладкого лимона» (как те бродвейские шоу, что я обсуждал в главе I), могут в более широкой перспективе отвлечь от благополучия.

предпочтения у него остались теми же. Предположим, однако, что новая иерархия — C, B, A, D, включая выбор C, которая могла сложиться в результате сверхадаптации к ограниченным возможностям. Токвиль утверждает, что таковым было особое качество француза: «Он выходит за пределы духа рабства, как только в них попадает». Скорее всего здесь мы имеем дело с общей тенденцией, наблюдаемой в статусных обществах. Кроме того, новая последовательность предпочтений может быть B, C, D, A. Если меня отвергают красивые женщины, я могу утешиться мыслью о том, что в силу их нарциссизма они наименее желанные партнеры.

Рассмотрим, наконец, случай F, в котором набор возможностей задается желанием агента. Это может происходить как посредством намеренного выбора, который я рассматриваю в главе XIII, так и через неинтенциональный причинно-следственный механизм. Как я заметил в главе IV, желание, направленное на состояния, являющиеся по преимуществу побочными, могут помешать вызвать их к жизни. Механизм может быть интрапсихическим, как в случае обреченного на провал желания заснуть, или интерперсональным, как в следующем примере. На университетском факультете, где я преподавал, было негласное правило: член факультета, который демонстрировал свое желание занять кафедру, тем самым лишал себя возможности получить эту должность. Карьера Токвиля может служить примером обеих частей этого утверждения. На ранних стадиях его политические амбиции были отвергнуты по причине их чрезмерной зримости (как в пословице «Кто идет на конклав папой, выходит кардиналом»). Позднее, комментируя свой успех на выборах в Национальное собрание в 1848 году, он писал: «Ничто так не приносит успех, как не слишком горячее желание его добиться».

Библиографические примечания

Теория непреодолимых желаний успешно развенчана Г. Уотсоном в «*Расстройстве аппетита: пристрастия, компульсии и зависимость*» (Watson G. *Disordered appetites: Addiction, compulsion, and dependence* // Elster J. (ed.). *Addiction: Entries and Exits*. New York: Russell Sage, 1999). Утверждение, что у индивидов одни и те же предпочтения, различающиеся только наличными возможностями, с которыми они имеют дело, в частности, ассоциируется со статьей Дж. Стиглера и Г. Бекера «О вкусах не спорят» (Стиглер Дж., Беккер Г. О вкусах не спорят // США: экономика, политика, идеология. 1994. № 2). Идеей адаптивных предпочтений и, в частности, сверхадаптации к ограничениям я обязан П. Вену и его работе «Хлеб и зрелища» (Veune P. *Le pain et le cirque*. Paris: Seuil, 1976). Документальные свидетельства спада в (общем) количестве самоубийств после перехода с угольного газа на натуральный в Британии приводятся в статье Н. Креймана «История

угольного газа» (Kreiman N. The coal gas story // British Journal of Social and Preventive Medicine. 1976. No. 30. P. 86–93). Степень, с которой сокращение доступа к одним средствам самоубийства вызывает увеличение использования других, обсуждается в статье К. Кантора и П. Баума «Доступ к средствам самоубийства: Какое воздействие?» (Cantor C., Baume P. Access to methods of suicide: What impact? // Australian and New Zealand Journal of Psychiatry. 1998. No. 32. P. 8–14). Использование Мэдисоном различия *возможности-желания* анализируется в книге М. Уайта «Философия, федералист и Конституция» (White M. Philosophy, The Federalist, and the Constitution. New York: Oxford University Press, 1987. Я обсуждаю использование этого различия у Токвиля в главе 4 «Политической психологии» (Political Psychology. Cambridge University Press, 1993).

Х. Люди и ситуации

Стыд и вина или презрение и гнев различаются, поскольку первая эмоция в каждой паре направлена на *характер* человека, а вторая — на *действие*, совершаемое этим человеком (глава VIII). Похожим образом гордыня основывается на убежденности в собственном превосходстве, а гордость — на вере в совершении какого-то выдающегося поступка. Но когда мы порицаем или хвалим какое-то действие, то не потому ли, что оно отражает характер человека? Действию каких других факторов можно это приписать?

В ЧЕМ ОШИБКА ФОЛК-ПСИХОЛОГИИ?

Эта книга не панегирик и не обличение, но *объяснение* человеческого поведения. В таком контексте проблема состоит в том, насколько характер может в этом помочь. Считается, что у людей есть индивидуальные особенности (интроверсия, робость и т. д.), а также добродетели (честность, мужество и т. д.) или пороки (семь смертных грехов и т. д.). Предположение фолк-психологии заключается в том, что эти черты стабильны: не меняются со временем и в зависимости от ситуации. Это допущение подтверждают пословицы, существующие во всех языках мира: «Раз солгал, а навек лгуном стал», «Кто лжет, тот и крадет», «Кто крадет яйцо, украдет и быка», «Верный в малом и во многом верен», «Пойманному с поличным доверия нет». Если бы фолк-психология была права, предсказание и объяснение поведения было бы легким делом. Одно действие раскрывало бы лежащую за ним черту характера или склонность и позволяло бы нам прогнозировать поведение в бесконечном количестве случаев, в которых эта склонность может проявиться. Это не тавтологическая процедура, какой она была бы, если бы мы сочли жульничество при сдаче экзамена свидетельством нечестности, а в дальнейшем использовали бы ее для объяснения обмана. На самом деле последний используется как подтверждение черты характера (нечестности), которая также может толкнуть человека на то, чтобы обманывать супруга. Если принять более радикальную фолк-теорию о сочетании добродетелей, обман можно использовать для предсказания трусости в бою или чрезмерного пристрастия к алкоголю.

Люди часто делают далеко идущие выводы на основании сдержанного поведения других. Один из членов Французской академии, как говорят, голосовал за де Голля из-за его достойной личной жизни, опираясь

на молчаливое допущение, что тот, кто изменяет жене, может изменить Родине. Во Вьетнаме коммунистические лидеры смогли завоевать сердца и умы народа благодаря своему неподкупному поведению, которое разительно отличалось от значительно меньшей жертвенности представителей других политических сил. В среде мафиози романы считаются признаком слабости и распушенности.

В некотором смысле фолк-психология работает как самоисполняющееся пророчество. Если люди *верят*, что другие будут предсказывать их поведение в ситуации типа А на основе поведения в ситуации типа В, они будут действовать в ситуации В с учетом ситуации А. Если вера в связь между частной и общественной моралью широко распространена (и это известно), это создает для политиков стимулы порядочно вести себя в личной жизни, предполагая, что любой недостойный поступок станет известен электорату. Предположим, что существует распространенное убеждение, что у людей одинаковый коэффициент дисконтирования во времени во всех ситуациях. Если они не настолько озабочены своим будущим, чтобы заботиться о собственном теле, они также могут (согласно фолк-психологии) нарушить договоренность, связанную с большой краткосрочной выгодой. Следовательно, чтобы давать достоверные обещания относительно долгосрочного взаимовыгодного сотрудничества, человек должен стремиться к стройному и здоровому внешнему виду.

В более широком смысле, однако, *фолк-психология очевидно ошибочна*¹. Если бы можно было устранить эффекты самой фолк-психологии с тем, чтобы не было стимула соответствовать ожиданиям кросс-ситуационной согласованности, такое последовательное поведение встречалось бы реже. Родители, наблюдающие за детьми дома, обычно удивляются, когда узнают, что те гораздо лучше ведут себя в школе или когда бывают в гостях у своих товарищей. Более того, меры по коррекции домашнего поведения не приводят к его улучшению в школе в сравнении с контрольными группами, не испытывающими никакого вмешательства². В лабораторных экспериментах большинство людей (около

¹ Не только фолк-психология. Социологи, утверждающие, что агент сигнализирует своим поведением, что он хороший или плохой тип, также переоценивают постоянство в поведении.

² В этих данных есть один поворот. Дети родителей, следовавших советам детского психолога, лучше успевали в школе, чем дети родителей, которые им не следовали, — факт, который иногда приводят как подтверждение эффекта «перелива» из дома в школу. Но эти результаты могут быть связаны только с тем, что предрасположенность подчиняться правилам является наследственной. У родителей, которые добросовестно следовали инструкциям детского психолога, более вероятно появление детей, которые будут добросовестно следовать инструкциям учителя.

двух третей испытуемых) можно склонить к безжалостному поведению, вплоть до того, что они подвергнут (как они думаю) своих товарищей по эксперименту электрошоку сильным разрядом (около 450 Вольт). И все же нет оснований полагать, что в основе их поведения лежит садизм, жесткость или безразличие к страданиям других людей. На самом деле многие из тех, кто так себя вел, были огорчены и расстроены тем, что делали. Дети с большей охотой ждут отложенного большего вознаграждения, так же как и меньшего, если оно скрыто. Любой академический ученый сразу поймет, что другие ученые добросовестны в своих исследованиях, но с определением добросовестности в преподавании или при выполнении административных задач дело обстоит сложнее. Разговорчивость за обедом, как оказывается, особенно не коррелирует с разговорчивостью в других ситуациях. Человек может тянуть время с уборкой дома, но никогда не откладывает дела на работе³.

В эссе «О непостоянстве наших поступков» Монтень противопоставляет поведение Катона Младшего поведению обычных людей, таких как он. «Взять, к примеру, Катона Младшего: тут тронь одну клавишу — и уже знаешь весь инструмент; тут гармония согласованных звуков, которая никогда не изменяет себе. А что до нас самих, тут все наоборот: сколько поступков, столько же требуется и суждений о каждом из них. На мой взгляд, вернее всего было бы объяснять наши поступки окружающей средой, не вдаваясь в тщательное расследование причин и не выводя отсюда других умозаключений». Он замечает: «Когда человек, падающий духом от оскорбления, в то же время стойко переносит бедность или боящийся бритвы цирюльника обнаруживает твердость перед мечом врага, то достойно похвалы деяние, а не сам человек».

Позвольте привести несколько примеров из области искусства и жизни отдельных творцов. В «Обретенном времени» Пруст писал, что «можно было подумать», что юноши, которым клиенты борделя Жюппена платили за причинение боли, должны были «по сути своей быть плохими, но они не только были прекрасными солдатами во время войны, настоящими героями, они также зачастую были добры и благородны в мирной жизни». Комментируя очевидное противоречие между тем,

³ В письме к этику Ренди Коэну (Randy Cohen) («Нью Йорк таймс мэгэзин». 2006. 15 января) один академический ученый спросил, должен ли тот факт, что не имеющий постоянного контракта (tenure) коллега требовал себе в факультетском клубе скидки, на которые он не имел права, служить основанием, чтобы голосовать против его постоянного контракта из-за «нечестности, которая потенциально может распространяться и на его научные исследования». Коэн ответил «нет» на том основании, что «люди, которые плохо ведут себя в одних ситуациях, часто хорошо себя ведут в других».

как Сван тонко скрывает приглашение в Букингемский дворец, и его похвальбу тем, что жена мелкого чиновника нанесла визит мадам Сван, Пруст писал:

Главной причиной была та (и она действительна для всего человечества), что даже наши добродетели не являются сугубо внешними, свободно парящими сущностями, которые всегда находятся в нашем распоряжении; на самом деле они так тесно связаны в наших умах с теми ситуациями, в которых возникала нужда ими воспользоваться, что необходимость заняться чем-то другим может застать нас врасплох, так что нам даже в голову не придет применить здесь наши добродетели.

Знакомый врач характеризовал джазового музыканта Чарли Паркера как человека, который живет текущим моментом, человека, движимого принципом удовольствия, музыкой, едой, сексом, наркотиками, возбуждением. Развитие его личности [*sic!*] остановилась на инфантильном уровне. У еще одного великого джазового музыканта, Джанго Рейнхарда, была еще более экстремальная сосредоточенность на настоящем в повседневной жизни. Он никогда не откладывал свои заработки, но тратил их на прихоти и дорогие автомобили, которые вскоре разбивал. В некотором отношении он был воплощением стереотипа цыгана. Однако вы не сможете стать музыкантом класса Паркера и Рейнхарда, живя лишь во всех отношениях настоящим моментом. Профессионализм требует долгих лет упорного труда и концентрации. Жизнь Рейнхарда изменилась драматическим образом, когда в пожаре он серьезно повредил левую руку, но в конце концов научился играть двумя пальцами лучше, чем многие играют четырьмя. Будь эти два музыканта вечно беззаботными и импульсивными — если бы их личности сводилась исключительно к инфантильности, они никогда бы не добились такого совершенства в музыке.

После 1945 года норвежский писатель Кнут Гамсун, сотрудничавший во время войны с нацистами, проходил психиатрическое обследование для того, чтобы определить, может ли он предстать перед судом (в тот момент ему было 86 лет). Когда профессор психиатрии попросил его описать свои самые характерные черты, Гамсун ответил следующее:

В период так называемого натурализма Золя и компания писали о людях с характерными чертами. Психологические нюансы были им ни к чему. У людей была одна доминирующая способность, которая управляла их действиями. Достоевский и другие научили нас несколько иному. Я с самого начала думал, что в моих произведениях нет ни одного персонажа с одной преобладающей способностью. Все они без так называемого характера — они противоречивы и фрагментированы, не плохие и не хорошие, одновременно и те и другие. Они полны нюансов, их сознание и поступки изменчивы. И, без сомнения,

я сам такой же. Очень возможно, что я агрессивен и что у меня есть некоторые черты из тех, что указал профессор, — ранимость, подозрительность, эгоизм, благородство, ревность, праведность, логика, чувствительность, холодность. Это все черты, свойственные человеку, и я не могу сказать, что какая-то из них во мне преобладает.

В главе XIV я коснусь темы характера или его отсутствия в художественной литературе. Здесь я замечу, что Гамсун не ссылается на возможность, что он, например, может быть последовательно благородным в одной ситуации и последовательно эгоистичным в другой. Перейдем теперь к этому вопросу.

СИЛА ОБСТОЯТЕЛЬСТВ

Безрассудное отношение к деньгам и преданность музыке, общительность за обедом или исследовательская добросовестность, конечно, являются чертами характера. Но это ситуативные и *локальные черты* (local traits), а не существенные личностные особенности, проявляющиеся всегда и в любых обстоятельствах. В противоположность фолк-психологии, систематические исследования показывают, что уровень кросс-ситуационного постоянства для черт характера очень низок. Хотя корреляции существуют, обычно они столь незначительны, что их нельзя различить «невооруженным глазом». Психопаты могут регулярно демонстрировать безрассудное поведение⁴, а Катон Младший мог неизменно вести себя героически, но от большинства индивидов, которые не относятся к этим крайностям, трудно ожидать такого постоянства. Более радикальное представление фолк-психологии о сочетании всех добродетелей не было тщательно обосновано, возможно, в силу очевидного неправдоподобия. И все же оно оказывает некоторое воздействие на умы, что проявляется в доверии, с которым мы относимся к докторам, обладающим большим врачебным тактом. В классической Античности широкого распространения достигла идея о том, что мастерство в одной области было безошибочным признаком или индикатором превосходства в других. Психологи называют это эффектом ореола (гало-эффектом).

Таким образом, причины поведения часто усматривают в *ситуации*, а не в *человеке*. Рассмотрим факт, что некоторые немцы шли на риск, чтобы спасти евреев от нацистского режима. Сторонники теории ха-

⁴ Поскольку умный эгоист, заботящийся о своем будущем, зачастую будет заинтересован в том, чтобы *имитировать* заботу о других (глава 5), окончательным объяснением психопатического поведения может быть чрезмерное дисконтирование во времени.

рактера предположили бы, что оказывавшие помощь обладали альтруистическим типом характера, к которому не принадлежали те, кто евреев не спасал. Однако, как выясняется, наибольшей объясняющей силой обладает ситуационный факт обращения или не обращения с *просьбой* о спасении кого-либо. Каузальная связь возникает двумя путями. С одной стороны, лишь благодаря обращению человек может получить *информацию* о том, что он должен выступить в роли спасителя; с другой — ситуация личного вопрошания может побудить человека согласиться только потому, что он будет испытывать *стыд* при отказе⁵. Первое объяснение предполагает альтруизм, но отрицает, что его достаточно для объяснения поведения; второе отрицает альтруизм и замещает социальные нормы моральными обязательствами. В обоих случаях тех, кто спасал, от тех, кто этого не делал, отличает ситуация, в которой они оказались, а не их индивидуальные черты.

Дело Китти Дженовезе — еще один пример из реальной жизни, подтверждающий власть обстоятельств. Бездействие не должно подводить нас к выводу о том, что все свидетели убийства были бессердечны и безучастны к человеческим страданиям. Возможно, многие из них думали, что кто-то другой вызовет полицию, или что, если никто не вызывает, дело не столь серьезное, как кажется («Вероятно, всего лишь домашняя ссора»), или что бездействие остальных свидетельствует о небезопасности вмешательства⁶. Такого рода рассуждение тем достовернее, чем больше число случайных свидетелей. Так, в одном эксперименте подсадное лицо симулировало по домофону эпилептический припадок. Когда испытуемые полагали, что это слышат только они одни, то 85 % оказывали помощь. Если они считали, что есть еще один свидетель, то вмешивались 62 %. Если полагали, что слышали еще 4 человека, то 31 %. В другом эксперименте вмешивались 70 % одиноких свидетелей и только 7 % из тех, кто находился рядом с безучастным подставным лицом. При двух наивных участниках жертва получала помощь только в 40 % случаев. Таким образом, с возрастанием количества свидетелей снижается вероятность не только вмешательства *любого* из них, но и того, что вмешаются

⁵ Сходным образом успех телевизионных благотворительных марафонов основан не на апелляции к альтруизму людей, а на том, что к человеку ко всему прочему стучат в дверь и обращаются с непосредственной просьбой. В данном случае объяснение, связанное с информацией, естественно является нерелевантным.

⁶ Люди, которые боялись физически вмешиваться в ситуацию, чтобы защитить женщину от нападавшего, могли позвонить в полицию. Но тогда полиция не принимала анонимных звонков, так что случайные свидетели могли опасаться неприятностей. В других ситуациях возможность *звонка* в полицию могла попросту отсутствовать.

хотя бы *некоторые*⁷. Другими словами, размывание ответственности, вызванное присутствием других, происходило так быстро, что оно компенсировалось увеличением количества людей, которые могли потенциально вмешаться.

В другом эксперименте студентам, изучающим теологию, предлагали самостоятельно подготовиться, чтобы провести короткую беседу в соседнем здании. Половине предложили обсудить притчу о добром самаритянине, остальным дали более нейтральную тему. Одной группе велели поторопиться, так как люди в другом здании их уже ждут, а другой сказали, что у них много времени. По пути в другое здание испытуемым попадался человек, колотивший в дверь и явно находившийся в бедственном положении. Среди студентов, которым велели торопиться, только 10% предложили помощь; в другой группе это сделали 63%. Группа, которой велели подготовить беседу о добром самаритянине, не проявила большей склонности поступать по примеру этого самого самаритянина. Точно так же поведение студентов не коррелировало с ответами в анкете, целью которой было выяснить, в какой мере их интерес к религии связан с желанием обрести спасение души или помочь другим⁸. Ситуационный фактор — спешка или ее отсутствие — обладает гораздо большей объяснительной силой, чем любой диспозиционный фактор.

Было бы неверно привести этот анализ в соответствие с тем, что обсуждалось в предыдущей главе, сказав, что студенты в торопящейся категории вели себя так из-за *ограничения во времени*. Это ограничение не было объективным или жестким; в частности, 10% студентов из этой группы нашли возможность оказать помощь. Поведение определяла скорее ситуация, влиявшая на отчетливость конкурирующих *желаний*. Просьба, высказанная при личной встрече, усиливает действие мотивов, связанных с другими людьми, тогда как требование торопиться его

⁷ Такова тенденция в многочисленных экспериментах такого рода. В упомянутом эксперименте (эпилептический припадок перед домофоном) выяснилось, что если мы предполагаем, что остальные слушатели были реальные, наивные участники, получившие одну и ту же информацию (а не только подсадные или придуманные самим экспериментатором), шанс, что хотя бы один из них вмешается, приблизительно постоянен и составляет 85%. В случае пяти участников (главный участник и четыре слушателя) вероятность, что кто-то из них не станет вмешиваться, составляет 0,69. Шанс, что никто не станет вмешиваться, равен 0,69⁵, или 0,156, что дает вероятность 0,844 вмешательства того, что по крайней мере один человек вмешается.

⁸ Так же как подопытные, которых побудили подвергнуть других людей электрошоку, те, кто прошел мимо человека в бедственном положении, явно очень переживали из-за этой встречи.

уменьшает. Возможность *увидеть* вознаграждение, которое можно получить сразу, делает его более привлекательным, чем отложенное вознаграждение, подобно тому как встреча с нищим на улице скорее вызовет щедрость, чем абстрактное знание о бедности. Случай с Китти Дженовезе меняет восприятие как издержек, так и выгод, которые может принести помощь. Желание следовать инструкциям бесстрастного экспериментатора, требующего продолжать подавать явно болезненные и, возможно, даже фатальные разряды тока, побеждает желание без необходимости не причинять другим людям боль.

Нет основного или общего механизма, посредством которого ситуация предопределяет поведение. Ситуации варьируют от непосредственных просьб спасти евреев до самых тривиальных событий, например когда человек находит 25 центов в отверстие для сдачи платного телефона, что поднимает ему настроение и побуждает помочь незнакомому (а на самом деле подсадному) человеку собрать бумаги, рассыпавшиеся по земле. Важный урок, который нужно вынести из этих наблюдений в реальной жизни и в лаборатории, заключается в том, что *поведение не более стабильно, чем ситуации, которые его формируют*. Человек может быть разговорчив за обедом, когда может расслабиться со своими старыми коллегами, но с незнакомыми людьми может быть замкнут. Человек может постоянно подавать нищим, но больше никогда не задумываться о бедняках. Человек может неизменно оказывать помощь в ситуациях, когда никто больше не может помочь, и вести себя пассивно в присутствии других потенциальных помощников. Мужчина может быть устойчиво агрессивен, делать едкие замечания в адрес жены и при этом быть спокойным и великодушным с другими людьми. Его жена может демонстрировать такую же двойственность. Агрессивность одного вызывает встречную, и наоборот⁹. Если они редко видят, как супруг общается с другими людьми (например, на работе), они могут полагать, что он или она агрессивны по своей природе, а не только в ситуациях, определяемых их собственным присутствием.

СПОНТАННОЕ ОБРАЩЕНИЕ К ДИСПОЗИТИВНОСТИ

В продолжение последнего примера. Специалисты по семейной терапии пытаются заставить обращающихся к ним за помощью супругов переключиться с языка характера на язык действия. Вместо того чтобы гово-

⁹ Говоря метафорически, они находятся в плохом психологическом равновесии. Но агрессия необязательно является наилучшим ответом на агрессию (как того требует понятие «равновесие» из теории игр). Это только психологически интеллигибельный ответ.

рить: «Ты плохой человек», тем самым не оставляя возможностей для надежды или изменения, они должны сказать: «Ты плохо поступил». Последняя формулировка оставляет открытой возможность, что указанное действие было вызвано специфическими ситуационными факторами, такими как провоцирующие замечания со стороны супруга или супруги. Одна из причин (среди многих), почему психологам редко удается преодолеть конфликты, состоит в том, что люди спонтанно отдают предпочтение объяснениям поведения, основанным на характере, а не на ситуации. Если мы узнаем, что кто-то подписался под обращением в защиту прав гомосексуалистов, мы склонны заключить, что это человек — *гомосексуалист* или либерал, а не что его *попросили* так, что трудно было отказать.

Проводя собеседование при приеме на работу, мы склонны объяснять то, что говорит или делает человек в свете той или иной предрасположенности, которую мы (слишком самоуверенно) ему или ей приписываем, а не в свете специфики ситуации собеседования. Сам язык отражает диспозиционный уклон. Определения, которые применяются к действию (*враждебное, эгоистичное или агрессивное*), обычно приложимы и к агенту, тогда как характеристик, которые также применимы к ситуации (*трудный* — это исключение), совсем немного.

Психологи называют неадекватное использование диспозиционного объяснения *фундаментальной ошибкой атрибуции*, то есть объяснением поведения, вызванного ситуацией, устойчивыми чертами характера агента. Когда подопытных попросили предсказать поведение студентов-теологов, столкнувшихся с человеком, попавшим в беду, они (ошибочно) подумали, что люди, чья религия основывается на желании помогать другим, с большей вероятностью будут действовать как добрые самаритяне, предположив (снова ошибочно), что большая или меньшая спешка здесь никакой роли не играют. Участники другого эксперимента прогнозировали более частое применение электрошока в отсутствие специфических ситуационных факторов в первоначальном эксперименте, тем самым показывая веру в диспозиционное объяснение. Когда преподаватель дал студенту задание написать эссе в защиту Кастро, другие, зная о том, как было сформулировано задание, все равно истолковали негативное по содержанию сочинение как положительную характеристику Кастро. Когда студентов попросили добровольно выполнить задания с высоким либо с низким вознаграждением и в результате было получено некое число добровольцев для каждого типа задания, то наблюдатели, зная о разнице в оплате, предсказали, что *все* добровольцы с большей вероятностью, чем недобровольцы, вызовутся выполнять неприбыльные задания. Иными словами, наблюдатели приписывали волонтерский поступок предрасположенности к волонтерству, а не структуре вознаграждений, которой характеризовалась ситуация.

Одни нации в меньшей степени, чем другие, склонны к фундаментальной ошибке атрибуции. Эксперименты показывают, что по сравнению с американцами азиаты при объяснении поведения придают меньшее значение характеру человека. Это различие проявляется и в реальных жизненных ситуациях. Так, в 1991 году потерпевший неудачу китайский студент-физик застрелил своего научного руководителя, нескольких соучеников, а потом застрелился сам. В том же году американский почтовый служащий, потерявший работу, застрелил своего начальника, нескольких коллег и случайных прохожих, а потом выстрелил в себя. Оба события получили широкий отклик в американской и китайской прессе: американская объясняла их в диспозиционных категориях («психически нездоровый», «психически неустойчивый», «тяжелый характер»), а китайская — с ситуационной точки зрения («доступность оружия», «потерял работу», «жертва навязанного стремления к успеху»). Другие данные подтверждают это различие. Это, однако, может быть связано с тем, что ситуационные факторы действительно имеют большее значение в детерминировании поведения азиатов. Они не столько более успешны в преодолении диспозиционного соблазна, сколько он в меньшей степени им свойственен, или же действуют оба фактора.

Преодоление фундаментальной ошибки атрибуции может принести раскрепощение. Студенты-первокурсники, которым сказали, что на первом курсе обычно учатся плохо, но потом оценки повышаются, и в самом деле лучше учатся на последующих курсах, чем те, кому этого не сообщили. Последние в большей степени склонны связывать плохие оценки со своими низкими способностями, чем с незнакомой и отвлекающей внимание обстановкой колледжа. Будучи уверены, что они не могут учиться лучше, они теряют мотивацию для улучшения результатов. Когда угнетенные группы отбрасывают веру в эссенциализм своих угнетателей, то есть идею того, что женщины, темнокожие или евреи по природе своей ниже остальных, им легче сбросить оковы.

Является фундаментальная ошибка атрибуции горячей или холодной, то есть это мотивированная ошибка или скорее нечто подобное оптической иллюзии? В той степени, в какой мотивация входит в процесс атрибуции, нет причины, по которой она должна постоянно приводить к переоценке значения характера. В своекорыстных интересах нам следовало бы приписывать успех устойчивым чертам нашего характера, а неудачи — трагическому стечению обстоятельств¹⁰. Если верить французским моралистам, мы должны относить успех других на счет их удач-

¹⁰ Иногда, впрочем, мы приписываем свои неудачи нашему характеру. Игрок или алкоголик может вполне счастливо уверять себя в том, что «это сильнее его», чтобы иметь оправдание продолжениям игр или возлияний.

ливости, а неудачи — на счет характера¹¹. С когнитивной точки зрения тенденция отдавать предпочтение человеку перед ситуацией может быть частным случаем того, что человек обращает больше внимания на движущийся передний план, чем на статический задний. Отсюда следует, что данная ошибка должна реже встречаться в культурах, где внимание более равномерно распределяется между планами (это, похоже, происходит в азиатских культурах).

РЕАБИЛИТАЦИЯ ЛИЧНОСТИ

Результаты исследований, которые я описал, подрывают то, что можно назвать грубым эссенциализмом в изучении личности. Неверно, что люди все время *являются* агрессивными, нетерпеливыми, экстравертами или разговорчивыми. В то же время эти результаты не означают, что понятие ситуации всемогуще в объяснении поведения. Мы должны разложить характер на ряд *случайных* реактивных тенденций. Вместо того чтобы описывать человека как альтруиста, мы можем сказать о ней или о нем: «помогает, когда об этом просят, но сам помочь не вызывается» или: «помогает при отсутствии стресса, в стрессовой ситуации становится невнимательным». Каждая из этих фраз может характеризовать один аспект человека и тем самым поддерживать более тонкую форму эссенциализма. Можно отчитывать супруга за то, что он никогда не убирает («Лентяй!»), или за то, что он занимается уборкой только когда попросят («Безответственный!»). В последнем случае супруг может быть проактивным скорее, чем реактивным, в других вопросах, например, он может следить за здоровьем детей в семье. Здесь не будет наблюдаться постоянная черта реактивности.

В такой перспективе объяснение основывается на отдельной ситуации, а также на специфическом для данного человека отношении между ситуацией и поведенческими наклонностями. Один человек может быть агрессивен с людьми, над которыми у него есть власть, и при этом исключительно дружелюбен с теми, кто имеет власть над ним, тогда как другой может демонстрировать противоположную модель поведения. Если мы наблюдаем, как оба ведут себя дружелюбно, мы можем заключить, что они *имеют* дружелюбный нрав. Но, как уже стало понятно, сходство поведения может быть вызвано различием в ситуации и в случайностях реакции, которые могут в точности взаимоаннулировать.

¹¹ Иногда, впрочем, приписываем успехи других людей *дефектам* их характера. Антисемитизм полагается на миф о том, что евреи добились успеха, потому что их аморальный характер позволяет им не стесняться в средствах для достижения цели.

Библиографические примечания

Тенденция к переоценке единства личности была четко описана в статье Г.Ичхайзера «Недоразумения в человеческих отношениях: Исследование ложного социального восприятия» (*Ichheiser G. Misunderstandings in human relations: A study in false social perception // American Journal of Sociology. 1949. No. 55. Supplement*). Последние работы, переносящие акцент с «характера», ведут свою родословную от книги У.Мишеля «Личность и ее оценка» (*Mischel W. Personality and Assessment. New York: Wiley, 1968*). Многие из изложенного выше основывают Л.Росс и Р.Нисбетт в работе «Человек и ситуация» (*Росс Л., Нисбетт Р. Человек и ситуация. Уроки социальной психологии. М.: Аспект-Пресс, 2000*) и Дж.Дорис «Бесхарактерность» (*Doris J. Lack of Character. Cambridge University Press, 2002*). Примеры с коммунистическими лидерами и мафиози взяты соответственно из книги С.Попкина «Рациональный крестьянин» (*Popkin S. The Rational Peasant, Berkeley: University of California Press, 1979*) и работы Д.Гамбетты «Неисповедимые пути доверия» (*Gambetta D. Trust's odd ways // Elster J. et al. (eds). Understanding Choice, Explaining Behavior: Essays in Honour of Ole-Jørgen Skog. Oslo: Academic Press, 2006*). Отсылка к эффектам вмешательства взята из Дж.Р.Харрис «Все разные» (*Harris J. R. No Two Alike. New York: Norton, 2006*). Готовность подвергнуть других людей электрошоку описана в классическом исследовании С.Милграма «Подчинение власти» (*Milgram S. Obedience to Authority. New York: Harper, 1983*). Информация о двух музыкантах почерпнута из книг «Жизнь птахи. Веселая жизнь и тяжелые времена Чарльза Паркера» Р.Рассела (*Russell R. Bird Lives: The High Life and Hard Times of Charlie (Yardbird) Parker. New York: Charterhouse, 1973*) и «Джанго: жизнь и музыка цыганской легенды» М.Дрегни (*Dregni M. Django: The Life and Music of a Gypsy Legend. Oxford University Press, 2004*). Заявление Гамсуна переведено из книги Г.Лангфельдт и О.Одегард «Судебно-психиатрическое заявление Кнута Гамсуна» (*Langfeldt G., Ødegård Ø. Den rettspsykiatriske erkloerengen om Knut Hamsun. Oslo: Gyldendal, 1978. P. 82*). Использование поведения в качестве индикатора в Античности обсуждается в книге П.Вена «Хлеб и зрелища» (*Veune P. Le pain et le cirque. Paris: Seuil, 1976. P. 114, 773*). См. также статью П.Вейна «Почему хотят, чтобы у государя были частные добродетели?» (*Veune P. Pourquoi veut-on qu'un prince ait des vertus privees? // Social Science Information. 1998. No. 37. P. 407–15*). «Характерологическое» объяснение готовности спасать евреев приводится в статье «Альтруизм и теория рационального действия. Спасители евреев в нацистской Европе» К.Монро, М.К.Бартона и У.Клингеманна (*Monroe K., Barton M. C., Klingemann U. Altruism and the theory of rational action: Rescuers of Jews in Nazi Europe // Ethics. 1990. No. 101. P. 103–122*). Ситуационное объяснение обосновывается в статье «Важно, чтобы тебя спросили. Спасение евреев в нацистской Европе» Ф.Варезе и М.Яйша (*Varese F., Yaish M. The importance of being asked: The rescue of Jews in Nazi Europe // Rationality and Society. 2000*).

No. 12. P. 307–324). В анализе случая с Китти Дженовезе, с точки зрения теории игр, А. Диксит и С. Скет, авторы «*Стратегических игр*» (Dixit A., Skeath S. *Games of Strategy*. 2nd ed. New York: Norton, 2004. P. 414–418), утверждают, что когда количество потенциальных участников растет, шансы того, что кто-нибудь из них действительно вмешается, падают. Скептическое замечание о склонности делать выводы о чертах характера на основании поведения прозвучало в «*Подозрении и диспозиционных выводах*» Дж. Л. Хилтона, С. Фейна и Д. Миллера (Hilton J. L., Fein S., Miller D. Suspicion and dispositional inference // *Personality and Social Psychology Bulletin*. 1993. No. 19. P. 501–12). Различие американцев и азиатов описано в «*Географии мысли*» Р. Нисбетта (Nisbett R. *The Geography of Thought*. New York: Free Press, 2004). То, что я называю «реабилитация личности», описывается в работе «*К интегративной науке о личности*» У. Мишеля (Mischel W. *Towards an integrative science of the person*. *Annual Review of Psychology*. 2004. No. 55. P. 1–22).

XI. Рациональный выбор

СТРУКТУРА РАЦИОНАЛЬНОГО ДЕЙСТВИЯ

Теоретики рационального выбора в качестве базового допущения при объяснении поведения выдвигают рациональность агентов. Это допущение включает гипотезу, что агенты формируют рациональные убеждения, включая представление о имеющихся у них вариантах выбора. Следовательно, нет необходимости различать субъективные (желания) и объективные (возможности) детерминанты поведения. Теория рационального выбора субъективна от начала до конца.

Структура объяснения при помощи рационального выбора представлена на рис. XI.1. Действие рационально, если оно отвечает *трем требованиям оптимальности*. Действие должно быть оптимальным, учитывая убеждения; убеждения должны быть как можно лучше обоснованы с опорой на факты; факты должны быть добыты в результате оптимальных вложений в сбор информации. Стрелки имеют двойное толкование с точки зрения как каузальности, так и оптимальности. Действие, например, должно быть вызвано желаниями и убеждениями, которые делают его рациональным; недостаточно совершить правильный поступок только благодаря счастливому стечению обстоятельств. Подобным образом убеждение не является рациональным, если оно является результатом двух противоположно направленных процессов, которые отменяют друг друга.

Приведем пример: курильщики и те, кто не курит, обрабатывают информацию об опасностях курения таким образом, что полагают, что эти

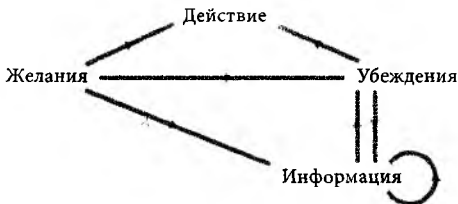


РИС. XI.1

опасности выше, чем на самом деле. В то же время курильщики склонны к своекорыстной тенденциозности, заставляющей их дисконтировать риски. Если в результате у них сформируется такое же убеждение, как у непредвзятого наблюдателя¹, это не доказывает, что они рациональны. В одном из самых влиятельных обсуждений рациональности в социальных науках Макс Вебер ошибочно выводил процессуальную рациональность из результирующей оптимальности, когда писал:

Для типологического научного исследования все иррациональные, эмоционально обусловленные смысловые связи, определяющие отношение индивида к окружающему и влияющие на его поведение, наиболее обозримы, если изучать и изображать их в качестве отклонений от чисто целерационально сконструированного действия. Так, для объяснения «биржевой паники» целесообразно сначала установить, каким было бы рассматриваемое поведение без влияния иррациональных аффектов, а затем ввести эти иррациональные компоненты в качестве «помех». Равным образом и при исследовании какой-либо политической или военной акции целесообразно установить, каким было бы поведение участников события при знании ими всех обстоятельств дела, всех намерений и при строго целерационально (в соответствии со значимым для нас опытом) ориентированном выборе средств. Лишь в этом случае возможно свести отклонения от данной конструкции к обусловившим их иррациональным факторам.

Хотя Вебер был прав, когда полагал, что отклонение от рационального хода событий — достаточное условие для того, чтобы заработала иррациональность, он ошибался, утверждая (в фразе, которую я процитировал), что оно является необходимым. Похожую ошибку совершают те, кто утверждает, что реакции инстинктивного страха рациональны, тогда как можно только сказать, что они являются *адаптивными*. Когда я вижу на тропинке нечто похожее на змею или на палку, следует сразу убежать, а не заниматься сбором дополнительной информации. Кажется, что люди жестко запрограммированы на такие действия. Такой побег не является в строгом смысле слова рациональным, поскольку он не вызван механизмом рационального принятия решений, и все же он имитирует рациональность в том смысле, что, будь названный механизм задействован в данной ситуации, он привел бы к такому же действию. Когда издержки, связанные со сбором информации (глава XII), велики, рациональный агент не будет собирать ее в большом количестве. Однако тенденция убежать зачастую вызвана не этими расчетами, но предшествует им².

¹ На самом деле вторая предвзятость курильщиков не полностью компенсирует первую.

² У крыс задержка между бездумной реакцией и реакцией рефлексивной составляет доли секунды.

ПРЕДПОЧТЕНИЯ И ПОРЯДКОВАЯ ПОЛЕЗНОСТЬ

При тщательном анализе становится ясно, что первое требование оптимальности заключается в том, что действие должно быть наилучшим средством удовлетворения желаний агента с учетом его представлений об имеющихся вариантах выбора и их последствиях. Что такое «наилучшее», определяется в категориях предпочтительности или предпочтения: наилучшее — то, лучше чего нет с точки зрения агента. Из этого не следует, что желания *эгоистичны*. Грубая ошибка — путать рациональность с эгоизмом, хотя этому способствует практика некоторых теоретиков рационального выбора. Точно так же нет необходимости в *стабильности* желаний, даже в минимальном смысле исключения изменений предпочтения во времени. Агент, который под влиянием эмоций или наркотиков выбирает *А* вместо *В*, действует рационально, выбрав *А*, даже если при других обстоятельствах он выберет *В*, а не *А*. Подобный случай (см. главу VI) происходит, когда значение, которое агент придает будущим последствиям настоящего выбора, в результате указанных влияний уменьшается.

Для начала следует дать хорошее определение «наилучшему». Его обеспечивают два условия³. Во-первых, предпочтения должны быть *транзитивными* (transitive). Предположим, что есть опции *А*, *В* и *С*. Если человек полагает, что *А* по меньшей мере не хуже, чем *В*, а *В* по меньшей мере не хуже, чем *С*, то он также должен полагать, что *А* не хуже, чем *С*. Если транзитивность не работает, например, если человек предпочитает строго *А*, но не *В*; *В*, но не *С*; *С*, но не *А*, у него может и не быть наилучшего выбора. Более того, другой человек может воспользоваться этим фактом, предложив агенту перейти от менее предпочтительного к более предпочтительному выбору в обмен на деньги. Поскольку предпочтения цикличны, эта операция может повторяться бесконечно, приведя человека к катастрофе путем пошаговых улучшений⁴.

Такая ситуация может возникнуть, если агент ранжирует варианты путем подсчета параметров (counting aspects). Предположим, я предпочитаю одно яблоко другому, если оно лучше по крайней мере по двум из трех параметров, а именно цене, вкусу и склонности к порче. Если яблоко *А* превосходит *В* по цене и вкусу, яблоко *В* превосходит *С* по цене

³ Я игнорирую техническое условие, согласно которому ряд возможностей должен быть компактным и закрытым. Компактность нарушается, если для каждого осуществимого выбора имеется один наилучший вариант, а закрытость — если ограничение последовательности осуществимых вариантов само по себе не является осуществимым (feasible).

⁴ Касательно еще одного способа «улучшить себя до смерти» см. примечание 1 к главе XVIII.

и склонности к порче, а яблоко *C* превосходит *A* по вкусу и склонности к порче, транзитивность нарушается. Хотя такая вероятность относительно неважна при индивидуальном выборе, где она отражает лишь несовершенство практического способа оценки, мы увидим (глава XXV), что она более значима при коллективном выборе.

Иная проблема возникает, когда безразличие не может стать транзитивным. Мне может быть безразлично, предпочесть *A* или *B* либо *B* или *C*, потому что различия в каждой паре слишком незаметны, но я могу остановить выбор на *C*, а не на *A*, так как разница между ними существенна. Есть вариант, который является наилучшим, а именно *C*, но агента все еще можно склонить к худшему выбору, сделав ему серию предложений (обменять *C* на *B* и *B* на *A*), от которых у него нет причин отказываться и которые он, соответственно, может принять. Применение к агенту с нетранзитивными предпочтениями определения *иррациональный* оправдывает не столько отсутствие наилучшего варианта, сколько тот факт, что он может принять предложения, которые ухудшат его положение.

Чтобы удостовериться в неизменной осмысленности идеи наилучшего выбора, можно также потребовать, чтобы предпочтения были *полными* (complete): для любых двух результатов агент должен иметь возможность определиться, предпочитает он первый или второй вариант, второй или первый либо они оба ему безразличны. Если он не в состоянии ответить ни на один из этих вопросов, он, возможно, не в состоянии определить, какой вариант является наилучшим. Я подробнее остановлюсь на неполноте в конце главы. Здесь я хочу отметить, что, в отличие от отсутствия транзитивности, отсутствие полноты не означает неудачу. Предположим, я хочу дать мороженое одному из двух детей, который получит от него наибольшее удовольствие. Чтобы сформировать *мое* предпочтение, мне нужно иметь возможность сравнить *их* уровни удовлетворенности, когда они получают мороженое. Однако часто это невыполнимая задача. Моя неспособность сделать это не является неудачей в том смысле, что я мог бы поступить правильнее, но просто отражает житейский факт.

Во многих случаях транзитивность и полнота предпочтений — это все, что нам требуется, чтобы определить рациональный выбор. Однако иногда удобно представить предпочтения численно, что часто называют значениями полезности (utility values), присваиваемыми вариантам выбора. Чтобы обеспечить такую возможность, мы налагаем на предпочтение еще одно условие — *непрерывность* (continuity). Если каждый вариант в последовательности A_1, A_2, A_3, \dots предпочтительнее, чем *B*, а последовательность стремится к *A*, тогда *A* предпочтительнее, чем *B*. Если *B* предпочтительнее, чем любой вариант в последовательности, *B* предпочтительнее, чем *A*. Контрпримером может послужить лексико-графическое предпочтение: набор из благ *A* и *B* в количествах (A_1, B_1) предпочтительнее дру-

ного набора (A_2, B_2) , тогда, и только тогда, когда либо $A_1 > A_2$ либо $(A_1 = A_2$ и $B_1 > B_2)$. В таком ранжировании предпочтений наборы $(1,1; 1)$, $(1,01; 1)$, $(1,001; 1)$..., предпочтительнее $(1; 2)$, который предпочтительнее $(1; 1)$. Можно сказать, что первый компонент набора благ несравнимо важнее второго, потому что никакое дополнительное количество блага B не может компенсировать малейшую потерю блага A ⁵. Проще говоря, компромисс невозможен. Таким образом, эти предпочтения не могут быть представлены кривыми безразличия. Если лексико-графические предпочтения редко применимы к обычным потребительским благам, они могут иметь значение при политическом выборе. Избиратель может предпочесть кандидата A кандидату B , потому что у первого более четкая позиция по вопросу о запрете абортов *либо* если у них одинаковая позиция по этому вопросу и A предлагает более низкие налоги, чем B . Для таких избирателей мирская ценность денег ничто в сравнении со священным даром жизни.

Если предпочтения агента полные, транзитивные и последовательные, мы можем представить их функцией непрерывной полезности u , которая присваивает число $u(A)$ каждому варианту (A) . Вместо того чтобы говорить, что рациональный агент выбирает наилучший осуществимый вариант, мы можем сказать, что агент *максимизирует полезность*. В данной фразе *полезность* — всего лишь условное обозначение для предпочтений с определенными свойствами. Чтобы это понять, можно отметить, что единственное требование для функции u , чтобы она могла представить порядок предпочтений, заключается в том, что A предпочтительнее B тогда, и только тогда, когда $u(A) > u(B)$. Если u всегда положительное, $v = u_2$ также может представить тот же порядок предпочтений, хотя v присваивает большие или (для < 1) меньшие значения, чем u . Абсолютные числа значения не имеют, только их относительные или *порядковые* (ordinal) величины. Таким образом, идея максимизации полезности не предполагает, что агент занят добычей максимально возможного количества некоего психического вещества. Она, однако, исключает такие иерархии ценностей, которые воплощены в лексико-графических предпочтениях. Они не могут быть представлены функцией полезности.

КОЛИЧЕСТВЕННАЯ ПОЛЕЗНОСТЬ И ОТНОШЕНИЕ К РИСКУ

Часто агенты сталкиваются с *рискованными* альтернативами, то есть с выбором, который при известных вероятностях может иметь более одного возможного варианта исхода. Может показаться, что рацио-

⁵ Таким образом, интуитивное представление о несопоставимости может быть расписано двумя способами: как неполное предпочтение или как непоследовательное предпочтение.

нальный агент выберет вариант с наибольшей *ожидаемой полезностью*, включающей как выигрыш от каждого исхода, так и вероятность его выпадения. Сначала для каждого варианта агент проведет оценку полезности каждого следствия с точки зрения его вероятности, суммирует все результаты, а затем выберет вариант с максимальной суммой.

Однако порядковая полезность *не позволяет* нам разъяснить эту идею. Предположим, есть два варианта — *A* и *B*. *A* может произвести результат O_1 или O_2 с вероятностью $\frac{1}{2}$ и $\frac{1}{2}$, тогда как *B* может дать результат O_3 или O_4 с вероятностями $\frac{1}{2}$ и $\frac{1}{2}$. Возьмем теперь функцию полезности *u*, которая присваивает значения 3, 4, 1 и 5 для соответственно O_1 , O_2 , O_3 , O_4 . Ожидаемая порядковая полезность для *A* составляет 3,5, а для *B* — 3. Если мы вместо этого будем использовать функцию $v = u^2$, значения будут составлять 12,5 и 13. И та и другая функции представляют предпочтения, но они выделяют в качестве наилучших разные варианты. Ясно, что такой подход бесполезен.

Можно добиться лучшего результата, но с некоторыми концептуальными издержками. Подход, который связывают с именами Джона фон Неймана (John von Neumann) и Оскара Моргенштерна (Oskar Morgenstern), показывает, что вариантам можно присваивать значения полезности, которые имеют количественное (или кардинальное), а не только порядковое (или ординальное) значение. Примером присваивания количественных значений может служить температура. Меряем мы температуру в градусах по Цельсию или по Фаренгейту, это не влияет на истинность высказывания «Средняя температура в Париже выше, чем средняя температура в Нью-Йорке» (если бы температура измерялась порядковым образом, высказывание не имело бы смысла). Истинность же высказывания «В Париже в два раза жарче, чем в Нью-Йорке», напротив, *зависит* от выбора шкалы измерения. Однако хотя истинность конкретных высказываний об интенсивностях зависит от шкалы измерения, для других высказываний это не так. Истинность высказывания «Разница температур между Нью-Йорком и Парижем больше, чем между Парижем и Осло», например, не зависит от выбора системы измерения. Сходным образом мы можем сконструировать количественные меры полезности, которые отражают (помимо прочего, как мы увидим) интенсивность предпочтений, а не только порядковое ранжирование вариантов. Они дают нам возможность сравнить выигрыш (или проигрыш) в полезности при переходе от *x* к $(x+1)$ и от $(x+1)$ к $(x+2)$, то есть позволяют говорить о возрастающей или снижающейся предельной полезности (*marginal utility*), лишенных смысла при измерении порядковой полезности.

Технические детали такого конструирования нас касаться не должны, поскольку для настоящих целей достаточно простой базовой идеи.

Мы начнем с допущения, что у агентов есть предпочтения по отношению не только к вариантам выбора, но к лотереям (lotteries) вариантов, включая выродившуюся лотерею (degenerate lottery), состоящую в безусловном получении базового выигрыша. Для любого конкретного набора вариантов или призов лотерея определяет вероятность получения каждого из них с суммарной вероятностью, равной 1. Предполагается, что агенты в таких лотереях имеют полные и транзитивные предпочтения. Кроме того, считается, что для предпочтений действует аксиома независимости: на предпочтения, оказываемые лотереям p и q , не влияет их комбинирование одним и тем же образом с третьей лотереей r . Эффект определенности, о котором говорилось в главе VII и который будет еще обсуждаться в главе XII, нарушает эту аксиому.

Наконец, предполагается, что предпочтения демонстрируют форму непрерывности, определяемую следующим образом. Предположим, базовые варианты включают лучший элемент A и худший элемент B . Мы произвольно присваиваем им значения полезности 1 и 0. Непрерывность означает, что для любого промежуточного варианта C есть вероятность $p(C)$, которая делает агента безразличным к тому, получить C наверняка или поучаствовать в лотерее, которая даст ему A с вероятностью $p(C)$ и B с вероятностью $1 - p(C)$ ⁶. Затем мы принимаем *кардинальную полезность* $u(C)$ равной $p(C)$. Это число является, конечно, произвольным, потому что такими же являются конечные полезности. Предположим, мы присвоим значения M и N соответственно элементам A и B ($M > N$). Определяем полезность C как ожидаемую полезность лотереи

$$pM + (1 - p)N = Mp - N - Np = (M - N)p + N.$$

Полученный класс функций полезности гораздо меньше, чем класс функций порядковой полезности⁷. Легко видеть, что если вариант X имеет большую ожидаемую полезность, чем Y , в соответствии с одной функцией, то это же он будет иметь и в соответствии с другой. Так, мы можем с уверенностью утверждать, что рациональный агент максимизирует ожидаемую полезность.

Функции количественной полезности имеют важное свойство *линейности по вероятности*. Давайте введем обозначение XpY , означающее лотерею, в которой предлагается вероятность p получения X и $(1 - p)$ по-

⁶ Выявление такой вероятности поднимает проблему укорененности, на которую указывалось во введении во второй части.

⁷ Две любые такие функции на самом деле относятся друг к другу, как шкалы температур Цельсия и Фаренгейта, которые присваивают разные значения (соответствующие в нашем случае величинам M и N) температуре кипения воды.

лучения Y . При использовании шкалы с конечной точкой 1–0 полезность $u(X)$ равна вероятности q , при которой агенту безразлично, выбрать X или лотерею AqB . Точно так же полезность $u(Y)$ равна вероятности r , при которой агенту безразлично, выбрать Y или лотерею ArB . XpY , таким образом, предлагает полезность, эквивалентную шансу p на получение A с вероятностью q и шансу $(1-p)$ на получение A с вероятностью r . Таким образом, полезность XpY составляет $pq + r(1-p)$, то есть p раз полезность X плюс $(1-p)$ раз полезность Y . Например, полезность вероятностной комбинации 3 шансов из 5 получить X и 2 из 5 получить Y равна $\frac{3}{5}q + \frac{2}{5}r$.

Здесь возможны следующие возражения. Предположим, у фермера есть выбор между двумя культурами: традиционной разновидностью, которая с равной вероятностью может дать хороший или средний урожай в зависимости от погоды, и современной, которая с равной вероятностью может дать отличный или плохой урожай. Предположим, количественная полезность 3 и 2 для более старой культуры, 5 и 1 для более новой. Поскольку ожидаемая полезность новой культуры больше, то фермер должен выбрать ее. Но — и в этом суть возражения — разве при этом не упускается из виду то, что фермер может не иметь склонности к риску и неохотно примет любой вариант, допускающий настолько низкий уровень полезности, как 1? Возражение, однако, включает двойной счет, так как нежелание рисковать уже инкорпорировано в построение количественных полезностей. При допущении, что A , B и C получают значения 100, 0 и 60, $u(C)$ вполне может составить 0,75 для человека, не склонного к риску, предполагая, что ему безразлично, получить 60 наверняка или участвовать в лотерее, в которой у него 25%-й шанс ничего не получить и 75%-й — получить 100. Подобный аргумент применим к присваиванию значений количественной полезности физическим объемам урожая.

В качестве еще одного примера рассмотрим получения разрешения опеки над ребенком (см. рис. XI.2). Горизонтальная ось может быть понята двояко — как предполагающая физическое разделение опеки (процент времени, проведенного с ребенком) либо вероятностное разделение (шансы получить по суду полную опеку). Количественная полезность равного разделения времени — AE , что больше, чем полезность AC 50%-й вероятности полной опеки. (Здесь мы апеллируем к тому факту, что количественная полезность линейна по вероятности.) Причина в том, что большинство людей в подобной ситуации демонстрируют нежелание рисковать. Они соглашаются на совместную опеку, потому что 50%-й риск не иметь возможности видеть ребенка слишком не приемлем. Только если родитель полагает, что его или ее шансы получить полную опеку выше, чем q процентов, судебное разбирательство стано-

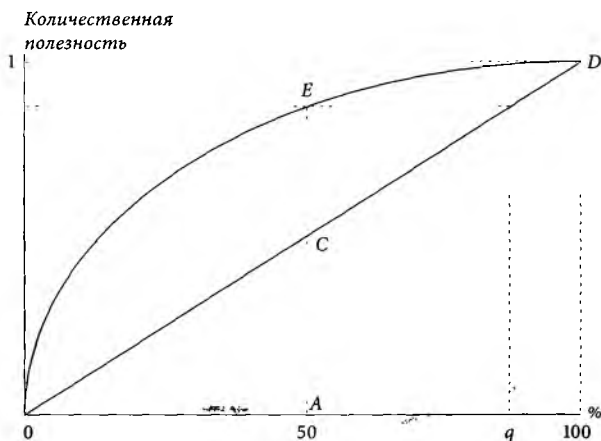


РИС. XI.2

вится предпочтительнее совместной опеки. Значительное число разбирательств об опеке в судах свидетельствует не о склонности родителей к риску, а о принятии ими желаемого за действительное, выражающемся в переоценке своих шансов на получение полной опеки над ребенком.

НЕПРИЯТИЕ РИСКА И УБЫВАЮЩАЯ ПРЕДЕЛЬНАЯ ПОЛЕЗНОСТЬ

Несмотря на верность общего хода рассуждений, предшествующее изложение может направить по неверному пути. Некоторые работы на эти темы склонны стирать различие между неприятием риска и уменьшающейся предельной полезностью. Для развития этой идеи мне потребуется ввести концепцию, которая, как подсказывает интуиция, имеет смысл, хотя ее (до сих пор) было нелегко измерить. Это идея *внутренней полезности* (intrinsic utility) некоего блага, отражающая интенсивность предпочтений агента. Интроспекция со всей очевидностью показывает, что некоторые блага или навыки дают огромное наслаждение, другие — простое удовлетворение, еще одни вызывают легкое раздражение, а какие-то просто ужасны. Очевидно, что представлять различия между ними в категориях количественных предпочтений («Я предпочитаю рай аду точно так же, как я предпочитаю четыре яблока трем») — значит пользоваться крайне обедненным представлением о благополучии или пользе. Тот факт, что у нас нет надежного способа присваивания число-

Действительная
полезность

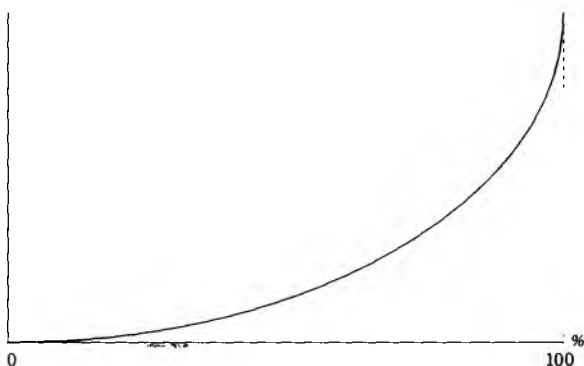


РИС. XI.3

вых значений внутренним уровням удовлетворенности или неудовлетворенности, не свидетельствует о безнадежности самой идеи. В равной мере наша неспособность квантифицировать и сравнивать уровни удовлетворенности разных индивидов не доказывает, что идея сравнения межличностного благополучия лишена смысла.

В этой перспективе может быть понята идея убывающей предельной полезности многих благ. Для бедного человека первые доллары имеют большую полезность, но каждый следующий доллар становится менее ценным с субъективной точки зрения. Каждый курильщик знает, что самая лучшая сигарета — первая, утренняя, что от курения получаешь больше удовольствия, если сдерживаешь себя и не куришь слишком часто. По сути выкуривание сигареты имеет два эффекта: производство удовольствия в настоящем и уменьшение удовольствия от сигарет, которые будут выкурены в будущем.

Второй эффект, однако, необязательно должен быть негативным. Обратимся еще раз к случаю с опекой над ребенком. Полдня, проведенные с ребенком один раз в две недели, могут принести родителю не столько радость, сколько фрустрацию. Полдня каждые выходные приносят в два раза большее удовольствие, потому что укрепление эмоциональной связи, благодаря более частым встречам, делает их ценными для обеих сторон. На другом конце временного спектра сверхудовлетворение от пребывания с ребенком в течение семи дней в неделю превосходит соответствующее субъективное восприятие пребывания с ним шесть дней и т. д.

за счет преимущества неограниченного планирования, которую обеспечивает полная опека. Фактически общение с ребенком имеет возрастающую предельную (действительную) полезность, как показано на рис. ХІ.3.

Горизонтальная ось интерпретируется как процент времени, проведенного с ребенком. По указанной выше причине каждый дополнительный час более ценен, чем предшествующий. *Это утверждение прекрасно совместимо с анализом, на котором основан рис. ХІ.2.* Предельная полезность времени, проведенного с ребенком, может снижаться, если понимается как количественная полезность, но возрастать — если понимается как действительная. Из факта, что только первая часть этого утверждения может быть измерена, не следует, что вторая его часть лишена смысла.

Хотя функции количественной полезности всегда порождаются двумя лежащими в их основе психологическими факторами — отношением к риску и действительной полезностью, их нельзя измерять раздельно. Строго говоря, мы не можем знать, является кривая *OED* на рис. ХІ.2 производной от нейтральности риска в сочетании со снижающейся предельной действительной полезностью времени, проведенного с ребенком, или от неприятия риска в сочетании с возрастающей предельной внутренней полезностью этого времени. В конкретном случае интуиция может подсказать нам, насколько та или иная интерпретация более правдоподобна. Некоторые родители относятся ко времени, проведенному с детьми, так же, как бабушки и дедушки, — небольшими дозами неплохо, но в больших количествах утомляет. В то же время такие родители могут не слишком беспокоиться о риске полного отсутствия общения с ребенком (нейтральное отношение к риску). Другие родители могут отличаться в обоих отношениях, порождая такие же функции количественной полезности. Повторю снова и снова: такие утверждения не могут быть строгими (пока), но очевидно, что они имеют смысл⁸.

РАЦИОНАЛЬНЫЕ УБЕЖДЕНИЯ

Мы заканчиваем обсуждение первого компонента рационального выбора — выбор наилучших средств для осуществления своих желаний с учетом убеждений. Ясно, что это только необходимое, но недостаточное условие для рациональности. Если я хочу убить соседа и считаю, что лучший способ кого-то убить — сделать его куклу и воткнуть в нее булавку, я буду действовать рационально (в том, что касается первого ком-

⁸ Таким образом, аналогия со шкалой температур верна только отчасти. Эти шкалы измеряют *только* интенсивность температуры. Функции количественной полезности измеряют общий результат интенсивности предпочтений и неприятия риска.

понента), если сделаю куклу соседа и воткну в нее булавку. Но, за исключением особых обстоятельств, такое убеждение вряд ли рационально⁹.

Рациональные убеждения формируются посредством обработки имеющихся данных с использованием процедур, которые в среднем должны с наибольшей вероятностью формировать истинные убеждения. Предположим, мы хотим получить представление о вероятности дождя 29 ноября, то есть ровно через неделю. Возможно, нам придется всего лишь просмотреть статистику выпадения осадков в предшествующие годы и предположить, что (ожидаемое) будущее будет похоже на прошлое. Но по мере приближения 29 ноября настоящая ситуация с осадками может заставить нас изменить ожидания. Если в ноябре часто идет дождь, а небо день за днем остается безоблачным, мы можем сделать вывод, что существует фронт высокого давления, который делает дождь 29 ноября менее вероятным.

Процесс пересмотра убеждений часто называют *обучением Байесовым методом* (названным так по имени священника XVIII века Томаса Байеса). Предположим, что мы имеем первоначальное (предварительное) субъективное распределение вероятностей, касающихся различных положений вещей. В приведенном примере предварительное распределение было выведено из частотности в прошлом. В других случаях это может быть просто догадка. На основании своей интуиции я, например, могу присвоить 60%-ю вероятность тому, что премьер-министр компетентен. Мы можем наблюдать за действиями, которые он предпринимает на должности, и за их результатами, например коэффициентом роста экономики. Предположим, мы можем сформировать оценку вероятности этих наблюдений с учетом компетентности премьер-министра. При компетентном премьер-министре у нас 80%-е ожидание хорошего результата, при некомпетентном — 20%. Байес показал, как скорректировать наши первоначальные вероятности, касающиеся компетентности премьер-министра, с учетом наблюдений.

Предположим, что есть только два возможных исхода — хороший и плохой, и что мы наблюдаем хороший. Если мы присваиваем $p(a)$ вероятности, что будет получено a , и $p(a|b)$ условной вероятности (conditional probability) получения a , если будет получено b , мы предположили, что

⁹ Вера в колдовство может быть самосбывающейся, если проклятый человек верит в эффективность проклятия и попросту теряет волю к жизни. В данном случае наблюдаемая эффективность проклятия может заставить поверить в рациональность колдовства, даже если (как с теорией действия на расстоянии) агент не может определить механизм, который приводит его в действие. Этот факт может делать колдовство наказуемым на основании его фактических последствий, а не только как *mens rea*, как предлагали Донн и Гоббс (см. введение ко второй части).

p (премьер-министр компетентен) = 60%, p (премьер-министр некомпетентен) = 40%, p (хороший исход | премьер-министр компетентен) = 80% и p (плохой исход | премьер-министр некомпетентен) = 20%. Мы пытаемся вычислить p (премьер-министр компетентен | хороший результат). Для обозначения мы используем буквы a и b , соответственно для компетентности и хорошего исхода. Тогда в начале мы можем отметить

$$p(a|b) = p(a \& b) / p(b) \quad (*)$$

Условная вероятность $p(a|b)$ равна вероятности получения a и b , деленной на вероятность b . Это вытекает из интуитивного предположения, что $p(a \& b)$ равна $p(b)$, умноженной на $p(a|b)$. Разделив обе части этого уравнения на $p(b)$, мы получаем уравнение (*).

Снова используя уравнение (*), поменяв местами a и b , мы имеем

$$p(b|a) = p(a \& b) / p(a),$$

что равнозначно

$$p(a \& b) = p(b|a) p(a).$$

Подставляя последнее выражение в (*), мы получаем

$$p(a|b) = p(b|a) p(a) / p(b). \quad (**)$$

Есть два способа, которым может произойти b (хороший исход) с компетентным или некомпетентным премьер-министром. Вероятность того, что случится одно из двух взаимоисключающих событий, является суммой вероятностей каждого события. Мы, таким образом, можем записать

$$p(b) = p(b \& a) + p(b \& \text{не-}a),$$

которое, согласно рассуждению в абзаце, следующим за (*), равнозначно

$$p(b|a) p(a) + p(b|\text{не-}a) p(\text{не-}a).$$

Если мы подставим это выражение для $p(b)$ в (**), получим *теорему Байеса*

$$p(a|b) = p(b|a) p(a) / [p(b|a) p(a) + p(b|\text{не-}a) p(\text{не-}a)].$$

Замещение числовых вероятностей в правой части этого уравнения показывает, что $p(a|b) = 80\%$, то есть наблюдение успешного исхода повышает вероятность того, что премьер-министр компетентен, с 60 до 80%. Наблюдение второго и третьего положительного исхода поднимут его до 91 и 97%¹⁰. Если, по первоначальным оценкам другого человека, $p(a) = 0,3$,

¹⁰ В этом рассуждении мы полагаемся на две предварительные установки: основное убеждение, касающееся типа премьер-министра, и фоновое — о вероятности того, что премьер-министр данного типа обеспечит хороший исход.

а не 0,6, три последующих позитивных наблюдения поднимут его оценку сначала до 0,53, потом до 0,75 и, наконец, до 0,89. Таким образом, возможно, что не имеет значения, надежными или ненадежными были первоначальные предположения, поскольку по мере поступления новой информации откорректированные убеждения становятся все более достоверными. Со временем первоначальная разница во мнениях может быть нивелирована новыми данными¹¹. Заметим (глава XXIII), что каждый новый фрагмент информации оказывает меньшее воздействие, чем предыдущий.

ОПТИМАЛЬНОЕ ИНВЕСТИРОВАНИЕ В СБОР ИНФОРМАЦИИ

Третий компонент рационального действия — оптимальное инвестирование ресурсов (таких как время или деньги) в добывание новой информации. Как показано на рис. XI.1, есть несколько детерминант этого оптимума. Во-первых, какое количество информации рационально получить, зависит от желаний агента¹². Например, агент, который не слиш-

Окончание сн. 10

Если фоновое убеждение не меняется на протяжении всего времени, то основное пересматривается по мере поступления новой информации. Мы также можем поменять передний план и фон местами, если заинтересованы в установлении, что действительно компетентные политики могут воздействовать на экономические результаты или же на них действуют в основном случайные или внешние факторы. Тогда мы могли бы начать с основного убеждения о вероятности того, что хороший политик принесет хороший исход, и фонового убеждения о его компетентности. Если наблюдение показывает, что под руководством политика, который с большей вероятностью может считаться компетентным, дела в экономике идут плохо, мы можем присвоить большую вероятность гипотезе о случайных и внешних факторах.

¹¹ Чтобы произошла конвергенция, последовательно получаемые фрагменты информации должны быть независимы друг от друга. В хрестоматийном примере Байесова формирования убеждений человек вынимает шары из урны, о которой известно, что она с одинаковой вероятностью может содержать 80% белых шаров и 20% черных или 80% черных и 20% белых, для того чтобы определить вероятность того, какая именно перед ним урна. Поскольку шары вынимаются наугад, а после возвращаются в урну, исход каждой попытки не зависит от предыдущей. В политических ситуациях, вроде той, что описана в тексте, может оказаться гораздо труднее проверить независимость. Конвергенция также предполагает, что лежащая в ее основе ситуация не меняется или, по крайней мере, не меняется слишком быстро. Если обстановка быстро меняется, процесс корректировки убеждений напоминает стрельбу по движущейся мишени (глава VI).

¹² Принятие желаемого за действительное, при котором «желание — мать мысли», естественно, иррационально. Но нет ничего иррационального в процессе, показанном на рис. XI.1, где желания фактически — дед и бабка убеждений.

ком заботится о наградах в отдаленном будущем, не будет много вкладывать в определение срока службы товара длительного пользования. Очевидно, что имеет смысл собирать больше информации при принятии важного решения, например при покупке дома, чем при выборе между двумя бутылками одинаково дорогого вина. В последнем случае можно просто подбросить монету, если предполагаемые издержки на определение того, что лучше, превосходят ожидаемую выгоду (основанную на предварительном знании качественного диапазона для вина в данной ценовой категории) от употребления вина более высокого качества.

Желания и предварительные представления совместно определяют выгоды, ожидаемые от новой информации. Иногда можно с высокой точностью сказать, сколько дополнительных человеческих жизней будет спасено благодаря определенному анализу на рак или, переводя все на уровень агента, насколько вероятно, что будет спасена его жизнь. Ценность жизни зависит от того, как ее обменивают на другие желаемые цели. Согласно одному расчету, для того чтобы подвинуть человека, занятого на опасной работе (например, при добыче угля), принять один шанс гибели от несчастного случая из тысячи на протяжении года, потребуется премия около 200 долларов в год. Таким образом, в тот момент, когда был проведен этот расчет, ценность жизни составляла около 200 тысяч долларов¹³. Предполагаемые издержки на получение новой информации, определяемые предшествующими убеждениями, также иногда могут быть определены. Для того чтобы диагностировать рак прямой кишки, принято проводить шесть недорогих анализов кала. Выгоды от первых двух анализов велики. Однако для каждого из четырех последних анализов издержки на определение (даже не на лечение) дополнительных случаев рака, как выяснилось, составляют соответственно 49 150, 469 534, 4 724 695 и 47 107 214 долларов.

Оптимальный поиск информации может зависеть от результатов самого поиска (на рис. XI.1 на это указывает петля). Когда тестируется новый медицинский препарат, принимается предварительное решение давать его лишь одной из двух групп пациентов. Однако когда препарат за короткое время показывает хорошие результаты, становится неэтичным скрывать его от контрольной группы. Этот аргумент применим и к отдельному рациональному агенту. Представим: я в лесу собираю ягоды. Ягоды растут группами, поэтому я готов потратить некоторое время на поиски прежде, чем начать сбор. Если мне повезет

¹³ При выполнении таких расчетов есть много подводных камней, но общую идею отрицать невозможно: все мы придаем жизни конечную ценность. В противном случае мы не предавались бы приятным или выгодным, но рискованным занятиям.

и я сразу найду ягодное место, то с моей стороны будет глупо продолжать поиск.

Мы можем рассматривать сбор информации как *теневое действие*⁴ (shadow action), сопровождающее основное. Прежде чем что-то делать, мы должны решить, сколько информации нужно собрать. Иногда *теневое и основное действия могут совпадать*, по крайней мере частично. Предположим, лидеры государства оценивают целесообразность начала войны против другого государства. В качестве примера можно взять вторжение Германии во Францию в 1940 году. Информация имела ключевое значение при принятии окончательного решения. Руководители страны должны были знать объективные возможности предполагаемого противника, а также «организацию, обычаи и привычки вражеской армии» (из немецкой инструкции «Обязанности сотрудников Генерального штаба»). Большая часть такой информации может быть собрана традиционными методами, включая шпионаж. Однако чтобы определить *моральное состояние* врага, его боевой дух, единственное средство — сразиться с ним.

НЕОПРЕДЕЛЕННОСТЬ

Два последних примера — сбор ягод и подготовка к войне — также помогают нам обнаружить *ограничения* теории рационального выбора, а точнее, одно из двух ее ограничений. В качестве инструмента объяснения эта теория может потерпеть неудачу одним из двух способов. С одной стороны, она может не суметь предсказать, что будут делать люди в конкретной ситуации. С другой стороны, люди могут не оправдать эти прогнозы, вне зависимости от того, насколько четко они были сформулированы. Второй тип неудачи, *иррациональность*, станет темой следующей главы. Первый — *неопределенность* — является предметом следующих замечаний.

Агент может не суметь определить наилучший элемент в наборе осуществимых возможностей по одной из двух причин. Потребителю может быть безразличен выбор между двумя вариантами, которые в равной и наибольшей степени хороши. В обыденных ситуациях это происходит, когда потребитель сталкивается с выбором между двумя идентичными банками супа в супермаркете. В нетривиальных случаях варианты могут различаться по нескольким параметрам, так что эти различия в точности взаимно уравнивают друг друга. Нетривиальные случаи редки, но все же существуют. Если мне предложено выбрать между двумя автомобилями, различающимися ценой, уровнем комфорта, внешним видом, скоростью и так далее, я, возможно, предпочту один, а не другой, будучи в действительности к ним безразличен.

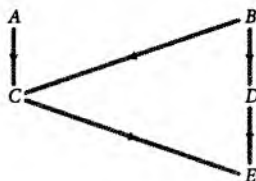


РИС. XI.4

Если бы я не был безразличен, то 5%-я скидка на одну машину должна была бы склонить в пользу этого варианта. Интуиция подсказывает, что такое вряд ли случится.

Предпочтения потребителя могут быть *неполными*. Предположим, я осмотрел пять моделей автомобилей — A, B, C, D, E, расставив их, как показано на рис. XI.4 (стрелки обозначают отношения предпочтений). Моя неспособность сравнить C и D не имеет значения, потому что я ни один из них не выберу.

Невозможность сравнить A и B, напротив, ставит меня в затруднительное положение. Да, я могу попытаться собрать дополнительную информацию, но как я определяю, что игра стоит свеч? Вскоре я вернусь к этой проблеме, однако прежде позвольте указать на другой и, может быть, более важный источник неполных предпочтений. Как правило, выбор того или иного варианта задан предпочтением определенного результата. Я склоняюсь к некоему варианту, потому что его итог, то есть его ожидаемая полезность в сравнении с другими вариантами, для меня предпочтителен. Если же это скорее неопределенная или неизвестная, а не рискованная ситуация (глава VII), я не имею возможности сравнить результаты¹⁴. Говоря бессмертными словами доктора Джонсона, «жизнь коротка, и не следует тратить слишком большую ее часть на размышления о том, как ее следует потратить: размышления, которые те, кто начинает их из благоразумия и продолжает с искусностью, должны после долгих затрат мысли закончить наугад. Чтобы предпочесть один будущий образ жизни другому на справедливых ос-

¹⁴ В ситуации принятия решения при неопределенности я могу иметь возможность сравнить варианты, если худший исход одного варианта лучше, чем лучший исход другого. Принимая решение в ситуации неизвестности, я лишен даже этой скромной возможности.

нованиях, требуются способности, которым Господь не благоволил нас наградить»¹⁵.

Дальнейшая неопределенность возникает из-за сложности определения оптимальных затрат на сбор информации. Когда я собираю ягоды на незнакомой территории, сколько времени я должен искать самое ягодное место и когда именно должен начать их собирать? Если только я сразу не напал на ягодное место, имеет смысл какое-то время продолжать поиски. При этом я не хочу заниматься поисками до ночи, потому что тогда наверняка уйду домой с пустой корзиной. Между нижним и верхним пределом во времени, которое нужно потратить на поиски, лежит обширный интервал неопределенности.

Иная проблема возникает в случае подготовки к войне. Если основное и теневое решения совпадают, то сторона, занятая подготовкой, все равно обречена (по крайней мере частично) оставаться в состоянии неопределенности. Теория рационального выбора не может служить верным ориентиром в таких случаях, будучи эффективной в четко очерченных ситуациях, в которых многое известно (таких как анализы на рак, например), но гораздо менее полезна в незнакомой обстановке.

Тратить меньше времени, определенного нижней границей, и больше, чем допускает верхняя, было бы иррационально. Ни один акт выбора, который агент совершает *внутри* этого интервала, не может быть охарактеризован как иррациональный. Можно подумать о том, чтобы отказаться от идеи рациональности в пользу идеи не-иррациональности. Эта откорректированная версия данной теории позволила бы нам объяснять больший диапазон вариантов поведения при более скромных возможностях прогнозирования. Большинство сторонников этой теории, думаю, не захотели бы пересматривать ее в таком ключе. То, что их прежде всего привлекает к этой теории, — это именно открываемая ею перспектива получения *однозначных* предсказаний. Ей это удастся благодаря элементарному математическому факту: любая удобная для анализа (*well-behaved*) функция полезности, определенная на удобном для анализа достижимом множестве, дает максимальное значение для един-

¹⁵ Я могу, однако, «закончить наугад», а потом придумать справедливые основания, например, придав большее значение тем атрибутам данного варианта, чье превосходство очевидно. Это может иметь нежелательные последствия. Предположим, у меня есть выбор пойти учиться юриспруденции или лесоводству. Будучи не в состоянии сделать обоснованный выбор, я более или менее наугад пойду учиться в юридическую школу и задним числом буду оправдывать мое решение, придавая больше значения доходу, который приносят эти профессии. С вновь возникшими предпочтениями я могу продолжать принимать решения, отличные от тех, которые я принял бы на основании предшествовавших выбору предпочтений.

ственного члена этого множества. Взаимодействие между достижимым множеством и кривыми безразличия на рис. IX.1 дает хорошее представление о простоте этой идеи, демонстрирующей все, на что она способна.⁴

Самый важный источник неопределенности убеждений возникает при стратегическом взаимодействии, когда каждый агент формирует свои убеждения о том, что с наибольшей вероятностью будут делать другие на основании *их* убеждений, зная при этом, что аналогичным образом они рассуждают в отношении него самого. В некоторых случаях, рассматриваемых в главе XIX, структура вознаграждения не позволяет агентам сойтись на общем наборе убеждений.

РАЦИОНАЛЬНОСТЬ НАСКВОЗЬ СУБЪЕКТИВНА

Позвольте завершить обсуждение теории рационального выбора, снова подчеркнув ее *радикально субъективную природу*. Наверняка кто-то понимает слово *рациональный* в объективном смысле, подразумевая, что рациональный агент тот, кто принимает решения, которые *делают его жизнь лучше* в соответствии с объективными критериями, такими как долгая продолжительность жизни, здоровье и доход. Однако при таком понимании эта идея не будет иметь никакой объясняющей способности. Как я подчеркивал, *последствия* решения не могут его объяснить. Только ментальные состояния, которые предшествуют решению, позволяют нам *объяснить* действия как оптимальные с точки зрения агента, а не характеризовать их как полезные или благотворные с точки зрения постороннего наблюдателя (или того же агента в следующий момент времени).

Предположим, я страдаю от глубокой неспособности откладывать удовлетворение, то есть не в состоянии принимать в расчет будущие последствия настоящего поведения. Допустим также, что ученые изобрели таблетку дисконтирования, которая усилит вес будущих вознаграждений в настоящих решениях. Если я приму таблетку, моя жизнь улучшится. Мои родители будут рады, если я ее приму. Оглядываясь назад, я сам буду доволен тем, что принял ее. Но если у меня есть выбор — принимать таблетку или не принимать, я откажусь от нее, если я рационален. Любое поведение, которое может вызвать таблетка, уже находится для меня в зоне достижимости. Я мог бы бросить курить, начать заниматься зарядкой или откладывать деньги прямо сейчас, но я этого не делаю. Поскольку я не хочу это делать, я не захочу принять таблетку, которая заставит меня это делать. Точно так же эгоистичный человек откажется от «таблетки альтруизма», а альтруист откажется от «таблетки эгоизма». Если я люблю свою семью и отчасти готов пожертвовать ради нее своим гедонистическим образом жизни, я отказался бы от таблетки с двухступенчатым эффектом понижения их благополучия, как отказался бы

от любого варианта (например, дорогого обеда для себя), который произвел бы тот же эффект в один этап.

Чтобы усилить этот аргумент, предположим, что человек потребляет x сегодня и y завтра и что его коэффициент дисконтирования за один период составляет 0,5 (ему безразлично, получить одну единицу полезности завтра или пол-единицы сегодня). Примем для простоты, что $u(x) = x$ и $u(y) = y$. Дисконтированная настоящая ценность его потребительского потока составляет $x + 0,5y$. Предположим, человек узнает, что завтра он будет испытывать боль, которая снизит полезность его потребления на 0,5. Теперь дисконтированная настоящая ценность его потребления $x + 0,25y$. Если рациональному агенту предложат бесплатный аспирин, который снимет боль, он его, естественно, примет, тем самым восстановив первоначальную настоящую ценность. Если бы он принял таблетку, которая индуцирует коэффициент дисконтирования, равный 1 (но не принял бы аспирин), исход был бы тем же в том смысле, что агенту был бы безразличен выбор между двухвременным потоком и одновременной полезностью $x + 0,5y$. Если бы агент принял аспирин и если эффект был бы таким же, как от таблетки дисконтирования, то почему бы ему не принять эту таблетку? Причина в том, что выбор такого варианта сдерживается потребностью в том, чтобы потребление, обусловленное приемом таблетки, превосходило таковое без нее, *оцененное с точки зрения первоначальных предпочтений*. Для выбора аспирина такого ограничения не существует, потому что нет разницы между предпочтениями до его приема и после. Даже без аспирина я предпочитаю иметь завтра возможность избавиться от боли. Когда это положение войдет в мой поведенческий репертуар, я решу его реализовать. Между тем поток полезности, вызванный таблеткой дисконтирования, уже присутствует в репертуаре, но я решаю им не пользоваться¹⁶.

¹⁶ При гиперболическом дисконтировании агент может принять названную таблетку. Используя числовой пример из главы 6, предположим, что эффект таблетки ниже значения k от 1 до 0,3. В тот момент, когда появляется меньшее вознаграждение, его настоящая ценность просто 10 (дисконтирования нет). Настоящая ценность большего вознаграждения в это время составляет $30 / (1 + 0,3 \cdot 5) = 12$. Отсюда следует, что предварительное связывание себя обязательством (precommitment) через прием таблетки даст агенту возможность действовать по спокойному и здравому рассуждению, тем самым предотвращая проявление слабости (в широком смысле). Это утверждение остается верным, даже если агент должен купить таблетку, если только ее стоимость (в единицах полезности) не превышает 2. Оно также остается верным, если предварительное связывание себя обязательствами будет иметь эффект сокращения ценности отложенной награды (возможно, таблетка дисконтирования имеет побочный эффект ослабления способности к наслаждению) до тех пор, пока потеря не превышает 5. Эти факты могут быть верными, если вместо таблетки дисконтирования мы возьмем психотерапию.

Другими словами, на выбор нужно смотреть глазами агента. Близорукому человеку, потерявшему свои очки, близорукость может мешать их найти. Он попал в ловушку¹⁷. Похожим образом рациональный агент может оказаться в ловушке убеждений, в которой он застрянет с ложной верой в то, что предполагаемые издержки на проверку убеждения слишком высоки. Таким образом, женщины, практикующие обрезание, могут находиться в ловушке убеждений. Бамбара из Мали верят, что клитор убьет мужчину, если вступит в контакт с пенисом во время сокоупления. В Нигерии некоторые племена считают, что если головка ребенка во время родов коснется клитора, ребенок умрет. В Польше существовало распространенное убеждение, что если кто-то, кому под кожу вшит дисульфид, выпьет, он умрет. На самом деле дисульфид фармакологически инертен. Ложное убеждение, тем не менее, удерживает людей от его проверки.

Рациональность верований и их истинность — это совершенно разные вещи. Если истина — это отношение между убеждением и миром, рациональность — это особенность отношения между убеждением и фактами, которыми обладает агент. Хотя рациональность требует от агента вложений в информацию, этот вклад всегда ограничен ожидаемыми (и *предполагаемыми*) затратами и прибылями. Если считается, что сбор дополнительной информации имеет слишком высокие *альтернативные издержки* (opportunity costs), как это бывает перед лицом неминуемой опасности, возможно, рациональнее воздержаться от такого инвестирования. Если считается, что это влечет высокие *прямые издержки* (direct costs), как в случае с проверкой веры в то, что с имплантированным дисульфидом пить смертельно опасно, только иррациональный человек будет делать подобные вложения. Обобщенно говоря, многие убеждения следует принимать за чистую монету из вторых рук, поскольку если мы возьмемся их проверять, то нам будет некогда заниматься собственной жизнью.

Любое объяснение поведения, основанное на выборе, субъективно. Хотя не все субъективные объяснения предполагают прозрачность агентов для самих себя и неустанные поиски оптимальности, составляющие отличительную особенность объяснений посредством рационального выбора. В следующей главе я буду рассматривать некоторые объяснения, которые отступают от теории рационального выбора в том или ином аспекте.

¹⁷ Если ему дадут его очки, он их наденет. Я утверждал, что если ему предложат таблетку дисконтирования, он ее не примет. Разница в том, что без таблетки дисконтирования он уже делает все, что мог бы делать после приема этой таблетки, тогда как существует множество вещей, которые он не может делать без очков, но делал бы, будь они на нем.

Библиографические примечания

Отношения между разумом (в том смысле, в каком это слово употребляется в главе IV) и рациональностью я развиваю в моей инаугурационной лекции в «Коллеж де Франс» *«Разум и доводы»* (Raison et raisons. Paris: Fayard, 2006). Подробнее о Вебере и рациональности можно прочесть в моей статье *«Рациональность, экономика и общество»* (Rationality, economy, and society // Turner S. (ed.). The Cambridge Companion to Weber. Cambridge University Press, 2000). Классическое изложение теории полезности можно найти в *«Играх и решениях»* Р.Д. Льюса и Х. Райфффа (Льюс Р., Райфффа Х. Игры и решения. М.: Иностранная литература, 1962). Стоит посмотреть и оригинальную работу Дж. фон Неймана и О. Моргенштерна *«Теория игр и экономическое поведение»* (Нейман Дж. фон, Моргенштерн О. Теория игр и экономическое поведение. М.: Наука, 1970). Выдающееся изложение теории рационального выбора (и ее проблем) есть в книге *«Рациональный выбор в неопределенном мире»* Р. Хести и Р. Доуза (Hastie R., Dawes R. Rational Choice in an Uncertain World. Thousand Oaks, CA: Sage, 2001). Я подробно обсуждаю опеку над детьми в главе 3 *«Соломоновых решений»* (Solomonic Judgments. Cambridge University Press, 1989). Отличное введение в теорию Байеса можно найти во *«Введении в байесовский вывод и Байесово решение»* Р. Винклера (Winkler R. An Introduction to Bayesian Inference and Decision. Gainesville, FL: Probabilistic Publishing, 2003). На мой аргумент о том, что рациональный человек не станет принимать таблетку дисконтирования, повлиял обмен мнениями с Гэри Бекером (Gary Becker) и Питером Даймондом (Peter Diamond); см. также *«Рассуждения о формировании терпения. Необходимость концептуальных различий»* О. Дж. Сгора (Skog O. J. Theorizing about patience formation: The necessity of conceptual distinctions // Economics and Philosophy. 2001. No. 17. P. 207–219). Идею убеждений как ловушки я взял из Дж. Макки *«Пеленание ног и инфибуляция»* (Mackie G. Ending footbinding and infibulation: A convention account // American Sociological Review. 1996. No. 61. P. 999–1017). Полезное исследование о важности шпионажа в подготовке к войне — книга *«Странная победа. Завоевание Гитлером Франции»* Э. Мей (Мэй Э. Странная победа. М.: АСТ, 2009). Информацию об имплантации дисульфирама в Польше я почерпнул из книги В. Осетински *«Алкоголизм: грех или болезнь?»* (Osiatynski W. Alcoholism: Sin or Disease? Warsaw: Stefan Batory Foundation, 1997), а данные о ее неэффективности из Дж. Йонсен и Дж. Мерленд *«Препарат дисульфирама замедленного действия: экспериментальные и клинические результаты»* (Johnsen J., Mørland J. Depot preparations of disulfiram: Experimental and clinical results // Acta Psychiatrica Scandinavica. 1992. No. 86. P. 27–30).

XII. Рациональность и поведение

ИГНОРИРОВАНИЕ ИЗДЕРЖЕК НА ПРИНЯТИЕ РЕШЕНИЙ

Идея рациональности обладает сильной нормативной привлекательностью. Для того, что мы делаем, мы *хотим* иметь основания — желания и убеждения, в свете которых наши действия выглядели бы рациональными. В действительности наше желание действовать с каким-то основанием, то есть наше преклонение перед рациональностью, может быть таким сильным, что приведет к иррациональному поведению¹. Мы можем определить *гиперрациональность* как склонность к поиску абстрактно оптимального решения, то есть решения, которое было бы оптимальным, если бы мы игнорировали издержки на сам процесс принятия решения. Это издержки троякого рода: (1) стоимость средств принятия решения, (2) стоимость побочных эффектов принятия решения, (3) альтернативные издержки принятия решения, то есть стоимость других вещей, которые могли бы быть сделаны, вместо того чтобы заниматься принятием решения. Позвольте кратко проиллюстрировать эти разновидности.

Гиперрациональность, возникающая из игнорирования издержек (1), может иметь место при *сравнительном шопинге*, когда (предполагаемая) экономия от того, что товар найден по самой низкой цене, меньше суммы, затраченной на поездку из одного магазина в другой. Туристы на юге Франции пересекают границу с Испанией, чтобы купить дешевые сигареты, как будто бензин ничего не стоит. Игнорирование издержек (2) может порождать гиперрациональность в случаях оспаривания опе-

¹ Действие с основанием не следует путать с действием в соответствии с разумным основанием, как это было определено в главе IV. Человек, который действует, руководствуясь своекорыстным интересом, действует с основанием, но не в соответствии с разумным основанием. И наоборот, человек, действующий в соответствии с категорическим императивом, действует на разумных основаниях, но не с основанием (см. примечание 15 к настоящей главе). Читатели, которые сочтут это утверждение контринтуитивным, возможно, имеют отличное от моего представление о том, что значить «действовать с основанием».

ки над ребенком. Суд может защищать интересы ребенка, определяя, кто из родителей больше подходит для опеки². Как только этот вопрос решен, у суда есть веские основания присудить опеку одному из родителей. Однако в юридически-психологическом процессе оценки относительной пригодности к опеке ребенку может быть нанесен непредвиденный ущерб. С учетом поставленной цели более рациональной процедурой могло бы быть подбрасывание монетки или сохранение традиционной презумпции материнской опеки.

Игнорирование альтернативных издержек можно проиллюстрировать замечанием доктора Джонсона в разговоре с Босуэллом о том, чему необходимо учить детей прежде всего. «Сэр, то, чему вы их прежде всего научите, имеет не больше значения, чем то, какую ногу вы первой вдеваете в брюки. Пока, сэр, вы стоите и обсуждаете, с какой ноги лучше начать, брюки так и остаются ненадетыми. Пока вы обдумываете, сэр, какой из двух вещей следует научить ребенка, другой мальчик уже выучит обе»³. Опять-таки подбросить монету, возможно, было бы рациональнее.

Или возьмем врача, который прибывает на место происшествия и решает, какие меры следует принять в первую очередь. Хотя сначала он должен осмотреть пациента, его поведение будет ошибочным, если он потратит на это столько времени, что пациент умрет у него на руках («Операция прошла успешно, но больной умер»). Некоторые сталкивались с таким опытом, собирая ягоды. Можно так долго искать ягодное место получше, что наступит ночь. Даже если экономия от сравнительного шопинга превышает расходы на транспорт, это поведение может оказаться иррациональным из-за стоимости потерянного времени.

НЕКОТОРЫЕ КАНОНИЧЕСКИЕ ПРИНЦИПЫ РАЦИОНАЛЬНОСТИ

Под головоломкой я буду понимать наблюдаемое поведение, которое, как кажется, не поддается объяснению с позиции рационального выбора. Хотя некоторые головоломки при ближайшем рассмотрении могут потерять свой головоломный характер, с другими этого, однако, не происходит. Эксперименты и поведение в реальной жизни демонстрируют многочисленные примеры поведенческих моделей, которые нарушают каноны рациональности. В приведенном ниже выборочном списке кано-

² «В наилучших интересах ребенка» — критерий, применяемый в большинстве законодательств об опеке над ребенком.

³ Джонсон, будучи исследователем Шекспира, мог иметь в виду замечание короля из «Гамлета»: «Как человек с колеблющейся целью, / Не знаю, что начать, и ничего / Не делаю» (пер. Б. Пастернака).

ческих принципов я начну с фундаментальных, а потом перейду к более специфичным. Я ограничусь индивидуальным выбором; отклонения при интерактивном выборе обсуждаются в пятой части. Этот список может быть дополнен когнитивными аномалиями, которых я касался в главе VII.

1. При выборе между действием и бездействием рациональный агент не будет действовать, если предполагаемые затраты на полезность действия превышают ожидаемые выгоды от этой полезности.
2. Выбирая из двух зол, рациональный агент выберет меньшее.
3. Рациональный агент придает одинаковый вес альтернативным (opportunity) и прямым (direct) издержкам.
4. Рациональный агент никогда не отдаст предпочтение подмножеству множества возможностей, если можно получить полное множество.
5. Если рациональный агент предпочитает X стакану, описываемому как наполовину полный, он также должен предпочитать X стакану, описываемому как наполовину пустой.
6. В чисто азартной игре рациональный игрок, делая ставки, не будет обращать внимания на исходы, выпавшие предыдущим игрокам.
7. Решая вопрос о том, продолжать проект или закрыть его, рациональный инвестор будет обращать внимание только на настоящее значение будущих потоков полезности от двух опций.
8. Если в момент 1 агент будет планировать совершить действие X в момент 2, он выполнит это действие в момент 2, если только к тому времени не изменятся его желания или убеждения.
9. При рискованном выборе рациональный агент будет выбирать средства в соответствии с ожидаемым исходом, а не с наихудшим (или наилучшим) сценарием.
10. На рынке, на котором играют рациональные агенты, окупаемость по всем активам должна быть (приблизительно) одинаковой⁴.
11. Если рациональный агент выбирает A из множества (A, B, C) , он также выберет A из множества (A, B) .

⁴ Есть две причины того, почему ожидаемое равновесие может быть только приближенным. Во-первых, внешние потрясения всегда вызывают его отклонения. Во-вторых, неприятие риска может вызывать более низкую стоимость (и соответственно более высокую доходность) высоковолатильных активов.

12. Рациональный агент не будет воздействовать на следствие, чтобы искоренить причину (при лечении пневмонии он будет принимать антибиотики, а не аспирин).
13. Если рациональный агент предпочитает верное получение вознаграждения X получению вознаграждения Y с вероятностью q , он также предпочтет получение X с вероятностью p получению Y с вероятностью pq (аксиома независимости теории количественной полезности).
14. Если рациональный агент делает X , когда знает, что будет иметь место обстоятельство C (или намеривается делать X , когда предполагается, что будет иметь место обстоятельство C), и делает X , когда обстоятельство C не будет иметь места (или намеривается делать X , когда не предполагается, что обстоятельство X будет иметь место), он будет делать или намериваться сделать X в условиях отсутствия информации об обстоятельствах.
15. Рациональный агент никогда не сделает предложение о сделке, если ее принятие раскроет информацию, в результате которой ожидаемая стоимость сделки станет отрицательной.
16. Если проступок вызывает желание отомстить, пострадавший, если он рационален, выдерживает время и наносит ответный удар при максимальной вероятности успеха или при минимальном риске для себя⁵.
17. При вызове на дуэль на шпагах рациональный агент будет брать уроки фехтования, если собирается принять вызов.
18. Прежде чем просить у кого-то руки, рациональный агент соберет информацию о поведенческих и эмоциональных склонностях избранницы.

НАРУШЕНИЯ КАНОНА

Эти нормативно-принудительные принципы, как выясняется, повсеместно нарушаются. Ниже приведены примеры (с номерами, соответствующими номерам принципов, которые в них нарушены)⁶.

⁵ Некоторые захотят заменить в этом предложении «или» на «и». Максимизировать эти две цели одновременно возможно лишь при счастливом стечении обстоятельств. Точнее, агент будет стремиться к оптимальному достижимому соединению этих двух целей, как это представлено множеством допустимых решений и кривых безразличия (глава IX).

⁶ Многие из примеров упоминались в предшествующих главах и приводятся здесь все вместе для удобства. Они взяты из разных источников — пословиц, классических произведений, мыслительных или лабораторных эксперимен-

1. *Парадокс голосования*. Поскольку итоги ни одного общенационального голосования не зависели от одного поданного голоса, индивидуальный выбор никак не влияет на результаты, тогда как поход на избирательный участок может доставить значительные неудобства. И все же большое число людей голосуют⁷.
2. *«Большую боль предпочитают меньшей»*. Участников эксперимента заставили прослушать две последовательности крайне неприятного шума. Обе включали 35-секундный очень громкий шум. В первом случае за этим следовало 15 секунд постепенно уменьшающегося (но по-прежнему крайне неприятного) шума. Когда их спросили, какую последовательность они готовы прослушать снова, участники выбрали последовательность, которая при всей очевидности была более неприятной.
3. *«Парадокс выкошенной лужайки»*. В небольшом пригородном сообществе мистер Х косит траву на своей лужайке. Сын его соседа мог бы скосить ее за 12 долларов. Мистер Х не стал бы косить траву на лужайке своего соседа за 20 долларов.
4. *«Головоломка рождественского клуба»*. Клиенты ежемесячно вносят сумму с низким процентом годовых или вообще без процента, которую они могут снять только в Рождество. В то же время у них имеются варианты с получением нормальных процентов или беспрепятственного снятия денег со счета.
5. *«Парадокс кредитной карточки»*. Когда появились кредитные карточки, поддерживающее их лобби предпочитало, чтобы разница между наличными и карточкой была представлена как скидка за наличный расчет, а не как доплата за использование кредитной карты. Хотя эти описания логически эквивалентны, потребители с большей охотой пользовались карточками, если разница описывалась как скидка за оплату наличными.

тов и жизненных наблюдений. Примеры в первых трех категориях основаны на теориях, которые разбираются в последующей части главы.

⁷ Этот парадокс возникает, когда единственная цель избирателей — выбрать кандидата на должность или ввести в действие законодательный акт. В его возникновении нет необходимости, когда требуется вдохнуть жизнь в демократическую систему или дать мандат кандидату, поскольку в этих случаях избиратели важны, даже если их число не имеет решающего влияния на результат. Но даже когда целью является поддержание демократии, объяснение может дать действие одного из нерациональных механизмов, обсуждаемых далее.

6. *Два заблуждения игроков.* Если пять раз подряд выпадало красное, примерно половина игроков полагают, что есть вероятность более 50%, что оно выпадет снова. Другая половина полагают, что есть вероятность менее 50%, что выпадет черное.

7. *Ошибка необратимых затрат.* «Завершение проекта, в который было вложено 1,1 миллиарда долларов, представляет собой бессовестное разбазаривание денег налогоплательщиков» (сенатор Дентон, 4 ноября 1981 года). Эту ситуацию иногда называют ошибкой Конкорда — по названию дорогостоящего англо-французского авиалайнера, или ошибкой Вьетнама — по ассоциации с нежеланием США выводить свои войска из этой страны. Если вы купили билеты на какое-то мероприятие и сильный снегопад затрудняет вам дорогу, вы все равно можете на него поехать, хотя и отказались бы, если бы вам предложили билеты бесплатно.

8. *«Головоломка дантиста».* 1 марта я записываюсь к зубному врачу на 1 апреля. 30 марта я звоню ему, чтобы сказать, что не смогу прийти из-за (выдуманных) похорон кого-то из родственников. За этот период не произошло никаких изменений, просто прошло время. В частности, зуб по-прежнему болит.

9. *Наилучший и наихудший сценарии.* Пациенты, страдающие от рака, на поздних стадиях заболевания часто переоценивают шанс на выживание. Вместо паллиативной терапии, облегчающей боль, они выбирают агрессивную и болезненную химиотерапию, не приносящую особой пользы. Когда людей спрашивают, сколько они готовы заплатить, чтобы уменьшить вероятность маловероятного несчастья, они готовы заплатить одинаковую сумму за снижение вероятности как до одного шанса на миллион, так и до одного шанса на 10 миллионов.

10. *Головоломка премии по акциям* (the equity premium puzzle). Исторически прибыль по акциям гораздо выше, чем прибыль по облигациям. Человек, инвестировавший 1 доллар в акции 1 января 1928 года, получил бы 1 января 1998 года инвестиционный портфель стоимостью 1800 долларов. Тот, кто вложил доллар в облигации, получил бы портфель стоимостью 15 долларов. Загадка, почему такое расхождение не привело к росту стоимости акций, который сблизил бы доходность по акциям по отношению к доходности по облигациям.

11. *Эффект нерелевантных альтернатив.* Если каждая из опций А и В превосходит друг друга по одному или двум релевантным параметрам, у людей может возникнуть затруднение

с выбором; они могут принять решение собрать дополнительную информацию о вариантах. Если в этом случае вводится третья опция С, которая (1) хуже А по обоим параметрам и (2) хуже В по одному параметру и лучше по другому, наблюдается тенденция выбирать А без дальнейших поисков.

12. *«Головоломка холодной воды»*. В эксперименте участники, которых заставили поверить, что время, которое они могут продержать руки в очень холодной воде, является наилучшим показателем их продолжительности жизни, держали руки в воде дольше, чем те, кто не получил эту (ложную) информацию.

13. *Эффект уверенности* (глава VII). В проведенных экспериментах большинство участников предпочитали получить шестидневную поездку по Англии наверняка, нежели иметь 50%-ю вероятность выиграть трехнедельный тур по Англии, Франции и Италии, но большинство также предпочитает 5%-й шанс второго варианта 10%-му первого.

14. *Эффект дизъюнкции*. Участники эксперимента, предполагающие, что выиграют в следующем раунде, на вопрос о том, будут ли они принимать в нем участие, склонны отвечать «да». Предполагая, что их ждет проигрыш, они склонны выражать такое же намерение. Если же они не знают, проиграют или выиграют, их намерения теряют определенность. Такой же эффект наблюдается в одноразовых «дилеммах заключенного»: человек в большей степени готов к сотрудничеству, если знает, что другой человек сотрудничал, чем когда ему известно, что тот отказался от сотрудничества, и (в этом и заключается эффект дизъюнкции) его готовность к сотрудничеству будет еще больше, если он не знает о выборе другого человека.

15. *«Проклятие победителя»*. В этом эксперименте участникам предлагается выставить на торгах свою цену за участок земли, точную стоимость которого знает продавец, тогда как участникам торгов известно только то, что эта стоимость лежит в рамках определенного диапазона, в котором все значения одинаково вероятны. Кроме того, покупателям сказано, что если они приобретут участок земли, для них он будет на 50% ценнее, чем для продавца, потому что они смогут более эффективно его эксплуатировать. Если предложение принимается, рациональные покупатели должны суметь сделать из этого вывод, что предполагаемая стоимость земли меньше той, которую они предложили. Если значения варьируют от 0 до 1000 и принимается предложение (например) 600, покупатель может из этого заключить, что реальная стоимость для продавца

лежит где-то между 0 и 600 с предполагаемой стоимостью 300. Таким образом, для покупателя эта предполагаемая стоимость была бы 450, что меньше предложенной им цены. Поскольку так можно рассуждать о любом принятом предложении на торгах, рациональные покупатели никогда не должны в них участвовать. Однако в экспериментах (основанных на реальных случаях) все неизменно делают свои ставки.

16. «Поспешная месть». Одна пословица гласит: «Мечь — это блюдо, которое нужно подавать холодным». Еще одна утверждает: «Чем дольше месть откладывают, тем сильнее удар». Предположительно обе пословицы возникли как реакция на скоропалительную месть, тем самым подтверждая существование подобного феномена.

17. *Пренебрежение эффе^ктивностью*. Монтень писал: «Честь сражения заключается в соперничестве сердца, а не умения; вот почему я видел, как некоторые мои друзья, мастера в данной дисциплине, выбирали для дуэлей оружие, которое лишало их возможности пользоваться своими преимуществами и которое полностью зависело от удачи и прямоты, чтобы никто не смог приписать их победу фехтовальному мастерству, а не их достоинству».

18. «Жениться впопыхах, каяться не спеша». Эта пословица применима не только к браку в прямом смысле слова. Когда люди влюбляются в дом, они иногда так торопятся подписать контракт, что не замечают скрытые недостатки.

АЛЬТЕРНАТИВЫ ТЕОРИИ РАЦИОНАЛЬНОГО ВЫБОРА

Для объяснения этих головоломок существует широкий диапазон альтернатив объяснению с позиций рационального выбора. Прежде чем обсуждать каждую из них, позвольте мне привести список ключевых механизмов для альтернативных объяснений, которые я буду рассматривать (в скобках указывается номер головоломки). Некоторые головоломки фигурируют более одного раза, потому что могут быть объяснены несколькими способами.

- Неприятие потерь (3, 5, 7, 10)
- Не согласующееся с теорией вероятности взвешивание исходов (13)
- Гиперболическое дисконтирование (4, 8)
- Эвристика (2, 6)

- Принятие желаемого за действительное (9, 12)
- Неспособность к проецированию (15)
- Желание действовать с основанием (11, 14)
- Магическое мышление (1, 12, 14)
- Категорический императив (1)
- Эмоции (3, 7, 14, 18)
- Социальные нормы (1, 3, 16, 17)

В обыденном мышлении, возможно, в наибольшей степени выделяют-ся механизмы неприятия потерь и гиперболического дисконтирования. По моему мнению, эмоции являются еще более важным источником ир-рационального поведения, независимо от того, действуют они напрямую или при посредничестве социальных норм. Хотя эмоции могут нарушать рациональность многими способами, вероятно, самый важный из них — порождение безотлагательности (*urgency*):

НЕПРИЯТИЕ ПОТЕРЬ определяется в соответствии с точкой отсчета с тем допущением, что люди оценивают изменения относительно этой базовой линии, а не с точки зрения конечных состояний, полученных после изменения. Как правило, за точку отсчета берется статус-кво, хотя субъектов могут заставить использовать и другие ориентиры. Неприятие потерь — это склонность придавать большее значение (в абсолютных категориях) потерям по отношению к базовому уровню, чем приобретениям того же размера⁸. Эмпирически считается, что коэффициент составляет 2,5 к 1. Этим коэффициентом я буду пользоваться в дальнейшем. Еще одно важное свойство функции ценности в том, что она вогнутая для выигрышей и выпуклая для потерь, что означает, что каждая дополнительная единица выигрыша ценится меньше, чем предыдущая, а каждая дополнительная единица потерь болезненнее предыдущей.

Две из головоломок могут быть объяснены тем, что потери кажутся более важными, чем выигрыши. Для решения «парадокса выкошенной лужайки» нам лишь следует отметить, что неприятие потерь предсказывает, что альтернативные издержки и расходы наличными оцениваются совершенно по-разному. Поскольку для домовладельца стоимость

⁸ Предполагая, что блага (включая деньги) имеют уменьшающуюся предельную полезность, стандартная теория рационального выбора предсказывает, что потери считаются более важными, чем приобретения такого же размера, рассматриваемые с той же самой базовой линии. Тем не менее величина эффекта, как правило, гораздо меньше. Кроме того, стандартная теория полезности предполагает, что выигрыш в полезности от перехода из *A* в *B* равен потере в полезности при переходе из *B* в *A*, поскольку эти различия возникают просто от сравнения уровней полезности двух состояний.

выигрыша 20 долларов равносильна стоимости потери 8 долларов, он полагает, что лучше отказаться от выигрыша, чем заплатить 12 долларов из своего кармана. Этим же рассуждением можно объяснить «парадокс кредитной карточки».

Головоломка премии по акциям требует дополнительного допущения, согласно которому люди выбирают свой набор акций и облигаций внутри краткосрочного горизонта. Поскольку доходность по акциям волатильна, тогда как облигации приносят устойчивый доход из года в год, мы можем рассматривать владение акциями в качестве принятия рискованной ставки. Допустим, мы предлагаем человеку ставку на акции, которая дает ему 50%-й шанс выиграть 200 долларов и 50%-й шанс потерять 100 долларов с фиксированной доходностью по облигациям в качестве точки отсчета. Если мы предполагаем неприятие потерь, то это отражается в высказывании, что стоимость денег равна x для $x > 0$ и равна $2,5x$ для $x < 0$. Поскольку стоимость потери 100 долларов равна (в абсолютном выражении) выигрышу 250 долларов, перспектива выиграть 200 долларов не может компенсировать вероятную перспективу потерять 100. Следовательно, человек отвергнет предложение. Допустим, что человеку предложили пакет из двух таких ставок, которые делаются последовательно. Эта смешанная игра предполагает 25%-й шанс выиграть 400 долларов, 50%-й шанс выиграть 100 долларов и 25%-й шанс проиграть 200 долларов. Если мы умножим проигрыш на коэффициент 2,5, чтобы сделать его сопоставимым с выигрышами и подсчитать ожидаемую прибыль⁹, легко показать, что он составит 25. По этой причине человек примет смешанную игру. Эмпирические исследования показывают, что у инвесторов имеется склонность слишком часто проводить переоценку инвестиционного портфеля. Это близорукая политика, заставляющая их слишком мало вкладывать в акции и слишком много в облигации¹⁰.

⁹ Поскольку теория перспективы предполагает, что вес решения отличается от вероятностей, такой расчет корректен лишь приблизительно. Ни это упрощение, ни допущение линейной функции ценности не имеют значения для завершения этого анализа.

¹⁰ Определение *близорукий* здесь имеет значение, отличное от того, которое в него вкладывалось в анализе дисконтирования во времени (глава VI). Оно указывает не на способ расчета агентом настоящей ценности будущих потоков дохода, но на склонность принимать последовательные решения по отдельности, а не связывая их вместе в одном общем выборе. Такая близорукость принятия решений может работать и в других контекстах. Таким образом, когда люди пытаются контролировать гиперболическое дисконтирование, группируя последовательные выборы (глава XIII), их успех может зависеть от количества выборов, которые они включают.

Ошибка необратимых затрат отсылает только к форме (вогнутой или выпуклой) функции ценности. Давайте рассмотрим следующий пример. Семья платит p долларов за билеты на матч, который будет проходить в 60 милях от их дома. В день матча начинается снежная буря. Они решают ехать, заметив между делом, что если бы билеты им достались даром, они остались бы дома. Записав v для функции ценности для выигрышей, получаем, что ценность похода на матч составляет $v(g)$. Определив v^* для функции ценности выгоды, ценность потери p долларов является отрицательным числом $v^*(-p)$. Стоимость поездки во время снежной бури составляет c . Допустим, что $v(g) = -v^*(-c)$, что предполагает, что если бы семья получила билеты бесплатно, ей было бы безразлично, оставаться дома или ехать на матч в снежную бурю. Но поскольку они уже заплатили p долларов, они предпочитают ехать. Чтобы это понять, отметим, что из-за выпуклости $v^* v^*(-(c+p)) > v^*(-c) + v^*(-p)$ ¹¹. Это может быть переписано как $v^*(-(c+p)) - v^*(-c) > v^* v^*(-p)$, что при только что указанном допущении равносильно $v^*(-(c+p)) + v(g) > v^*(-p)$. Поскольку левосторонний член в последнем неравенстве есть чистый выигрыш или потеря от поездки на игру, а правосторонний член — потеря от отказа от такой поездки, они предпочитают поехать.

НЕ СОГЛАСУЮЩЕЕСЯ С ТЕОРИЕЙ ВЕРОЯТНОСТИ ВЗВЕШИВАНИЕ ИСХОДОВ. Неприятие потери следует из влиятельной альтернативы теории рационального выбора, называемой *теорией перспективы*. Другое следствие этой теории заключается в том, что люди склонны взвешивать исходы иначе, чем активы теории ожидаемой полезности. Согласно этой теории полезность линейна по вероятности (глава XI). Теория перспективы, наоборот, утверждает, что люди наиболее чувствительны к изменениям вероятности вблизи естественных границ 0 (невозможный исход) и 1 (гарантированный исход). Эффект уверенности демонстрирует нелинейность вокруг 1. Создатели теории перспективы Даниэль Канеман и Амос Тверски приводят следующий пример нелинейности вокруг 0, (который они приписывают Ричарду Цекхаузеру (Richard Zeckhauser)).

Предположим, вас заставляют сыграть в русскую рулетку, но вам дана возможность купить право вынуть одну пулю из заряженного пистолета. Вы бы заплатили за то, чтобы сократить число пуль с четырех до трех, столько же, сколько за то, чтобы сократить это число с единицы до нуля? Большинство людей ощущают, что гораздо больше заплатили бы за то, чтобы сократить шанс гибели с $\frac{1}{4}$ до нуля, чем за его уменьшение с $\frac{1}{4}$ до $\frac{3}{4}$. Экономические соображения [то есть теория

¹¹ Поскольку это сравнение двух отрицательных чисел, неравенство устанавливает, что первое ближе к нулю, то есть меньше по абсолютному значению.

ожидаемой полезности] заставили бы заплатить больше в последнем случае, в котором ценность денег предположительно уменьшается значительной вероятностью того, что человек не доживет до того, чтобы ими насладиться.

ГИПЕРБОЛИЧЕСКОЕ ДИСКОНТИРОВАНИЕ обсуждалось в главе 6. Здесь отметим лишь тесную связь между головоломками (8) и (4). Причина того, почему люди все-таки вступают в рождественский клуб, предположительно в том, что они знают, что, положив сбережения на обычный счет и намериваясь продержаться там до Рождества, они изменят решение и снимут деньги.

ЭВРИСТИКА. Эвристика (эмпирические правила) может сбить людей с толку. Вера игрока в то, что у колеса рулетки есть память, может возникнуть из эвристики как репрезентативности («Пора выпасть красному»), так и доступности («Сейчас работает красное»). Предпочтение более неприятного шума возникает из-за применения эвристики «пик—конец», согласно которой прошлые опыты оцениваются с точки зрения того, что они собой представляли на максимуме (положительном или отрицательном) и в конце, а не по их общей шкале приятности или неприятности. Такая эвристика может заставить людей предпочесть опыт, который объективно является худшим, поскольку у него лучше конец и точно такой же пик.

ПРИНЯТИЕ ЖЕЛАЕМОГО ЗА ДЕЙСТВИТЕЛЬНОЕ. Феномен принятия желаемого за действительное обсуждался в главе VII. Оно может быть спровоцировано простым желанием, как когда люди с хорошо оплачиваемыми рискованными профессиями преуменьшают риски, которым подвергаются. Еще больше вероятность его возникновения, когда у истоков желания лежит сильная эмоция, например, когда раковые больные выбирают лечение, единственный эффект которого — причинение им лишних страданий. Принятие желаемого за действительное, вызванное желанием подольше жить, также может действовать в «головомомке холодной воды».

НЕСПОСОБНОСТЬ К ПРОЕЦИРОВАНИЮ. В целом ряде ситуаций люди принимают неудачные решения из-за неспособности проецировать себя в будущее. Я имею в виду отсутствие способности вообразить, какие у самого человека или других людей будут основания для убеждений или стимулы к действию. «Проклятие победителя» может быть объяснено этой неспособностью. В качестве другого примера можно вспомнить объявление президентом Шираком досрочных выборов в июне 1997 года.

Возможная причина поражения его коалиции в том, что избиратели понимали, что если Ширак захотел провести досрочные выборы, значит, он знал что-то, чего они не знали, но что давало ему основания опасаться поражения, если он будет тянуть время. Объявив досрочные выборы, он раскрыл *то*, что знал, или по меньшей мере раскрыл, что знает нечто неблагоприятное, и тем самым дал им повод голосовать против себя. Опросы отдавали победу Шираку, но *опросы — это не выборы*, поскольку проведение опроса не раскрывает перед респондентами убеждения лица, заказавшего опрос¹².

ЖЕЛАНИЕ ДЕЙСТВОВАТЬ С ОСНОВАНИЕМ. Я приводил несколько примеров этого механизма в начале главы. В эффекте 11 это желание заставляет агента менять свое поведение при появлении варианта, который, без сомнений, хуже одного из вариантов, которые были в его распоряжении¹³. В других случаях добавление вариантов может помешать агенту принять решение. Поведение, подобное поведению буриданова осла, который умер с голоду, потому что не мог решить, какой из двух идентичных стогов сена начать есть, находило подтверждение в реальных жизненных условиях. Психолог, поставивший на Бродвее прилавок для торговли джемом, выяснил, что если на прилавке было большое разнообразие марок, прохожие больше их рассматривали, но меньше покупали, в сравнении с прилавками, где был представлен меньший выбор марок. При большем числе вариантов сложнее с уверенностью сказать себе: «Этот лучший». Те, кто нуждается в том, чтобы иметь достаточные основания для выбора, от него воздержатся.

Как подсказывает эффект (14), для того чтобы действовать с основанием, нужно иметь это основание, а не просто *знать*, что оно есть. Таким образом, представим, что я достоверно знаю, что обстоятельства *p* или *q* действительны, но не знаю, какое из них действует. Если действительно обстоятельство *p*, я делаю *X*. Если *q*, то у меня тоже есть основание сделать *X*. Соответственно в любом случае у меня есть основание делать *X*, но по-

¹² В пятой части я обсуждаю поведение Ширака как пример того, что можно назвать «синдромом младшего ребенка».

¹³ Такая формулировка может в какой-то степени вводить в заблуждение, потому что в экспериментах присутствовала не одна и та же группа участников, которым сначала предоставили множество (*A*, *B*), а потом множество (*A*, *B*, *C*), но две разные группы, которым то или иное множество вариантов давалось произвольно. Однако было бы естественно предположить, что участники в группе (*A*, *B*, *C*) повели бы себя так же, как участники в группе (*A*, *B*), будь им предоставлен выбор из двух вариантов. Таким образом, модификация относится к условной, а не к реальной базовой линии. Это замечание применимо ко многим экспериментам, упоминаемым в книге.

сколько я не знаю, какое именно основание действительно, я воздерживаюсь от совершения X . Эффект (14) предлагает пример такой аномалии¹⁴. В качестве еще одного примера можно привести тот факт, что в старом английском законодательстве с обвиняемого снимали обвинения, если улики не указывали, что именно он совершил кражу или растрату, хотя, будь одно из этих обвинений доказано, ему вынесли бы приговор. Чтобы это сделать, необходимо признать его виновным либо в p , либо в q . То, что его признают виновным в p или q , будет недостаточно. Еще один возможный выход из подобного затруднения — отложение вынесения решения. Предположим, что американское правительство обсуждает вопрос об установлении торговых отношений со страной, в которой скоро пройдут президентские выборы. Если выиграет проамериканский кандидат, у США будет причина для установления отношений. Если выиграет антиамериканский кандидат, будет набор достаточных оснований, чтобы продолжать ту же политику. Но США могут подождать до выборов, прежде чем определяться со своей политикой. В некоторых обстоятельствах это может оказаться дорогостоящей и бесполезной отсрочкой.

МАГИЧЕСКОЕ МЫШЛЕНИЕ. Механизм магического мышления (глава VII) может объяснить «головоломку холодной воды» и некоторые случаи эффекта дизъюнкции. Тот факт, что люди в большей степени склонны к взаимодействию в «дилемме заключенного», если не знают, как повел себя другой человек, возможно, связан со сверхъестественной верой в то, что своей готовностью сотрудничать они могут подвигнуть партнера действовать так же. «Он такой же, как я, значит, он и действует, как я.» Сходным образом могут складываться намерения избирателя. Если я иррационально полагаю, что мой выбор служит не только индикатором предпочтений других, но некоторым образом влияет на вероятность того или иного исхода, повышенная эффективность моих действий заставляет участие в голосовании казаться рациональным.

КАТЕГОРИЧЕСКИЙ ИМПЕРАТИВ. Есть тесная связь между этим последним примером магического мышления и кантовским категорическим императивом (в повседневной версии), согласно которому должно делать A , а не B , если известно, что всем будет лучше, если все будут делать A , а не B ¹⁵. Тем не менее действие, согласно категорическому императиву, иррационально. Рациональность велит мне делать выбор в зависимости

¹⁴ Точно так же максимальная сумма, которую люди готовы заплатить за потерю из двух вариантов, как было показано, меньше, чем максимум, который они готовы дать за наименее привлекательный из них.

¹⁵ В оригинале: «Поступай так, чтобы максима твоей воли могла стать осно-



РИС. XII.1

от того, что случится, если я сделаю А, а не В¹⁶. Категорический императив заставляет меня делать выбор в зависимости от того, что случится, если *все* сделают А, а не В. На общенациональных выборах даже те, кому не свойственно магическое мышление, могут воздержаться от воздерживания, подумав: «А что если все это сделают?»

ЭМОЦИИ. Для сравнения поведения, основанного на эмоциях, с рациональным поведением мы можем модифицировать рис. II.1, включив в него (выделенное жирными линиями) воздействие эмоций на каждый элемент схемы (рис. XII.1).

Утверждалось, что эмоции могут влиять на *действие* напрямую в случаях слабоволия (глава VI). Медея, убивая детей, чтобы отомстить Ясону, знает, что, *совершая это*, поступает вопреки здравому суждению. Я отмечал, что у меня есть некоторые сомнения на сей счет, но этого нельзя исключать. Эмоции воздействуют на *желания* двояко. Во-первых, посредством связанной с ними тенденции к действию, которая может вызвать временное изменение предпочтений. Если ситуация предполагает задержку между временем принятия решения и непосредственным действием, последнее может так и остаться несовершенным. Приведенный в главе VIII пример показывал, как возросший интерес к службе в армии после 11 сентября 2001 года не привел к увеличению количества призывников. Если действие может быть совершено незамедлительно, то позднее, когда эмоции стихают, оно иногда меняется на про-

вой всеобщего законодательства» (Кант И. Основы метафизики нравственности. М.: Мысль, 1999). — Примеч. ред.

¹⁶ Рациональность не заставляет меня делать выбор в зависимости от того, что случится со мной, если я сделаю А, а не В (см. главу XI), что согласуется с некоторыми видами морали, учитывающей интересы других, но не с моралью, которая представлена категорическим императивом.

тивоположное. Так, из 200 тысяч человек, дезертировавших из армии Конфедерации во время Гражданской войны в США предположительно из-за страха, 10% добровольно вернулись обратно. В некоторых случаях, однако, действие, предпринятое на основании порожденных эмоциями предпочтений, является необратимым (ситуация «ловушка для лобстера»¹⁷). Когда молодые мужчины и женщины на пике энтузиазма вступают в партизанское движение, а потом хотят из него выйти, эта возможность может оказаться для них закрытой. Во-вторых, эмоции могут вызвать временный рост коэффициента дисконтирования во времени, в результате чего ранее не столь привлекательные варианты с плохими долгосрочными последствиями могут показаться предпочтительными.

Эмоции могут воздействовать на убеждения напрямую, через механизмы принятия желаемого за действительное, и через обратное явление. Эмоция гордыни, основанная на представлении о собственном превосходстве, будет сопротивляться признанию совершенной ошибки¹⁸. Этим в некоторых случаях можно объяснить предрасположенность к ошибке необратимых затрат. Кроме того, гордыня позволяет понять, почему люди склонны перекладывать вину на своих жертв¹⁹. Так как их представление о себе не позволяет им признать вину, они находят в другом недостаток, который не только оправдывает то, что было сделано в его отношении, но и мотивируют дальнейшее причинение вреда. Как и в случае с необратимыми затратами, неготовность признать прошлые ошибки приводит к худшим последствиям, чем сами ошибки. Как и для эмоций и принятия желаемого за действительное, мы рассмотрим несколько примеров паники, вызванной страхом, в главе XXIII.

Наконец, в силу безотлагательности (глава VIII) эмоции могут мешать оптимальному сбору информации и тем самым косвенно воздействовать на убеждения. Гнев (головоломка 16) и любовь (18) заставляют нас совершать вещи, которые мы не совершили бы, будь наша политика сбора информации более рациональной. Сенека сформулировал это

¹⁷ Способ финансовой самозащиты компании-эмитента от нежелательного поглощения конвертируемых ценных бумаг, состоящий в установке ограничений на их конвертацию после достижения определенного уровня концентрации в одних руках. — *Примеч. ред.*

¹⁸ В этом контексте любопытно, что животные не склонны к ошибке необратимых затрат. Возможно, потому, что им не нужно заботиться о том, как они выглядят в собственных глазах.

¹⁹ Как я замечаю ближе к концу главы XV, существует тенденция винить жертву действия, совершенного третьими лицами, подобно тому как люди, не имеющие отношения к изнасилованию, валят вину на его жертву. В данном случае действует не гордыня, а вера в то, что, поскольку мир по сути своей справедлив, жертва каким-то образом заслужила свою судьбу.

следующим образом: «Гнев совершенно не уравновешен, он то заходит слишком далеко, то останавливается скорее, чем нужно». Часто безотлагательность вызывает пренебрежение долгосрочными последствиями выбора в силу того, что *само их вычисление требует много времени*. Таким образом, отсечение временного горизонта иногда порождается не более высоким дисконтированием известных последствий, но тем, что некоторые из них даже не возникают на ментальном экране агента.

СОЦИАЛЬНЫЕ НОРМЫ. Эмоции презрения и стыда играют важную роль в поддержании социальных норм (глава XXI). «Поспешная месть» может быть следствием безотлагательности гнева, но может быть связана с социальной нормой, признающей трусом всякого, кто ее откладывает. Нежелание дуэлянтов выбрать оружие, которым они лучше владеют, также основано на опасениях, что их сочтут озабоченными выживанием, а не вопросом чести. Наконец, «парадокс выкошенной лужайки» может быть объяснен не действием социальных норм, а феноменом неприятия потерь. Постоянному жителю даже в голову не придет косить лужайку соседа, потому что в пригородных сообществах существует социальная норма, запрещающая взрослым заниматься этим за деньги. Так просто *не делается*. Голосование тоже может отражать действие социальных норм, если другие наблюдают за самим актом голосования и осуждают тех, кто в нем не участвует. Участие в больших анонимных выборах можно с большей достоверностью считать результатом действия *моральных* норм, которые сами по себе могут иметь иррациональные аспекты (см. предшествующее обсуждение).

Библиографические примечания

Критическую оценку теории рационального выбора с точки зрения, насколько отличной от приведенной выше, можно найти в книге Д. Грина и И. Шапиро «Патологии теории рационального выбора» (Green D., Shapiro I. *Pathologies of Rational Choice Theory*. New Haven, CT: Yale University Press, 1994), полезным образом дополненной «Спором вокруг рационального выбора» (Friedman J. (ed.). *The Rational Choice Controversy*. New Haven, CT: Yale University Press, 1996). Касательно идеи рациональности как нормы см.: Д. Фёллесдал «Статус допущения рациональности в интерпретации и объяснении действия» (Føllesdal D. *The status of rationality assumptions in interpretation and in the explanation of action* // Martin M., McIntyre L. (eds). *Readings in the Philosophy of the Social Sciences*. Cambridge, MA: MIT Press, 1994). Я указываю на гиперрациональность принципа «Действий в интересах ребенка» в главе 3 «Соломонова решения» (Solomonic Judgments. Cambridge University Press, 1989). Полезное обсуждение некоторых смежных вопросов можно найти в работе Дж.

Уинера «Управление ятрогенными рисками риск-менеджмента» (Wiener J. Managing the iatrogenic risks of risk management // Risk: Health. Safety and Environment. 1998. No. 9. P. 39–82). Большинство последующих головоломок обсуждается в статьях, перепечатанных в источниках, указанных в библиографическом примечании к главе VII. Исключение составляет парадокс (1), касательно которого см.: А. Блэйс «Голосовать или не голосовать. Заслуги и ограничения теории рационального выбора» (Blais A. To Vote or Not to Vote: The Merits and Limits of Rational Choice Theory. Pittsburgh: University of Pittsburgh Press, 2000), а также головоломка (2), о которой см.: Д. Канеман «Объективное счастье» (Kahneman D. Objective happiness // Kahneman D., Diener E., Schwartz N. (eds). Well-Being: The Foundations of Hedonic Psychology. New York: Russell Sage, 1999); головоломки (16) и (17), по поводу которых см. главу 3 моей книги «Алхимия разума» (Alchemies of the Mind. Cambridge University Press, 1999), и головоломка (16), которая обсуждается в главе VIII. Большинство альтернатив отражено в тех же источниках, за исключением «принятия желаемого за действительное» (глава VII), «эмоций» (глава VIII) и «социальных норм» (глава XXII). Нужно заметить, что эффект определенности тесно связан с самой первой головоломкой, представленной (в 1953 году) в качестве вызова теории рационального выбора, — «парадоксом Алле». «Головоломка Ширака» вместе со многими другими подобными примерами обсуждается в книге А. Смита «Назначение даты выборов» (Smith A. Election Timing. Cambridge University Press, 2004). Утверждение о том, что животные не совершают ошибку необратимых затрат, защищается в статье Х. Аркса и П. Эйтона «Необратимые затраты и эффект Конкорда. Правда ли, что животные рациональнее людей?» (Arkes H., Ayton P. The sunk cost and Concorde effects: Are humans less rational than lower animals? // Psychological Bulletin. 1999. No. 125. P. 591–600). Пример с продажей джема на Бродвее взят из С. Йенгар и М. Леппер «Когда выбор лишает мотивации. Может ли хорошего быть слишком много?» (Iyengar S., Lepper M. When choice is demotivating: Can one desire too much of a good thing? // Journal of Personality and Social Psychology. 2000. No. 79. P. 995–1006). Ссылка на растрату и кражу в старом английском законодательстве взята из книги Дж. Ф. Стифена «Уголовное право Англии» (Стифен Д. Уголовное право Англии в кратком начертании. СПб., 1865). Данные о вступлении в армию Конфедерации взяты из работы Д. Косты и М. Кана «Дезертиры, социальные нормы и миграция» (Costa D., Kahn M. Deserters, social norms and migration), которая готовится к печати в Journal of Law and Economics. Свидетельства о вызванном эмоцией росте дисконтирования во времени приводятся в работе Д. Тайса, Е. Браславски и Р. Баумайстера «Регулирование эмоциональных расстройств одерживает верх над контролем импульсов» (Tice D., Braslavsky E., Baumeister R. Emotional distress regulation takes precedence over impulse control // Journal of Personality and Social Psychology. 2001. No. 80. P. 53–67).

XIII. Отвечая на рациональность

РАЦИОНАЛЬНОСТЬ «ВТОРОГО СОРТА»

В двух предшествующих главах я рассматривал идеал рационального поведения и частые отклонения от него. Такие отклонения, сколь бы широко они ни были распространены, не являются неизбежными. Если мы понимаем свою предрасположенность к совершению ошибок, мы можем принять меры предосторожности, чтобы уменьшить вероятность их совершения или снизить наносимый ими ущерб. Как я уже не раз отмечал, мы *хотим* быть рациональными, тогда как стратегии предосторожности могут показаться нам формой несовершенной, или «второсортной», *рациональности*. Их следует отличать от простого обучения, которое происходит, когда предрасположенность просто исчезает в результате улучшения понимания. Известно, например, что когда люди осознают бессмысленность голосования, они менее активно участвуют в выборах¹. Когнитивные заблуждения, напоминающие оптические иллюзии, также могут быть преодолены обучением. Точно так же, как мы учимся игнорировать вид палки, которая в воде кажется сломанной, некоторые игроки на собственном тяжелом опыте приходят к пониманию того, что игральные кости не имеют памяти. Здесь, однако, меня интересуют склонности, которые не исчезают с течением времени.

Для того чтобы справиться с нашей склонностью вести себя иррационально, мы можем использовать либо интрапсихические, либо экстрапсихические приемы (*предварительное связывание себя обязательствами*, *precommitment*). Сначала я покажу, как эти технологии используются для борьбы с гиперболическим дисконтированием и порождаемым им противоречивым поведением, а затем коснусь темы их применения в контроле над эмоциональным и аддиктивным поведением. Эти разнообразные стратегии необязательно рациональны, хотя многие из них таковыми являются.

¹ Особенно часто так поступают студенты-экономисты.

БУДУЩИЕ «Я» КАК СОЮЗНИКИ

Агент, осуществляющий гиперболическое дисконтирование и знающий об этом, — *искушенный агент*. В отличие от наивного агента, который снова и снова меняет свое решение, не понимая стоящего за этим механизма, искушенный агент одновременно и осознает свою склонность, и сожалеет о ней. Предвосхищая будущие ситуации выбора между получением небольшой награды сразу и большим вознаграждением позднее, он предпочтет последнее вопреки склонности выбрать первое. В некоторых случаях он может рассматривать свои будущие «Я» как союзников в совместной попытке преодолеть искушение. В других случаях он может видеть в них врагов и пытаться ограничить урон, который они могут нанести его теперешнему «Я». Этот язык, разумеется, является метафорическим, но эта метафоричность будет в дальнейшем устранена.

Рассмотрим сначала случай, когда выбор между более быстрым небольшим вознаграждением и отложенным большим возникает снова и снова. Агент может выбрать отложенное вознаграждение, *сгруппировав* («связав в пучок») различные варианты.

Позволю себе привести пример из тех времен, когда я жил на холме поблизости от университета, в котором преподавал. Каждый год я брал велосипед, чтобы добираться до кампуса. На обратном пути приходилось преодолевать небольшой отрезок, идущий в гору, так что каждый день я боролся с соблазном слезть с велосипеда и пойти пешком, вместо того чтобы заставлять себя крутить педали. При выезде из кампуса у меня было твердое намерение всю дорогу не слезать с велосипеда, но в разгар подъема меня посещала соблазнительная мысль: «Не пройтись ли сегодня пешком, а на велосипеде можно будет и завтра поехать». Затем, к счастью, появляется следующая: «Что такого особенного произойдет завтра? Если я поддамся искушению сегодня, не значит ли это, что я поддамся ему и завтра, и послезавтра, и в дальнейшем?» Последняя мысль позволяла мне не слезать с велосипеда.

Этот интрапсихический прием предполагает *рефрейминг* ситуации. Вместо того чтобы рассматривать поездки домой как *серию выборов*, я начинаю видеть их как *выбор между двумя сериями*: всегда подниматься на холм на велосипеде или всегда идти с велосипедом пешком. Сказав себе, что мое поведение в один из дней идеально предсказывает поведение в другие дни, я запускаю внутренний эффект домино, который поднимает ставки и заставляет меня выбрать отложенную награду улучшения здоровья вместо более быстрой награды устранения дискомфорта. Обратившись к рис. VI.1, можно показать, что если мы разместим множество пар вознаграждений, идентичных А и В, на горизонтальной оси, а потом прочертим две кривые, одну для суммы кривых настоя-

щей стоимости всех небольших вознаграждений, другую, аналогичную — для больших вознаграждений, последняя кривая окажется выше первой в момент совершения первого выбора, при условии что количество последующих выборов велико². Другими словами, в результате группирования вариант постоянного выбора большей награды может оказаться предпочтительнее варианта постоянного выбора меньшей награды. Разумеется, выбор меньшей награды сегодня и больших наград во всех будущих случаях еще лучше, но считается, что агенту не дан этот вариант в его множестве допустимых решений³.

Оправдано ли такое допущение? Является ли мое сегодняшнее поведение индикатором завтрашнего? В случаях, предполагающих реальный каузальный эффект, это может быть верно. Езда на велосипеде сегодня укрепит мои мускулы, так что я смогу и завтра поехать на велосипеде⁴. Однако в моем случае я полагался скорее на магическое мышление, чем на каузальную эффективность. Подобно тому как мысль «кто, если не я?» заставляет многих людей голосовать или жертвовать на благотворительность, меня удерживала на велосипеде мысль «когда, если не сейчас?», точнее, если привести ее в более развернутом виде «В сегодняшнем дне нет ничего особенного. Если я слезу с велосипеда, то при-

² Для примера предположим, что настоящая ценность 1 единицы в момент времени t в будущем — $1/(1+t)$ и что агент дважды сталкивается с выбором между маленькой наградой в 3 единицы и большой наградой в 10. Маленькие награды можно получить в моменты 0 и 6, большие — в моменты 3 и 9. В момент времени 0 настоящая ценность первой большой награды составляет $10/(1+3) = 2,5$, что меньше настоящей (немедленной) ценности маленькой награды. Если выбор делается только на основании этого сравнения, предпочтение будет отдано маленькой награде. Тот же самый выбор по тем же самым причинам будет сделан в момент времени 6. Поскольку сумма настоящих ценностей двух маленьких наград составляет $3 + 3/(1+6) = 3,43$, а сумма настоящих ценностей двух больших наград составляет $10/(1+3) + 10/(1+9) = 3,5$, группирование заставит агента предпочесть две большие награды.

³ Если агент страдает близорукостью при принятии решений (глава XII), группирование может не сработать. Предположим, агент группирует свои выборы задолго до того, как станет доступна первая небольшая награда. В момент группирования настоящая ценность потока больших наград выше, чем ценность потока небольших наград, и агент намеревается дождаться первой большой награды. По мере того как он приближается к моменту, когда станет доступна первая награда, его намерение может сохраниться или не сохраниться. Такое переворачивание предпочтений происходит не из-за гиперболического дисконтирования, но из-за близорукости в принятии решений.

⁴ В качестве аналогии предположим, что, проголосовав, я мог бы повлиять на многих других, кто в противном случае не пошел бы на выборы и не отдал свой голос. Заметим, что при таком допущении магическое мышление никак не задействовано, а работает эффект каузальной мультипликации.

чины, заставившие меня это сделать, будут действовать и завтра и вызовут то же самое поведение. Если я сейчас не сделаю усилия, я не сделаю его никогда». При отсутствии реального каузального эффекта, однако, такого вывода не последует. Если я *могу* сегодня не слезать с велосипеда, но решаю все-таки слезть, я могу не слезать с него и завтра. Хотя это ложное рассуждение, оно убедительно и, как я полагаю, очень широко распространено. Оно показывает, что мы можем призвать одну форму иррациональности (магическое мышление), чтобы бороться с другой (гиперболическое дисконтирование)⁵.

Чтобы хорошо работать, такие стратегии должны быть оформлены как бинарные оппозиции: *всегда* делать это или *никогда* не делать этого. Многим людям полное воздержание дается легче, чем умеренность. Босуэлл замечал: «Джонсон, хотя и мог практиковать строгое *воздержание*, не выказывал *умеренности* ни в еде, ни в питии. Он мог полностью от них воздерживаться, но не мог потреблять умеренно». Такая же проблема возникает, когда вместо того чтобы просто ограничить потребление в каждом отдельном случае, пытаются ограничить количество случаев, когда представляется возможность себе потакать. Стратегия, заранее намечающая законный повод, легко размывается. Когда люди решают не пить вина до обеда, может получиться, что они начнут сдвигать обед на все более ранний час. Если человек взял себе за правило пить вино только в ресторане и не пить дома, он может начать чаще посещать ресторан. Кантовское правило выкуривать после завтрака только одну трубку (глава IV) не было настолько недвусмысленным, чтобы обеспечивать ему полную защиту; с течением времени он стал покупать трубки все бóльших размеров. Когда это осуществимо, правило «Никогда этого не делать!» может оказаться единственным твердо соблюдаемым. Поскольку к еде такая политика неприменима, ожирение оказывается более устойчивым феноменом, чем приверженность к любым частным правилам.

При этом рамка с бинарной оппозицией может вызвать крайне негибкое поведение. Предположим, я сказал себе, что никогда не буду делать никаких исключений из правила ежедневно чистить зубы. В какой-то момент я остаюсь без зубной щетки и решаю пройти пять миль в метель, чтобы ее купить. Чтобы поддержать это решение, я говорю себе, что если я нарушу правило в этот раз, то вступлю на скользкую дорожку, которая приведет меня к его нарушению и по более мелким поводам, в конце концов никакого правила не останется и зубы у меня выпадут. Некото-

⁵ При голосовании цель магического мышления в том, чтобы помочь нам преодолеть склонность к социально вредному рациональному поведению, а не противодействовать иррациональности.

рые люди конструируют весьма изощренные системы такого рода, в которых нарушение одного правила предсказывает нарушение других, тем самым все больше поднимая ставки⁶. Поскольку частные правила могут иметь подобный эффект, сводящий на нет их действенность, иногда даваемое ими лекарство хуже самой болезни. На языке Фрейда (глава IV), жесткий контроль импульсов, осуществляемый «сверх-Я», может нанести больший урон, чем сами исходящие от «Оно» импульсы.

БУДУЩИЕ «Я» КАК ПРОТИВНИКИ

Рассмотрим теперь случай, в котором агент сталкивается с выбором между наградами (наказаниями) в один из нескольких дней в будущем. (В отличие от предыдущего случая, предполагается, что выбор проводится разово.) Он может воспользоваться интрапсихическим приемом стратегической реакции на известную склонность будущих «Я» гиперболически дисконтировать будущее. Предположим, я — гиперболический откладыватель, который всегда переносит поступки на завтра, а затем, когда завтра наступает, продлевает отсрочку до послезавтра. Как только я понимаю, что у меня есть такая склонность, мое оптимальное поведение меняется. Предположим, я могу выполнить данную неприятную задачу в любой из трех периодов времени и издержки на ее выполнение со временем возрастают. Если я наивен, я могу сказать себе, что выполню эту задачу завтра; если искушен, я знаю, что завтра я снова отложу ее до последнего момента. Понимание того, что если я не сделаю это прямо сейчас, за это, возможно, придется заплатить очень высокую цену, может заставить меня действовать⁷.

Искушенность в данном случае помогает. В других случаях, наоборот, полезной может оказаться наивность. Предположим, что вы можете получить награду в любой из трех последующих периодов времени и что

⁶ Они могут ошибочно полагать (глава X), что кросс-ситуационная последовательность может вызвать кросс-ситуационное провоцирование срывов.

⁷ Для примера предположим, что настоящая ценность 1 единицы полезности в момент времени t в будущем составляет $1/(1+t)$ и что, откладывая визит к стоматологу, я буду только усугублять мои страдания: если я пойду сегодня, то испытаю боль $-2,75$, завтра — боль, равную -5 , а через день -9 . Из сегодняшней перспективы настоящие ценности составляют соответственно $-2,75$, $-5/(1+1)=2,5$ и $-9/(1+2)=-3$. Таким образом, может показаться, что оптимальным сегодняшним решением было бы перенести визит на завтра. Однако в силу искушенности я знаю (сегодня), что завтра настоящая ценность завтрашнего визита будет составлять -5 , а послезавтрашнего $-9/(1+1)=-4,5$, включая предпочтение подождать до послезавтра. Но поскольку сегодня я предпочитаю сегодняшний визит визиту послезавтра, я пойду сегодня.

награды с течением времени растут. Возьмем человека, которому в награду обещано вино, в течение трех лет улучшающее, а потом ухудшающее свои качества. У наивного человека может возникнуть намерение подождать до третьего периода, а затем он меняет решение и выпивает вино во второй период. Искушенный человек будет знать, что он никогда не дожидается третьего периода, так что фактически он имеет дело с выбором между наградой первого периода и наградой второго. При таком выборе более быстрая награда может победить⁸. По словам некоторых алкоголиков, они тоже склонны иногда так рассуждать: «Я знаю, что все равно поддамся искушению, так почему бы не сделать это прямо сейчас?» Точно так же наивные курильщики, первый раз бросающие курить, могут добиться большего успеха в сравнении с более опытными и искушенными курильщиками, которые знают, как мало у них шансов на успех. Хотя соскальзывание обратно в зависимость необязательно предполагает гиперболическое дисконтирование. Общее положение таково: если вы можете предсказать, что отклонитесь от наилучшего плана, вы можете отклониться от него даже в большей степени или раньше, чем если бы не осознавали возможность неудачи.

ЭКСТРАПСИХИЧЕСКИЕ ПРИЕМЫ

На практике изошренное планирование против своих будущих «Я», вероятно, менее важно, чем стратегии группирования и приемы предварительного связывания себя обязательствами, к которым я сейчас перейду. Они предполагают такое воздействие на внешний мир, которое не может быть сразу или безболезненно устранено, тем самым снижая вероятность того, что человек в будущем выберет более быструю небольшую награду. Здесь выделяются пять стратегий: *устранение выбора* более быстрой награды из множества допустимых решений, *введение штрафа* за выбор более быстрой награды, *добавление премии* за выбор

⁸ Для примера предположим, что настоящая ценность 1 единицы полезности в момент времени t в будущем составляет $1/(1+t)$ и что я могу все больше увеличивать свою выгоду, откладывая открытие бутылки вина. Если я выпью его в этом году, я получу удовольствие, равное 2,75, в следующем году оно составит 5, а через год — 9. Из сегодняшней перспективы настоящая ценность составляет соответственно 2,75, $5/(1+1) = 2,5$ и $9/(1+2) = 3$. Таким образом, может показаться, что оптимальный выбор, с точки зрения первого года, — выпить вино в следующем. Однако благодаря искушенности я знаю (в этом году), что в будущем году настоящая ценность выпивания вина в следующем составит 5 единиц, а еще через год она составит $9/(1+1) = 4,5$, тем самым возникает предпочтение выпить вино в следующем году. Однако поскольку в первый год я предпочитаю выпить его в первый, а не на следующий год, я выпью его сегодня.

отложенной награды, *введение отсрочки* между выбором и фактическим получением награды и *уклонение от сигналов*, которые могут вызвать переворачивание предпочтений. В качестве примера действия первых четырех вариантов можно привести откладывание денег. Если я начну их откладывать на Рождество, но потом окажется, что я снимаю деньги с накопительного счета, вместо того чтобы их там хранить, я могу присоединиться к рождественскому сберегательному клубу, который не допускает досрочный отзыв вклада (глава XII). Или же я могу поместить свои сбережения на счет с высоким процентом годовых, который предполагает санкции за досрочное закрытие, тем самым объединяя в себе премию и штраф. Если я хочу откладывать деньги на старость, я могу установить отсрочку между решением воспользоваться сбережениями и тем моментом, когда деньги будут в наличии, например, вложив их в неликвидные активы, а не в акции и облигации. Пятый вариант можно проиллюстрировать случаем, когда тяга к десерту провоцируется у человека визуальными сигналами. Трюк состоит в том, чтобы пойти в ресторан, в котором не возят между столами тележку с десертами, а нужно заказывать десерт по меню. Противоположную позицию занимает человек, испытывающий проблемы с десертами из-за гиперболического дисконтирования. Для него наилучшее решение — пойти в ресторан, где десерт нужно заказывать в начале обеда или ужина.

Люди, записывающиеся на еженедельный фитнес, обычно бросают его через неделю-две. Чтобы это предотвратить, они могут (по крайней мере теоретически) подписать с фитнес-центром договор, чтобы заплатить вдвое больше положенного и получать переплаченную сумму по частям каждый раз, когда они приходят на занятия. Люди, записывающиеся на программы для похудения, должны оставлять залог, который они смогут получить только тогда, когда сбросят указанное количество килограммов, иногда с оговоркой, что если им не удастся этого сделать, деньги пойдут на поддержку самой ненавистной им политической силы. Когда я начал читать лекции, положенные в основу этой книги, я заранее ограничил себя обязательством, сказав студентам, что в конце каждой недели буду давать им тезисы главы. Если бы не смог сдержать этого обещания, я страдал бы из-за их насмешек. Если я опасаясь, что отменю поход к стоматологу, я могу попросить его взять с меня денег в два раза больше обычного в случае, если я не явлюсь в назначенный день. В ситуации с вином, которое улучшается со временем, вы можете попросить продавца поддержать его у себя, тем самым предохраняя себя от преждевременного получения удовольствия. Если вы боитесь, что прочтете последний роман вашего любимого автора детективов слишком быстро, пропуская какие-то места, чтобы быстрее добраться до развязки, вам нужно купить аудиOVERSIYU (и плеер, в котором нет функции промотки

вперед), который не оставляет вам другого выбора, кроме как слушать каждое слово.

Примеры в двух последних абзацах предполагают предварительное самоограничение в отношении двух видов искушения. С одной стороны, это *прокрастинация* (откладывание на потом), включая неспособность экономить, проходить болезненные лечебные процедуры, делать физические упражнения или писать книгу; с другой стороны — *преждевременное удовлетворение*, когда слишком рано выпивают вино или пропускают страницы в книге. Эти искушения напрямую происходят от гиперболического дисконтирования. Все происходит исключительно из-за того, что проходит время. В следующей категории случаев, *неумеренном поведении*, гиперболическое дисконтирование может взаимодействовать с инстинктивными мотивациями. Сюда входит избыточная реакция, пристрастие к азартным играм и аддиктивное поведение. В таких случаях иногда трудно понять, вызвана инверсия предпочтений структурой дисконтирования или другими факторами. Решение поститься, принятое на сытый желудок, может развеяться, как только человек почувствует голод. Намерение бросить курить может пошатнуться при виде другого человека, зажигающего сигарету. Это феномен *зависимости от сигналов* — потребности, провоцируемые визуальными образами, ассоциирующимися с потреблением вызывающих привыкание веществ. Намерение бросить азартные игры, вызванное чувством вины игрока перед своей семьей, может ослабнуть, как только эмоция начнет терять силу (глава VIII). Кроме того, иногда трудно определить, имеем мы дело с откладыванием на потом или с инстинктивными факторами. Решение регулярно принимать лекарства может быть подорвано спадом сильных эмоций, которые прежде заставили обратиться к врачу.

Как только агент понимает, что находится под воздействием этих механизмов, он может связать себя предварительным обязательством, чтобы предупредить их действие. Чтобы помешать голоду подорвать решимость сесть на диету, он может принять таблетку, приглушающую потребность в еде. Еще более радикальным стало бы решение перевязать челюсть так, чтобы можно было потреблять только жидкую пищу. Если агент знает, что его потребность в десерте зависит от сигналов, он не пойдет в ресторан, где есть тележка с десертами⁹. Бывшие наркоманы с героиновой зависимостью будут воздерживаться от посещения мест, где они в свое время принимали наркотики. Бывшие игроки учатся не ходить в казино, «чтобы только взглянуть на то, как играют дру-

⁹ Если мы сравним этот случай с тем, когда десерт нужно заказывать в начале обеда, мы увидим, что в обоих случаях речь идет о защите себя от эффекта близости искушения — временной или пространственной.

гие». Если агент может предсказать, что его гнев утихнет и он не захочет наказывать обидчика, он может исполнить наказание немедленно. Такое поведение можно было наблюдать в Бельгии после войны. На основании опыта времен Первой мировой многие бельгийцы считали, что через какое-то время всеобщее желание вынести суровый приговор тем, кто сотрудничал с врагом, уступит место безразличию, поэтому они хотели, чтобы суды прошли как можно быстрее, до того как эмоции утихнут.

При борьбе с зависимостями очень распространена стратегия принятия на себя издержек. Когда генерал де Голль хотел бросить курить, он рассказал об этом своим друзьям, чтобы увеличить цену возврата к этой привычке. В центре по лечению кокаиновой зависимости в Денвере врачам предоставляется возможность написать письмо с саморазоблачениями, адресованное Национальному совету медицинских экзаменаторов, в котором они признаются, что принимают наркотики, и просят отозвать у них лицензию. Письмо посылается автоматически, после положительного анализа на наркотики. Некоторые бывшие алкоголики пытаются сохранять трезвость, принимая дисульфирам — лекарство, которое при принятии алкоголя вызывает болезненные ощущения¹⁰.

Добровольные отсрочки также могут быть эффективными в борьбе с чрезмерными желаниями. Чтобы помешать импульсивному пьянству, я могу запереть алкоголь в сейф, оборудованный таймером, или же могу не держать алкоголь дома, с тем чтобы мне пришлось идти за ним в магазин. Технология дисульфирама фактически сочетает принятие издержек с установлением отсрочки, поскольку после принятия таблетки должно пройти два дня, прежде чем вы сможете выпить и не заболеть. В лечении от кокаиновой зависимости издержки сочетались с отсрочками. Такой метод позволял людям разорвать контракт, подав нотариально заверенное заявление. Существует двухнедельная отсрочка. Всякий, кто подает заявление, через две недели может отозвать компрометирующее письмо. Но если в течение этих двух недель он решает вернуться обратно, то выдача письма откладывается еще на две недели. Хотя многие пациенты заявляли о выходе из договора, все они в течение двух недель отзывали свое заявление.

Озабоченность предварительным самоограничением в борьбе с временной непоследовательностью и неуравновешенностью — сравнительно новое явление. Писавшие на эту тему классики сосредоточивались

¹⁰ Самая распространенная форма приема — оральная, в результате чего прием алкоголя вызовет у вас неприятное состояние. Имплантация под кожу (неэффективная физически, но эффективная психологически) менее распространена.

на самоограничении *страстей*, понимаемых широко, с включением в них опьянения и психотических состояний¹¹. В «Одиссее» Гомера приводится пример предварительного связывания себя обязательствами, ставший классическим: Одиссей заставляет привязать себя к мачте корабля так, чтобы он не смог ответить на зов сирен. В работе «О гневе» Сенека писал: «Пока мы здоровы, пока в своем уме, давайте попросим спасения от мощного зла, которому часто сами потакаем. Те, кому вино ударяет в голову, так что они боятся, что после станут грубыми и дерзкими, пусть попросят друзей увести их с пира. Те, кто понял, что теряет разум, будучи болен, пусть отдадут приказ, чтобы их не слушались во время их болезни». В романе мадам де Лафайет «Принцесса Клевская» героиня оставляет двор, чтобы избежать искушения поддаться ухаживаниям герцога Немурского; даже позднее, когда ее муж умер и она может снова выйти замуж, она не возвращается ко двору. «Зная, как обстоятельства влияют на самые мудрые решения, она не хотела подвергать себя риску изменить свое или вернуться туда, где живет человек, которого она любила». В романе Стендаля «Люсьен Левен» госпожа де Шастеле из предосторожности встречается с Люсьеном только в присутствии компаньонки, чтобы ей пришлось слишком дорого заплатить, если она все-таки уступит любовному порыву.

Эти стратегии широко распространены. Иногда люди сжигают за собой мосты в стратегических целях (глава XIX), но иногда они делают это, чтобы не дать себе возможности уступить страху. Я могу сторониться вечеринок в офисе, потому что знаю по прошлому опыту, что скорее всего выпью и раскрепощающий эффект алкоголя подтолкнет меня к агрессивному поведению или ухаживаниям, о которых я позднее пожалею. В качестве альтернативы я могу взять с собой жену, чтобы повысить издержки такого поведения. *Простое решение* не пить (или сдерживать эмоции, если пью) менее эффективно, учитывая силу обстоятельств (глава X). Точно так же управление гневом при помощи интрапсихического приема, предписывающего сосчитать до десяти, прежде чем ответить, предполагает такую отстраненность, которая, как правило, отсутствует в запале. Общий совет, который подсказывает практика самосовершенствования, — «как можно раньше разорвать цепь», а не полагаться на самообладание перед лицом искушения или провокации. Марк Твен говорил: «Легче не попадать в беду, чем пытаться потом из нее выбраться». В качестве последнего примера хочу привести заголовок в «Нью-Йорк Таймс» (5 апреля 1996 года): «Техас согла-

¹¹ Гораздо реже встречается кратко упомянутый в тексте феномен связывания себя предварительными обязательствами, будучи в состоянии страсти и из страха того, что она утихнет.

сил на проведение хирургической операции для растлителя: мужчина, готовящийся выйти на свободу, хочет кастрации, чтобы умерить свои сексуальные порывы».

Стратегии откладывания могут показаться самыми многообещающими в борьбе с иррациональностью, основанной на эмоциях. Поскольку эмоции недолговечны, эффективным лекарством могло бы стать любое препятствование немедленной реализации побуждения к действию. Как я отмечаю далее, власти действительно полагаются на эту особенность эмоций, когда требуют от людей подождать какое-то время, прежде чем принимать важное решение. Однако случаи, когда люди *сами для себя* устанавливают отсрочки, чтобы противодействовать страсти, наблюдаются редко. Они могут попросту не владеть соответствующей технологией. В качестве примера можно привести институт договорного брака, который существует в трех американских штатах (Арканзасе, Аризоне и Луизиане) как опциональная форма, которую труднее заключить и расторгнуть, чем обычный брак. Как правило, пара, вступившая в договорный брак, может получить развод лишь через два года после того, как она разошлась; для сравнения: срок ожидания развода для обычного брака — полгода. Немногочисленное меньшинство (менее 1%) брачующихся, пользующихся этим вариантом, предположительно делает это, чтобы продемонстрировать друг другу свою решимость и защититься от мимолетных страстей и искушений.

Предварительное связывание себя обязательствами предполагает помощь со стороны других лиц, организаций или властей. Они же в свою очередь должны сохранять независимость от инструкций агента, так как в противном случае он может их отменить. Чтобы побороть пристрастие к опиуму, Сэмюэль Кольридж нанял человека, который мешал бы ему заходить в аптеку. Когда этот человек попытался его удержать, Кольридж сказал: «Ерунда! Возникло чрезвычайное, непредвиденное обстоятельство. Неважно, что я говорил вам раньше. Сейчас я говорю вам, что если вы не уберете руку с дверного косяка этого почтенного аптекаря, я получу основания для того, чтобы подать на вас в суд за нападение и избиение». Похожим образом Мао Цзэдун распоряжался игнорировать все приказы, которые он может отдать после приема снотворного. Когда, приняв снотворное, он приказал помощнику послать приглашение американской команде по настольному теннису посетить Китай (начало американо-китайских отношений), а его помощник спросил: «Ваши слова имеют значение?», он ответил: «Да, имеют. Давай побыстрее, а то не будет времени».

Организации являются более надежным инструментом для связывания себя обязательствами. Денверский наркологический центр или рождественские клубы, специально созданные для того, чтобы помочь лю-

дям справиться с этой проблемой¹², предлагают варианты добровольного взятия на себя обязательств, которые индивиды не могут принять самостоятельно, и препятствуют отмене данных инструкции. В Норвегии есть Закон о защите психического здоровья, позволяющий индивидам *по собственной воле, но без права пересмотра решения* проходить трехнедельное лечение в психиатрической лечебнице. Но эта система, кажется, не работает, потому что врачи имеют право (но не обязанность) удерживать таких лиц в клинике. Чтобы закон стал действенным, следует предоставить пациентам право подавать на клинику в суд, если она по их требованию отпустила их раньше положенного срока.

С 1996 года в штате Миссури практикуется программа самоисключения для игроков в азартные игры: человеку пожизненно запрещается входить в любое казино, расположенное на теплоходе. Если он попытается обойти этот запрет и сыграть на одном из таких теплоходов, его должны будут оттуда удалить, «держатель лицензии будет сотрудничать с агентом комиссии, сообщив о происшествии соответствующей карательной инстанции и требуя предъявления обвинений... в уголовном правонарушении, проступке класса В». Исключенным игрокам следует отказывать в получении выигрышей, если им удалось каким-то образом проникнуть на борт теплохода, сделать ставки и выиграть. Государство может играть еще более активную роль, вводя отсрочки на аборт, покупку оружия или развод (и оформление брака!) и предоставляя потребителям трехдневный или недельный период ожидания, в течение которого они могут отменить покупки, сделанные в приступе энтузиазма. Это, однако, примеры *государственного патернализма*, а не *самопатернализма при поддержке государства*.

Иногда конституции понимаются в качестве таких механизмов по предварительному связыванию себя обязательствами или как форма *коллективного самопатернализма*. Джон Поттер Стоктон писал в 1871 году: «Конституции — это цепи, которыми люди приковывают себя, будучи в здравом уме, чтобы спастись от собственного самоубийственного порыва в дни безумия». Еще одна распространенная метафора: конституция — это обязательства, которые трезвый Питер налагает на пьяного Питера. Часто в качестве примера предварительного связывания себя обязательствами приводят двухпалатную структуру парламента: когда закон проходит через обсуждение в двух палатах, это позволяет выиграть время на то, чтобы страсти улеглись и возобладал разум (или интерес!). Этот же аргумент приводился в оправдание отсрочек при вступлении конституционных поправок в силу. Однако если предвари-

¹² В отличие от сейфов с таймером, не предназначенных для того, чтобы помогать людям бороться с алкоголизмом.

тельное связывание себя обязательствами понимается как самоограничение, его распространение с индивида на коллектив и с членов одного поколения на разные поколения довольно сомнительно. Получается тогда, что это не сообщества связывают себя обязательствами, а большинство налагает обязательства на меньшинства или настоящее поколение — на будущее. Более того, поскольку конституции обычно пишут в беспокойные времена, основатели и сами часто находятся в пылу страсти. Будучи сами «в опьянении», они могут проглядеть необходимость принятия мер против «пьянства». Так, 7 сентября 1789 года, когда в Национальном учредительном собрании Франции обсуждали вопрос о том, каким должен быть парламент — однопалатным или двухпалатным, депутат Адриан Дюкнун сделал следующую запись в своем дневнике: «Если бы можно было оценить вероятность, ясно, что большинство собрания никогда не проголосовало бы за двухпалатную систему. Подобный исход мог бы иметь большие недостатки, но ситуация такова, и все умы так взбудоражены, что иначе и быть не могло; возможно, через несколько лет это можно будет изменить. Тогда мы придем к пониманию, что единая ассамблея для народа, столь пылкого, как мы, может иметь самые ужасные последствия».

Библиографические примечания

Интрапсихический прием группирования широко обсуждался в работах Дж. Эйнсли (Ainslie G.), в частности в «Пикоекономике» (Picoeconomics. Cambridge University Press, 1992). Попытка дать рациональные объяснения таким «частным правилам» предпринимается в статье Р. Бенабу и Дж. Тироля «Сила воли и личные правила» (Bénabou R., Tirole J. Willpower and personal rules // Journal of Political Economy. 2004. No. 112. P. 848–886). Обсуждение близорукости при принятии решений основывается на работе О.-Дж. Скога «Гиперболическое дисконтирование, сила воли и зависимость» (Skog O.-J. Hyperbolic discounting, willpower, and addiction // Elster J. (ed.). Addiction: Entries and Exits. New York: Russell Sage Foundation, 1999). Стратегические реакции искушенных индивидов, которые осознают свою склонность к гиперболическому дисконтированию, обсуждаются в работе Т. О'Донохью и М. Рабина «Сделать сейчас или потом» (O'Donoghue T., Rabin M. Doing it now or later // American Economic Review. 1999. No. 89. P. 103–124). Идея предварительного и добровольного связывания себя обязательствами для борьбы с иррациональными склонностями обсуждается в работе Т. Шеллинга «Эгономика, или Искусство управления собой» (Schelling T. Egonomics, or the art of self-management // American Economic Review: Papers and Proceedings. 1978. No. 68. P. 290–294) и в некоторых более поздних публикациях этого автора. Я касался этого в «Одиссее и сиренах» (Ulysses and the Sirens. Rev. ed. Cambridge University Press,

1984), в «Освобожденном Одиссее» (Ulysses Unbound. Cambridge University Press, 2000) и в «Не сжигайте ваш мост, пока по нему не пройдет. Сложности и неясности предварительного связывания себя обязательствами» (Don't burn your bridge before you come to it: Ambiguities and complexities of precommitment // Texas Law Review. 2003. No. 81. P. 1751–1788). Существует книга, целиком посвященная проблеме приема лекарств, — Ж. Риш «Зачем лечиться» (Reach G. Pourquoi se soigne-t-on. Paris: Editions de Bord de l'Eau, 2005). Историю о Кольридже я нашел у Томаса де Квинси в «Исповеди английского морфиниста» (Quincy T. de. Confessions of an Opium Eater. London: Penguin Books, 1968. P.145). История о Мао Цзэдуне взята из книги «Мао: неизвестная история» Дж. Чана и Дж. Халлидея (Chang J., Halliday J. Mao: The Unknown Story. New York: Knopf, 2005. P. 580–581). Если в «Одиссее и сиренах» я полон энтузиазма по отношению к идее конституции как механизма предварительного связывания себя обязательствами, то в «Освобожденном Одиссее» я отрекаюсь от этой точки зрения.

XIV. Некоторые следствия для интерпретации текстов

По общему признанию науки делятся на три различные части: гуманитарные, социальные и естественные. Для одних целей это разделение научного поля полезно, но для других жесткое разделение может помешать взаимобогащению. В четвертой части я обосновываю возможность извлечения социальными науками пользы из биологических исследований человека и животных. В этой главе я отстаиваю гораздо большую, чем принято считать, общность гуманитарных и социальных наук. В частности, я попытаюсь показать, что *интерпретация* произведения искусства и *объяснение* — это тесно взаимосвязанные операции. Понять произведение искусства — значит объяснить его с точки зрения предшествующих ментальных состояний ее создателя. *Успешное* произведение искусства то, которому может быть дано объяснение с точки зрения рационального выбора. Но я собираюсь оспорить то, что называется интерпретацией с точки зрения последствий и позднее получит объяснение. Объяснение, которое я собираюсь предложить, охватывает не все виды искусства. Даже внутри литературы, которой я хочу ограничиться, оно имеет смысл только для классических (созданных до 1850 года) романов и пьес, в которых действует негласная конвенция, что описываемые персонажи *могли быть реальными*¹.

Сначала рассмотрим рациональность как мотив *героев* беллетристики и драматургии. Классическая проблема в литературоведении — почему Гамлет откладывает месть за смерть отца. Предлагалось множество объяснений. Некоторые из них апеллировали к иррациональности, будучи даны с точки зрения слабости воли или клинической депрессии. Однако есть простое объяснение с точки зрения рационального выбора. Хотя Гамлет поначалу верит, что тень его отца сказала ему правду о Клавдии, позднее он решает *собрать дополнительную информацию*, разыграв пьесу, чтобы «обнажить нечистую совесть короля». Когда реакция Клавдия подтверждает его предположения, он лишен *возможности* осуществить свое желание и обрець врага на вечные адские муки. Ему представляется случай убить короля, когда тот молится. Если бы он это сделал, то, согласно поло-

¹ Раннее нарушение этой конвенции можно найти ближе к концу «Пер Гюнта» Ибсена (1867), когда Пер боится утонуть и «странный пассажир» говорит ему, что «в середине пятого акта не умирают».

жениям современной ему теологии, он обеспечил бы королю скорее спасение, а не проклятие. Позднее он убивает Полония за занавесом, *ошибочно, но не иррационально* полагая, что это король. Учитывая информацию, которая у него была, его вера в то, что за пологом прятался король, была рациональной. Более того, у него не было причин собирать *большее* количество информации, поскольку он мог сделать вполне разумный вывод, что если некто прячется за пологом в присутствии королевы, то это король.

Я не утверждаю, что это правильная интерпретация (я еще не сказал, что следует понимать под правильностью интерпретации). Я просто хочу сказать, что три упомянутых мною эпизода, на первый взгляд, не противоречат той идее, что Гамлет рационально движется к цели отмщения за убийство отца. Другой вопрос — совместима ли эта идея с его постоянными самообвинениями в отсутствии решимости отомстить? Многие комментаторы истолковывают эти знаменитые монологи как знак слабо-волия и считают, что в основе двух первых эпизодов лежат обманчивые оправдания для бездействия (четвертый эпизод труднее подогнать под эту теорию). Но хотя слабоволие и самообман нарушают каноны рациональности, они совершенно *интеллигибельны* (см. третью часть). Что касается внутреннего развития произведения искусства, для интерпретации гораздо полезнее идея интеллигибельности, чем идея рациональности.

В противовес внутренней точке зрения, мы можем занять внешнюю позицию автора. Тогда на вопрос: «Почему Гамлет откладывает свою месть до пятого акта?» — мы можем ответить: «Смерть короля должна произойти в пятом акте»². Это вопрос драматургии, а не психологии. Сам по себе такой ответ был бы неудовлетворительным. Если бы Шекспир ввел месть через серию произвольных событий или совпадений только ради того, чтобы она произошла в конце пьесы, мы сочли бы это его драматургическим провалом. Точнее, это был бы случай *авторской иррациональности*.

Авторская рациональность подобна Божественной. Автор запускает процесс, в котором каждое событие может быть *объяснено дважды* — сначала каузально, а потом телеологически. Я взял эту идею у Лейбница, который писал:

Это как бы два царства: одно — причин действующих, другое — конечных. Но с точки зрения их общего происхождения одного из них недостаточно без другого, ибо они вытекают из одного источника, в котором соединены сила, производящая причины действующие, и мудрость, управляющая причинами конечными³.

² В отличие от Ибсена, говорившего устами «странного пассажира» (см. предыдущее примечание), Шекспир не мог заставить Гамлета сказать: «Я не могу убить короля до пятого акта».

³ Лейбниц Г. В. Сочинения: в 4 т. М.: Мысль, 1982–1985. Т. 1. С. 373. — Примеч. ред.

Цель бога — создать лучший из возможных миров. Если конкретизировать эту идею, включив в нее временное измерение, ее можно истолковать как идею создания лучшей из возможных последовательностей. Хотя переход из одного состояния универсума в другое происходит посредством обычной физической каузальности, изначальное состояние и сами законы каузальности были избраны таким образом, чтобы максимизировать общее совершенство последовательности.

В классической драме или романе задача автора в том, чтобы развивать сюжет посредством диалогов и поступков героев, зачастую реагирующих друг на друга, чтобы при этом максимально повышалась эстетическая ценность. Каждое действие или заявление того или иного героя может быть объяснено дважды — одновременно как реакция на предшествующие действия и высказывания (или внешние события) и как генератор удивления, напряжения и, наконец, разрешения этого напряжения в читателе. Первое объяснение покоится на интеллигибельности героев, второе — на рациональности автора, понятой в том смысле, который я сейчас попытаюсь разъяснить.

Тот факт, что авторы часто создают множество черновых вариантов, пока не будут полностью удовлетворены или прежде чем не положат перо, является неоспоримым доказательством того, что они вовлечены в процесс *выбора* и что у них есть эксплицитный или имплицитный критерий *лучшего*. Тот факт, что черновики обычно содержат *небольшие* вариации, подсказывает, что они нацелены на *локальный максимум* формы лучшего, к которому они стремятся. Однако различие между автором и тем, кто просто взбирается на один градус вверх, в том, что у автора *творческий процесс* выходит за рамки простого выбора⁴. Литературное произведение не может быть сведено к рациональному выбору, потому что количество осмысленных последовательностей слов слишком велико, чтобы один человек мог их все изучить и выбрать наилучшее. Хотя рациональный творец может упростить задачу, намеренно исключив некоторые последовательности (в этом одна из функций метра и рифмы в стихе); все равно останется слишком много вариантов для того, чтобы выбор мог стать действующим механизмом отбора. Вместо этого автору придется полагаться на ассоциативный механизм своего бессознательного.

Рациональное творчество, таким образом, состоит в том, чтобы получить право на вторую цифру после запятой или, если воспользоваться другой метафорой, взобраться на вершину ближайшего холма. Если продолжать говорить метафорами, то это задача левого полушария. За-

⁴ По той же самой причине обречены на провал попытки найти объяснение техническим изменениям с точки зрения рационального выбора.

дача правого, состоящая в том, чтобы правильно определить первую цифру после запятой или найти холм, возвышающийся над другими, находится вне сферы рациональности. И все же даже будучи редуцирована до тонкой настройки, авторская рациональность имеет значение. Как явствует из фразы «маленький шедевр», иногда лучше найти вершину небольшого холма, чем застрять на склонах холма повыше. Я не хочу проводить сравнений, но «Хроника объявленной смерти»⁵ и «Взгляни на дом свой, ангел»⁶ служат для иллюстрации этих двух возможностей.

Позвольте мне перечислить некоторые требования, которые рациональность предъявляет к автору. Во-первых, поступки и слова героев должны быть интеллигибельны. Во-вторых, автор должен соблюдать требования *полноты* и *экономии*. В-третьих, произведение должно течь *равным потоком*, то есть должно быть минимизировано обращение к случайностям и совпадениям. В-четвертых, оно должно предлагать психологически удовлетворительную модель создания и разрешения напряжения.

Интеллигибельность может быть абсолютной или относительной, а во втором случае еще и глобальной или локальной. Вопрос абсолютной интеллигибельности: может ли *хоть одно* человеческое существо вести себя подобным образом? Вопрос относительной глобальной интеллигибельности: не противоречит ли поведение вымышленного персонажа его (или ее) общему характеру, как он был ранее изображен в произведении? Если требования абсолютной или относительной локальной интеллигибельности являются ключевыми ограничениями для авторской рациональности, требование относительной глобальной интеллигибельности не является таковым. Несоблюдение этого последнего требования может рассматриваться как серьезный эстетический недостаток.

В некоторых случаях абсолютная интеллигибельность может нарушаться избытком рациональности. Обратимся к Медее Еврипида или к Федре Расина. Обе прекрасно осознают разрушительный характер своих страстей. Обе изображаются как жертвы слабости воли в строгом смысле этого слова, понимающие: то, что они делают, противоречит здравому суждению, которое они выносят *в момент действия*. Хотя страсть приводит к отклонениям от этого суждения, оно не подвержено его воздействию. Гермiona Расина — более правдоподобный персонаж. Поскольку ее суждение затуманено эмоциями, она жертва скорее самобмана, чем слабости воли. Я предполагаю (и это не более чем предположение), что одновременное присутствие сильной эмоции и полной когнитивной ясности противоречит человеческой природе.

⁵ Роман Габриэля Гарсии Маркеса 1981 г., издан в России под названием «История одной смерти, о которой знали заранее». — *Примеч. ред.*

⁶ Роман Томаса Вулфа (1929).

Если чрезмерная рациональность может быть неинтеллигибельной, иррациональность в свою очередь может оказаться совершенно интеллигибельной. Что может быть понятнее, чем реакция господина де Ренала в «Красном и черном» Стендаля, который, получив недвусмысленные указания на то, что у его жены роман с Жюльеном Сорелем, предпочитает оставаться в неведении? Есть более парадоксальные случаи, когда желание, чтобы жена была верна, вызывает веру в ее неверность вопреки фактам. В «Отелло»: «Ведь для ревнивца вздор — /Такой же сильный довод, как святое/Писание»⁷. Первый случай — это короткое замыкание, второй — касание проводов (глава III).

Относительная интеллигибельность, нарушаемая персонажем пьесы или романа, который действует вопреки своему характеру, поднимает разные проблемы. Во-первых мы должны учесть аргументы психологов, согласно которым черты характера скорее *локальны*, чем *глобальны* (глава X). Хотя многие писатели (Гамсун упоминает Золя) соглашались с фолк-психологией, предполагающей кросс-ситуационное постоянство; хорошие писатели (Гамсун приводит в пример Достоевского) так не поступают. Они могут разочаровать читателей, ожидающих, что персонажи будут вести себя согласно своему образу, но их произведения предназначены для другой аудитории. Как мы вскоре увидим, даже на хороших авторов могут действовать ограничения, вызванные психологией их читателей, но им не следует поддерживать веру в глобальные черты характера. Читатели, тем не менее, вправе ожидать локальной непротиворечивости. Если автор загоняет себя в угол, так что единственный способ развития сюжета по намеченному пути — позволить персонажу совершить локально противоречивое действие, он нарушает негласный договор с читателями. Сюжет должен развиваться подобно тому, как вода стекает вниз, а не форсироваться автором, чтобы течь вверх.

Позвольте мне проиллюстрировать эту идею замечаниями, сделанными Стендалем на полях рукописи его незаконченного романа «Люсьен Левен», опубликованного посмертно. Герой, чьим именем назван роман, влюбляется в молодую вдову госпожу де Шастеле. Она отвечает ему взаимностью, но он не решается принять ее любовь. Утонченность его натуры, дававшая ему превосходство над «самым удачливым Дон Жуаном» и тем самым способная внушить любовь, ставила его ниже любого «не столь хорошо воспитанного парижанина», который сразу сообразил бы, как нужно вести себя в такой ситуации. Чтобы двигать сюжет вперед, Стендаль должен как-то свести героев, но он не знает, как. Он записывает на полях: «После чего хроникер говорит: невозможно ожидать от добродетельной женщины, что она сама отдастся, ее нужно взять.

⁷ Перевод Анны Радловой. — *Примеч. ред.*

Лучшая охотничья собака может только принести подстреленную дичь. Если охотник не стреляет, собака бесполезна. Романист для его героя, как эта собака». Этот комментарий прекрасно иллюстрирует потребность в том, чтобы поведение персонажей в романе оставалось в рамках образа.

Стендалю все-таки удается создать ситуацию, в которой любовь Люсьена и госпоже де Шастеле может быть показана, а не декларирована. Но на этом его проблемы не заканчиваются. Стендалевский план романа следовал диалектическому голливудскому рецепту: мальчик встречает девочку, мальчик и девочка расходятся, мальчик и девочка сходятся снова. Как мы только что видели, у него возникли проблемы с воплощением этого тезиса. Чтобы предложить антитезис, Стендаль использует смешной и явно телеологический прием, заставив Люсьена поверить, что мадам де Шастеле, которую он каждый день видел вблизи, произвела на свет ребенка. Но по-настоящему он споткнулся на синтезе. Хотя мы не знаем, почему он так и не написал третью часть романа, в которой герои должны были снова сойтись, есть предположение, что их союз был бы неправдоподобен. Во второй части романа после разрыва Люсьен становится в некоторой степени циником, человеком порядочным с точки зрения не слишком строгих норм Июльской монархии, но все-таки очень далеким от того застенчивого и утонченного молодого человека, которого полюбила мадам де Шастеле. Стендаль мог решить, что если он заставит ее полюбить преобразившегося Люсьена, он нарушит относительную интеллигентность.

Аристотель писал: «Сказание [*mûthos*] [...] должно быть [подражанием действию] единому и целому, а части событий должны быть так сложены, чтобы с перестановкой или изъятием одной из частей менялось бы и расстраивалось целое, ибо то, присутствие или отсутствие чего незаметно, не есть часть целого»⁸. Мы можем прочесть этот пассаж как выражение эстетических идеалов *полноты* и *экономии*. Читатель имеет право считать, что автор дал ему всю информацию, необходимую, чтобы понимать развитие сюжета⁹. И наоборот, читатель имеет право предполагать, что если автор сообщает о дожде, который шел в момент выхода

⁸ Аристотель. Сочинения: в 4 т. М.: Мысль, 1983. Т. 4. С. 655.

⁹ Разумеется, некие детали, потенциально относящиеся к делу, могут быть сознательно пропущены, чтобы дать место воображению читателя. Рациональное творчество может даже потребовать, чтобы некоторые пробелы читатель заполнил сам. Однако если автор будет переоценивать воображение своей аудитории, он обречен на провал. Предположим, романист пытается выразить несовместимость темпераментов героя и героини, присвоив домам, в которых они живут, номера, соответствующие простым числам, не имеющим общего делителя. За исключением особых случаев, он не может рассчитывать, что читатель сможет уловить данный факт.

героя из дому, то этот факт понадобится ему позднее; также он вправе рассчитывать, что произносимые героем речи должны что-нибудь говорить о самом персонаже или служить предпосылкой к действию других¹⁰.

Ранее я упоминал о том, что любой хороший сюжет обладает текучестью, используя в качестве примера поведение, вписывающееся в образ. Говоря более обобщенно, в хороших сюжетах не должно быть невероятных событий, случайностей и совпадений. В «Миддлмарч»¹¹ встреча между Раффлзом и господином Балстродам (ключевой момент в развитии сюжета) настолько притянута за уши, что отвлекает от в остальном идеально-плавного течения романа. Разумеется, в книге должно быть место случайностям. Случайная смерть родственника может послужить отправной точкой или повлиять на развитие сюжета, равно как и гибель обоих родителей в одном несчастном случае. Но если сюжет требует, чтобы они погибли в разных несчастных случаях, достоверность снижается. Пришедшаяся кстати смерть супруга, позволяющая герою или героине вступить в брак с любимым человеком, тоже является признаком предосудительной авторской лени.

Но психология читателей не так легко согласуется с теорией вероятности. Предположим, перед автором стоит выбор — добраться из сюжетного пункта А в сюжетный пункт В за два шага или за шесть. Чтобы сузить задачу, предположим, что два шага потребуют введения событий, вероятность которых соответственно 0,9 и 0,2, тогда как каждое из шести событий может произойти с вероятностью 0,75. Если допустить, что события в каждой последовательности не зависят друг от друга, последовательность из двух шагов обладает большей вероятностью (0,18 против 0,178), но при этом последовательность из шести шагов будет обладать ценным качеством плавного течения. Общая убедительность сценария гораздо больше зависит от достоверности самых слабых звеньев, чем от их количества. Я полагаю, что автор должен уважать капризы читателей, поскольку это мешает ему воспользоваться легкими, но маловероятными совпадениями.

Однако даже у текущего вниз водного потока могут быть свои повороты и изгибы, прежде чем он спокойно доберется до моря. Если он лишен этого, наблюдение за его течением мало что дает. Следовательно, автор обязан заготовить необходимые сюрпризы для своих читате-

¹⁰ Конечно же, излишней информации не всегда нужно избегать, поскольку у нее может быть эстетическая функция. Для того, чтобы передать скуку, многословие может быть более эффективным средством, чем просто авторская декларация. Но даже в этом случае может случиться, что повторение наскутит *читателю*, а не будет отражением состояния героя.

¹¹ Роман Джордж Элиот, вышедший частями в 1871–1872 годах. — *Примеч. ред.*

лей и зрителей, а также препятствия для персонажей, чтобы поддерживать интерес публики. Набор таких стратегий слишком широк, чтобы его можно было анализировать или даже классифицировать. Некоторые из них тесно связаны с жанром. В театре (комедии, драме или трагедии) в распоряжении автора имеются разные средства. Если комедия, чтобы создать напряжение, часто полагается на *недоразумения*, драма, а еще больше трагедия могут полагаться на *неведение*. Как только недоразумение проясняется, наступает счастье. Когда же снимается покров неведения, происходит катастрофа. Романисты могут присоединять свой голос к голосам героев для усиления неопределенности, если только ониумышленно не вводят читателей в заблуждение.

Теперь я могу определить, что я подразумеваю под правильной интерпретацией текста. Как я сказал в самом начале, это вопрос объяснения. Поскольку все объяснения каузальны (включая те, которые в качестве причин называют намерения) и причина должна предшествовать следствию, *непосредственное* восприятие произведения читательской аудиторией отношения к делу не имеет. В свою очередь сознательные восприятия могут быть частью объяснения. Авторский замысел — неединственное, что имеет значение среди причин, предшествующих произведению. На него могут оказать влияние бессознательные авторские установки. Так, на Жюль Верна, автора «Таинственного острова», могли повлиять как его антирасистский замысел, так и его расистские предубеждения. Однако для краткости ограничусь только сознательными намерениями.

Итак, работа по интерпретации литературного произведения утверждает соотносимость важных его особенностей с решениями, принятыми автором с целью увеличения эстетической ценности опыта, который некоторая специфическая читательская аудитория должна вынести из произведения. Отстаивая этот тезис, литературоведы должны действовать так же, как ученые из других областей. Они могут обращаться к черновикам, если таковые существуют, и к высказываниям автора по поводу своего творчества, например к стэндалевским заметкам на полях. Они могут обращаться к другим произведениям того же автора, чтобы посмотреть, возникает ли там такая же модель авторских решений. Они могут ссылаться на современные автору произведения, чтобы выявить конвенции, которые позволяют отличить рамки, в которых делается выбор, от самого выбора. Они могут обращаться и к современным источникам, чтобы выявить читательские предпочтения, которые могли повлиять на автора.

В данном случае их подход ничем не отличается от метода других историков. Как и представители исторической науки, литературоведы и критики сталкиваются с проблемой конечности объема информации, поскольку эксперименты над прошлым невозможны. Как и другие историки, они могут минимизировать соблазн добычи данных, ограничив-

шись треугольником старых источников, поиском новых и выяснением того, какие новые следствия можно извлечь при проверке интерпретаций фактами. Они могут отличаться от других историков тем, что их интерпретация часто (хотя и не всегда) идет рука об руку с *оценочным суждением*. Добился ли автор успеха в создании локального максимума эстетической ценности? Разумеется, есть авторы, которые не ставят перед собой такой цели. Их главной заботой могут быть только деньги или пропаганда — цели, к которым предъявляются иные требования с точки зрения рациональности. Но если достоверна гипотеза, что у автора были преимущественно эстетические претензии, имеет смысл спросить, как в случае с любой другой целью, насколько успешно он их оправдал.

Я отмечал, что авторские неудачи могут быть интеллигибельны. Я утверждал, что на авторов оказывается двойное давление: они должны продвигать сюжет вперед и пользоваться для этого интеллигибельными высказываниями и поступками персонажей. Мы можем их упрекнуть, если они пожертвовали последней целью ради первой, то есть пожертвовали каузальностью ради телеологии, но мы все же можем *понять*, зачем они это сделали. Даже будучи недостоверным с точки зрения каузальности, тот факт, что Гамлет тянет время, может показаться телеологически интеллигибельным, учитывая потребность Шекспира отложить его месть до конца пьесы.

Это тоже пример интерпретации. Хотя очевидно, что такая интерпретация отличается от интерпретации задержки с точки зрения психологии и обстоятельств Гамлета, она все-таки отвечает на вопрос: почему происходит задержка? Хотя в хорошем литературном произведении все должно быть объяснено дважды, у несовершенных произведений может быть только одна интерпретация. Позвольте закончить примером того, как интерпретация может нарушать или игнорировать объяснение. Некоторые современные исследователи утверждают, что Фанни Прайс в «Мэнсфилд-парке»¹² плетет интриги и что ее напускная скромность — только способ заполучить Эдмунда Бентрама. Более того, они подчеркивают, что само ее имя означает «секс за деньги». Эти утверждения *не проходят две проверки интенциональностью*. Во-первых, в романе нет свидетельств того, что Фанни Прайс плела интриги. Хотя ее скромность действительно вознаграждена, это *следствие* ее поведения не может объяснить само поведение¹³. Во-вторых, нет свидетельств того, что, по замыслу Джейн Остин, читатели должны считать Фанни Прайс наполовину проституткой. Хотя

¹² Роман Джейн Остин (1814).

¹³ Кроме того, гипотеза о корыстности Фанни Прайс не объясняет факт, что она ответила отказом на предложение гораздо более высокородного Генри Крауфорда.

текст может вызвать подобные ассоциации у части современной аудитории, указанные исследователи не приводят доказательств намерений Остин натолкнуть читателей на параллель между Фанни и героиней порнографического романа Фанни Хилл или фамилией «Прайс» (то есть «цена») и платой за секс. Такие интерпретации с точки зрения последствий имеют много общего с функциональными объяснениями в социальных науках. Они полагаются на произвольные методы, ограничивающиеся не фактами, но только пределами изобретательности исследователей. В четвертой части мы рассмотрим некоторые более respectable разновидности объяснения с точки зрения последствий.

Библиографические примечания

Общий подход, которым я пользуюсь в данной главе, часто обвиняют в интенциональности. Я согласен с ответами на эту критику, данными Н. Кэрролом, в частности, в его работах *«Искусство, намерение и разговор»* (Carroll N. Art, intention and conversation // Iseminger G. (ed.). *Intention and Interpretation*. Philadelphia: Temple University Press, 1992) и *«Заблуждение интенциональности: в защиту моих взглядов»* (The intentional fallacy: Defending myself // *Journal of Aesthetics and Art Criticism*. 1997. No. 55. P. 305–309). В *«Герменевтике и гипотетико-дедуктивном методе»* (Hermeneutics and the hypothetico-deductive method // Mart M., McIntyre L. (eds). *Readings in the Philosophy of the Social Sciences*. Cambridge, MA: MIT Press, 1994) Д. Фёллесдал (Føllesdal D.) предложена похожая интерпретация «Пер Гюнта», за исключением того, что эта пьеса, возможно, не подпадает под действие условия, согласно которому все события и герои могли быть реальными. Замечанием о том, что промедление Гамлета может быть вызвано озабоченностью Шекспира построением пьесы, я обязан статье Э. Вагенкнехта *«Идеальная месть — промедление Гамлета. Новое прочтение»* (Wagenknecht E. The perfect revenge — Hamlet's delay: A reconsideration // *College English*. 1949. No. 10. P. 188–195). Идею произведений искусства как локальных максимумов я обсуждаю в главе 3 *«Освобожденного Улисса»* (Ulysses Unbound. Cambridge University Press, 2000). Эта глава также включает более подробное обсуждение «Люсьена Левена». Источником сравнения сюжетов как потоков, текущих вниз или вверх, стала работа Канемана и Тверского *«Эвристика симуляции»* (Канеман Д., Тверски А. *Эвристика симуляции* // Канеман Д., Словик П., Тверски А. *Принятие решений в неопределенности. Правила и предубеждения*. М.: Генезис, 2005). Толкования «Мэнсфилд-парка», которые я критикую, можно найти в статье Дж. Хейдт-Стивенсон *«Доходит до смешного. Непристойный юмор и политика тела в романах Джейн Остин»* (Heydt-Stevenson J. «Slipping into the ha-ha»: Bawdy humor and body politics in Jane Austen's novels // *Nineteenth-Century Literature*. 2000. No. 55. P. 309–339) и в книге Дж. Дэвидсона *«Лицемерие и политики вежливости»* (Davidson J. *Hypocrisy and the Politics of Politeness*. Cambridge University Press, 2004).

УРОКИ ЕСТЕСТВЕННЫХ НАУК

4

Ч А С Т Ь Ч Е Т В Е Р Т А Я

Разные научные дисциплины могут находиться друг с другом в одном из двух видов отношений — *редукции* или *аналогии*. Редукция принимает форму объяснения феноменов одного уровня в иерархии наук в терминах феноменов более низкого уровня (см. рисунок).

Редукционистская программа часто бывает противоречивой. В течение долгого времени многие биологи со страстью утверждали, что редукция биологии к химии не работает, но в действительности это не так. Со времен Эмиля Дюркгейма многие оспаривали редукцию социальных наук к психологии. Центральная идея данной работы состоит в том, что она возможна.

Между двумя этими редукциями есть редукция психологии к биологии. Релевантные биологические дисциплины в данном случае — генетика, физиология, биология развития и эволюционная биология. Две первые дисциплины изучают непосредственные причины изменений в строении и поведении организмов, две последние — косвенные причины, влиявшие на историю развития индивидуальных организмов или видов. Изучение структуры и изучение поведения связаны друг с другом, поскольку структура создает возможности для поведения и накладывает на него ограничения. Тот факт, что у нас две почки, а нужна только одна, дает нам возможность стать донором почки для своего потомства и делает возможными социальные нормы, санкционирующие (или запрещающие) возникающую в этой связи практику. Но две почки у нас не для того, чтобы обеспечить возможность трансплантации от живого

Социология, экономика, политология, антропология

Психология



Биология



Химия



Физика

донора. Многие структуры существуют только ради того, чтобы мы могли что-то делать, но это не тот случай. Как бы то ни было, зачастую трудно сказать, являются следствия какой-то структуры, создающие возможности, случайными или объясняющими.

Значимость биологии для социальных наук должна быть очевидной, так как области их изучения пересекаются. Но все же многие ученые противятся биологическим объяснениям, потому что считают их редукционистскими. Это странное обвинение для тех, кто, как я, полагает, что редукционизм — двигатель прогресса в науке. Но если редукционизм сопровождается такими эпитетами, как *поспешный*, *грубый* и *спекулятивный*, это возражение может иметь основания.

Поспешный редукционизм наблюдается, когда ученые, убежденные в необходимости перехода от объяснения более высокого уровня к более низкому уровню, пытаются делать это, когда у них еще отсутствуют необходимые средства измерения, концепции и теории. Классический пример — механистическая физиология Декарта, которую Паскаль комментировал следующим образом: «Декарт. Мы должны кратко выразиться так: Это образовано фигурой и движением, ибо это верно. Но сказать это и создать машину — это смешно. Ибо бесполезно, неопределенно и болезненно». Сегодня в похожей ситуации находятся те, кто предлагает алгоритмы для распознавания образов и автоматического перевода. Задачи по узнаванию человеческих лиц и распознаванию имеющих смысл предложений, которые мы выполняем безо всяких усилий, до сих пор недоступны искусственному интеллекту.

Грубый редукционизм наблюдается, когда ученые пытаются объяснить специфическое поведение в биологических категориях, вместо того чтобы объяснять *способность* или *тенденцию* к такому поведению, которая в конкретном случае может реализоваться или не реализоваться. Примером могут служить попытки объяснить политическое поведение в категориях территориального императива, встречающегося у низших животных. Еще один пример — идея, что тяжелая атлетика может быть объяснена как результат сексуального отбора, то есть по аналогии с павлиньими перьями или большими рогами у оленей. Можно привести другие примеры.

Спекулятивный редукционизм наблюдается, когда ученые производят истории в духе «просто так», которые дают объяснение того, как могло возникнуть данное поведение, не показывая, что оно действительно возникло именно таким образом. Подобными примерами изобилует социобиология и тесно с нею связанная эволюционная психология. Например,

ученые этих направлений утверждают, что самообман развился в силу эволюционных преимуществ или что послеродовая депрессия у женщин возникла как инструмент торговли (глава XVII).

Сказать, что плохой редукционизм плох, — значит ничего не объяснить. Желание редуцировать сложные явления к более простым может принимать форму упрощения, но то же самое может произойти с любой стратегией в науке. Подавляющее большинство примеров из истории науки показывает, что редукцию низм прогрессивен, а антиредукционизм является сдерживающей силой. История также показывает, что использование *аналогий* между разными научными дисциплинами для производства гипотез может оказаться рискованным делом. Само по себе использование аналогий безобидно: о научных гипотезах судят по их потомкам (экспериментально верифицируемым результатам), а не по их предкам. Но когда мышление аналогиями приводит к тому, что ученые отдают предпочтение одним гипотезам в ущерб другим, результат обычно попадает в кунсткамеру научной мысли. Например, аналогия между обществом и биологическими организмами использовалась для поддержания идеи о том, что общества, подобно организмам, являются саморегулирующимися структурами со встроенными механизмами гомеостатической коррекции (например, революции). В XIX веке ученые обсуждали, какой элемент общества соответствует *клетке* организма, не задаваясь вопросом о том, есть ли какие-либо основания для такой аналогии. Другие исследователи использовали физические (а не органические) аналогии и искали социальные эквиваленты законов Ньютона или силы тяготения. Ученые, утверждающие, что социальные науки могут воздействовать на свой объект, привычно цитируют принцип неопределенности Гейзенберга, как будто глубина этого принципа может превратить их банальность в нечто, столь же глубокое.

В главе XV я рассматриваю некоторые результаты физиологии и науки о мозге, которые обещают дать редукционистское объяснение таких явлений, как страх, доверие и поспешность с выводами. В главах XVI и XVII я обращаюсь к теории естественного отбора, которая использовалась как для редукции, так и для аналогии. Я попытаюсь доказать, что хотя некоторые попытки редукционизма состоятельны, другие — поспешны, грубы или спекулятивны. Использование естественного отбора в качестве аналогии для социальных явлений имеет более чем сомнительную ценность. Одна из причин того, почему в главе XVI я описываю естественный отбор настолько подробно, что это может показаться излишним для книги о социальных науках (при этом слишком поверхностно для полного раскрытия предмета), заключается в стрем-

лении показать, что в социальном мире нет ничего, хотя бы отдаленно сопоставимого с этим феноменом. Полным энтузиазма разговорам социологов о социальном отборе и социальной эволюции не хватает точности и оптической резкости, чтобы их можно было воспринимать всерьез.

Библиографические примечания

По поводу критики спекулятивной социобиологии и эволюционной психологии см. работу Ф. Китчера «*Перепрыгивая амбиции*» (Kitcher P. *Vaulting Ambition*. Cambridge, MA: MIT Press, 1987), а также Д. Баллера «*Приспосабливающийся мозг*» (Buller D. *Adapting Minds*. Cambridge, MA: MIT Press, 2005). О неоправданной поспешности нейронаук в объяснении поведения и об открытой защите теории «черного ящика» см.: Дж. Р. Стэддон «*Адаптивная динамика: теоретическое объяснение поведения*» (Staddon J. R. *Adaptive Dynamics: The Theoretical Explanation of Behavior*. Cambridge, MA: MIT Press, 2001). Исследование роли биологических метафор в изучении общества можно найти в книге Ж. Шлангер «*Метафоры организма*» (Schlanger J. *Les métaphores de l'organisme*. Paris: Vrin, 1971).

XV. Физиология и нейронаука

Поиски физиологических (часто нейрофизиологических) оснований сложного человеческого поведения в последние годы достигли нового пика, главным образом благодаря появлению новых технологий измерения и наблюдения. Без сомнения, у этой линии исследований большое будущее, даже если ее применение иногда может оказаться поспешным, грубым или спекулятивным. Я укажу на три вида открытий, которые, как представляется, имеют непосредственное отношение к целям этой книги.

СТРАХ

В главе VIII я утверждал, что эмоции, как правило, порождаются убеждениями или когнитивными antecedentes. Например, я рассержусь, если сочту, что вы столкнулись со мной намеренно, из-за неосторожности или невнимательности, но не буду сердиться, если посчитаю, что это произошло случайно в результате толчка третьего лица или резкого движения поезда. В «Пленнице» Пруста чувство ревности у рассказчика растет, но утихает вместе с его предположениями о том, чем могла заниматься Альбертина в его отсутствие. Однако могут спросить: не вызываются ли подчас эмоции *ощущениями*, которые, в отличие от убеждений, не имеют причин? Декарт полагал, что удивление или изумление «могут возникнуть еще до того, как мы узнаем, подходит нам эта вещь или нет».

Нейрофизиологические исследования страха (у крыс) это подтверждают. Есть два разных маршрута, ведущих от сенсорного аппарата в таламусе к миндалевидному телу (часть мозга, вызывающая как инстинктивные, так и эмоциональные реакции). Подтверждая традиционное представление о том, что интеллектуальная деятельность всегда предшествует эмоциям и вызывает их, один путь лежит от таламуса к неокортексу (мыслящей части мозга) и от неокортекса к миндалевидному телу. Организм получает сигнал, формирует представление о том, что он означает, и реагирует эмоционально. Однако есть и прямой путь от таламуса к миндалевидному телу, который полностью обходит мыслящую часть мозга. По сравнению с первым, второй путь «простой и грязный». С одной стороны, он быстрее: у крысы акустический стимул достигает миндалевидного тела через таламус всего за 12 миллисекунд,

через кору головного мозга он идет в два раза дольше. С другой стороны, второй путь не так хорошо дифференцирует поступающие сигналы. Если кора головного мозга может понять, что тонкий изогнутый предмет на лесной тропинке скорее палка, чем змея, миндалевидное тело не может их различить. Хотя, с точки зрения выживания, лучше принять палку за змею, чем наоборот.

Неизвестно, можно ли экстраполировать эти открытия в исследовании страха на другие эмоции. Гипотетически это может быть отчасти верно для гнева. При столкновении с чем-то, что может быть истолковано как нападение, альтернативные издержки выяснения, действительно ли это так, могут оказаться слишком высокими. Возможно, естественный отбор заложил реакцию «сначала стрелять, потом задавать вопросы». Если я поспешу с выводами, а потом выяснится, что я не был жертвой нападения, я, тем не менее, смогу придумать историю, оправдывающую мое поведение. Этот тонкий механизм, связанный с нашей потребностью в самоуважении и резюмируемый половицей «Обидевший простить не может», по-видимому, взаимодействует с нейрофизиологическим механизмом, который есть и у нас, и у лишенных самоуважения животных. Это похоже на некий паттерн, воспроизводящийся во многих открытиях физиологии и нейронауки. Почти автоматические реакции, которые объединяют нас с другими видами, становятся предметом тенденциозных толкований и построений, доказывающих исключительную принадлежность их человеку. Подобные рационализации нетривиальны, поскольку могут заставить нас упорствовать в агрессивном поведении, вместо того чтобы признать свою вину.

Еще одно открытие, сделанное в рамках той же исследовательской программы, предлагает альтернативу теории Фрейда о вытеснении воспоминаний. Часто у жертв изнасилования или у прошедших войну солдат в сознательной памяти не остается пережитого опыта. Является не-вспоминание (необходимость этого нейтрального термина будет прояснена ниже) мотивированным процессом, как думал Фрейд, или же мотивация не играет в нем роли? Для поиска ответа можно обратиться к связи между стрессом и формированием (или отсутствием) воспоминания. Небольшой стресс обостряет память о стрессовом событии, но сильный и продолжительный стресс может так поднять уровень стероидов в надпочечниках, что на гиппокамп (часть мозга, в которой образуются сознательные воспоминания) будет оказываться обратное воздействие. Воспоминания о травматическом событии будут не просто вытесняться, они даже не возникнут. Из этого вывода, если он верен, не следует, что травматические события не оставляют следов в психике. Предположим, что вы попали в аварию и сработал

автомобильный гудок. Позднее автомобильный гудок станет стимулом, сразу поступающим в миндалевидное тело и вызывающим телесную реакцию, обычно возникающую в ситуации опасности. Снова получается, что панику вызывает восприятие, а не убеждение.

ДОВЕРИЕ

Предвосхищая обсуждение доверия в главе XXI, я хотел бы указать на три экспериментальных открытия, заставляющих о многом задуматься. Все они основываются на игре «Доверительное управление», в которой участвуют два игрока — «инвестор» и «доверительный собственник». Первый получает определенный вклад, который он полностью или частично может перевести второму. Затем устроители эксперимента увеличивают размер перевода, так что сумма, которую получает «доверительный собственник», в несколько раз превышает ту, что была переведена первоначально. Теперь «собственник» должно решить, делать ли обратный перевод «инвестору» и каков будет его размер. В некоторых экспериментах «инвесторы» получали право наказывать партнера, если считали, что обратный перевод недостаточно щедр. Если оба игрока рациональны, эгоистичны и знают об этом, при одноразовых анонимных интеракциях не будет проводиться никаких переводов, однако все эксперименты обычно показывают положительный уровень прямых и обратных переводов. Еще один вариант этой игры будет разбираться в главе XXI. Здесь я остановлюсь только на некоторых открытиях, связывающих поведение «инвесторов» с гормональным уровнем и активацией центров удовольствия в мозге.

В первом эксперименте изучался размер инвестиций в зависимости от наличия или отсутствия гормона окситоцина. Известно, что этот гормон стимулирует у грызунов просоциальное поведение и способствует образованию грудного молока у женщин, но то, что он стимулирует просоциальное поведение или доверие у людей, стало неожиданностью. После получения гормона процент «инвесторов», которые целиком перевели свои вклады доверенным лицам, вырос с 21 до 45%. Еще более интригующими являются три следующих результата. Во-первых, «доверительные собственники», получившие гормон, не увеличили размер обратных переводов. Во-вторых, у «инвесторов», которым давали гормон, не изменились представления о надежности партнеров (то есть не изменились ожидания касательно обратных трансферов) по сравнению с теми, кому гормон не давали. В-третьих, когда «инвесторы» знали, что обратные трансферы генерируются случайным механизмом с таким же распределением прибылей, как в случае игры против реального человека, окситоцин не оказывал никакого воздействия

на уровень трансферов¹. Этому факту можно дать естественное толкование: влияя на поведение, гормон снижает в «инвесторах» неприятие предательства, но не неприятие риска. Важность неприятия предательства подтверждается и другими экспериментами, не основанными на физиологических манипуляциях.

Второй эксперимент, позволявший «инвесторам» наказывать неблагодарных «доверительных собственников», должен был позволить установить, что происходит у них в мозге, когда они осуществляют наказание. В этой игре у «инвесторов» есть выбор: перевести все 10 денежных единиц вклада или не перевести ничего. Если «инвестор» делает перевод, он увеличивается в 4 раза, то есть партнер получает в итоге 50 денежных единиц — первоначальный вклад в 10 единиц, плюс 40 единиц, полученных от инвестиции. Затем «доверительный собственник» должен выбрать, переведет он 25 из 50 единиц обратно «инвестору» или не переведет ничего. Иными словами, есть три возможных исхода: (10:10), (25:25) и (0:50) (рис. XV.1)

После того как «доверительный собственник» принял решение, оба игрока получали дополнительный вклад 20 единиц. «Инвестор» мог использовать свой вклад, чтобы наказать контрагента при выполнении одного из двух условий. При дорогостоящем условии «инвестор» может дать другому участнику до 20 штрафных очков, за каждое очко «инвестор» будет терять 1 денежную единицу, а «доверительный собственник» — 2 единицы. Таким образом, при максимальном штрафе «инвестор» может сократить выигрыш партнера с 70 (50 + 20) до 30 единиц, тогда как его собственный выигрыш упадет с 20 до 0². При бесплатном условии штраф налагался только на доверенное лицо.

Все 15 «инвесторов», кроме одного, последовательно выбирали выполнение трансфера. Экспериментом манипулировали таким образом, чтобы каждый «инвестор» играл против 7 «доверительных собственников»,

¹ Для определения выигрышей в случае, когда игрокам сообщали, что они играют против случайного механизма, использовались реальные выигрыши инвестора, игравшего против реального доверительного собственника. Если предположить, что ожидания участников игры рациональны, то есть у них правильные убеждения относительно распределения выигрыша, который они могут получить, они будут иметь дело с одинаковым распределением в обеих ситуациях. Такое допущение можно поставить под вопрос.

² В суммах, в действительности выплаченных участникам эксперимента, каждое штрафное очко соответствовало 0,1 швейцарскому франку. Сумма может показаться ничтожной, но другие эксперименты показывают, что в такого рода играх тот же самый качественный результат можно получить, когда ставки соответствуют месячной зарплате. Такие результаты получают при проведении экспериментов в странах третьего мира на гранты, полученные в развитых странах.



РИС. XV.1

из которых 3 сделали обратный трансфер, а 4 все оставили себе. В центре эксперимента были эти эгоистичные «доверительные собственники». После того как один из них заявлял о том, что все оставляет себе, у «инвестора» была всего одна минута, чтобы решить, будет ли он наказывать контрагента и насколько сурово. В этот момент ему проводили сканирование мозга, чтобы выявить соответствующую активность в разных его областях. Одна область — хвостатое ядро (caudate nucleus) — тесно связана с обработкой наград. Другие области — глазнично-дробная кора (orbitofrontal cortex) и предлобная кора — тесно связаны с интеграцией отдельного когнитивного процесса, например с выработкой компромисса между издержками и доходами. Активность в каждом из этих регионов подтвердила гипотезу о мотивации наказания, которую я собираюсь сейчас изложить.

И в дорогостоящих, и в бесплатных условиях есть корреляция между активацией участков мозга, связанных с наградами, и назначенными реальными денежными штрафами. Эта корреляция может означать, что решение о наказании вызывает удовлетворение или что ожидаемое от наказания удовлетворение обуславливает решение его применить. Чтобы разграничить эти две гипотезы, экспериментаторы изучили 11 участников, которые применили максимально возможное наказание в бесплатных условиях. Те из них, у кого были сильнее всего активированы участки мозга, связанные с наградами, применяли самые суровые наказания в дорогостоящих условиях. Поскольку их больше возбуждает наказание, они готовы тратить на него больше, тем самым подтверждая вторую гипотезу. Эта интерпретация также подтверждается фактом, что активация коры головного мозга выше в дорогостоящих условиях, когда участники должны договариваться друг с другом о материальных затратах и психических выгодах, чем в бесплатных.

Эти результаты, как представляется, подтверждают «эффект теплого свечения» этой формы альтруистического поведения (глава V). Хотя наказание косвенным образом оказывается альтруистическим, поскольку дает преимущества третьим лицам за счет материальных издержек агента, у его истоков отнюдь не альтруистические мотивации. Пока рано говорить, насколько устойчивы эти результаты и распространяются они на прямые формы альтруизма (жертвование денег бедным) или на наказание третьих лиц (*А* наказывает *В* за вред, несправедливо причиненный *С*). Справедливое наказание третьих лиц может не принести столь же сладкого удовлетворения, как месть за личную обиду. Тем не менее ясно, что сейчас наука дает средства, помогающие сделать выбор между двумя конкурирующими мотивационными гипотезами, выходящие за рамки чистой интроспекции. Если наказание мотивировано удовольствием, ожидаемым от наказания, такая связь не может быть сознательной. Все, кроме садистов, наказывают других, так как полагают, что те *заслужили наказание*. Можно взять еще более прозрачный пример: было бы концептуально непоследовательно считать, что *единственной* мотивацией пожертвований бедным является «теплое свечение» от таких пожертвований, поскольку сам этот эффект паразитирует на вере человека в то, что он делает доброе дело для других, а не только для себя. Посредством сканирования мозга сегодня мы имеем доступ к состояниям, недоступным для сознания непосредственно.

Назначение наказания отличается от *наблюдения* за заслуженным наказанием. В третьем эксперименте участнику делали сканирование в тот момент, когда он наблюдал, как двоих других (подсадных) участников эксперимента подвергали болезненному электрошоку. Они оба играли с ним в игру «Доверительное управление», при этом только один из них поступал честно. Когда подопытный наблюдал, как наказывают честного игрока, была зафиксирована активация зон мозга, связанных с болью («Я чувствую твою боль»). Наблюдение за наказанием нечестного игрока вызвало меньшую активацию в зонах, связанных с болью, и большую — в областях, связанных с удовольствием (злорадство, *Schadenfreude*). (Важное, но необъяснимое открытие: женщины испытывают большую боль и меньшее удовольствие, чем мужчины, когда наказывают нечестного игрока.) Злорадство — своего рода косвенный гнев: *А* испытывает удовольствие, наблюдая за тем, как *В* наказывает *С*, если *С* причинил вред *А*. Было бы интересно исследовать мозговую активность при косвенном негодовании: получит ли *А* такое же удовольствие от наблюдения за тем, как *В* наказывает *С*, если *С* причинил вред *Д*? Кроме того, можно сравнить косвенный гнев с прямым: насколько удовольствие от наблюдения за тем, как *В* наказывает *С*, когда *С* причинил ущерб *А*, сравнимо с удовольствием от того, когда *А* сам наказывает *С* за причиненный себе ущерб? Эти элементарные эмоции, хорошо известные по литературе, драматургии и поэзии, теперь стали доступными для научных изысканий.

ЗАПОЛНЕНИЕ

Читатели, пользующиеся проверкой орфографии, знают, что программа способна давать подсказки или заполнять то, что может быть пропущено в неправильно написанном слове. Иногда это приводит к появлению неверного слова, которое (если его не проверить) может вызвать комический или катастрофический эффект, но в общем это надежный и полезный инструмент. Мозг тоже обладает способностью заполнения пробелов в опыте. Круглая фигура с небольшими пробелами будет восприниматься как полная и завершенная окружность. Если таких пробелов много, она, возможно, не будет восприниматься как круг в прямом смысле слова, но все равно будет казаться, что ее видимые части лежат на невидимой окружности. В верхней части рис. XV.2 мы видим круги

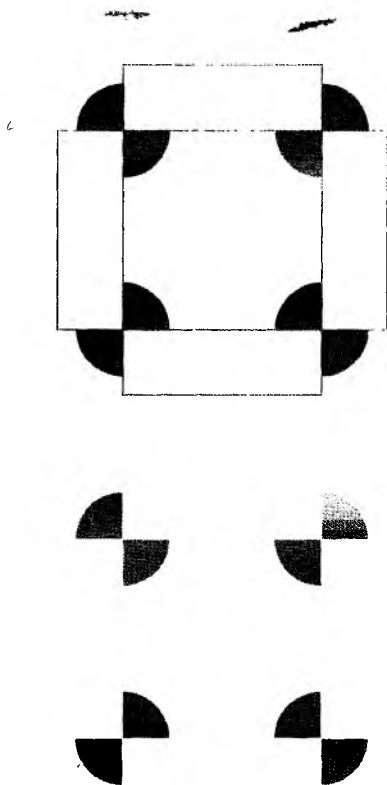


РИС. XV.2

под прямоугольниками, но если прямоугольники убрать, как это сделано в нижней части, круги также исчезнут.

Заполнение — это повседневное явление. На днях я гулял в Леваду-Пере, парижском пригороде, и увидел название улицы, которое я прочитал как «улица Огюста Бланки». Я удивился, почему в состоятельном и предположительно консервативном районе улица названа в честь самого пламенного французского революционера XIX века. Когда же я присмотрелся, увидел, что улица на самом деле названа в честь безупречно респектабельного французского государственного деятеля Аристиды Бриана. Случилось так, что мой мозг заметил первые буквы имени и фамилии, а также их приблизительную длину, а поскольку я гораздо лучше знаком с Бланки, чем с Брианом, заключил, что улица названа в память о нем. Не то чтобы я интерпретировал эти знаки, сознательно дополнив отсутствующие буквы тем или иным способом. Заполнение произошло бессознательно, а в сознательном опыте я *видел* два полных слов. В другой раз я искал булочную («boulangerie») и заметил вывеску магазина с таким названием, но потом выяснилось, что это вторая половина названия «Ресторан л'Оранжери» (Restaurant l'Orangerie). В отличие от случая с Бланки-Брианом, это может быть случаем *мотивированной ошибки восприятия*, а не угадывания мозгом того, что он видит.

Угадывание можно также назвать поспешностью с выводами. К счастью для нас, мозг занимается этим постоянно. Механизмы заполнения, подобно проверке орфографии, полезны; более того, они жизненно необходимы. Однако иногда поиски паттернов превращаются в навязчивые поиски значения. Мозг от природы является приверженцем теории заговоров. Согласно некоторым теориям, правое полушарие выполняет функцию последовательного структурирования информации, которая на нас постоянно обрушивается. Этим можно объяснить некоторые случаи анозогнозии, или отрицания болезни, теми, кто от нее страдает. Это не столько *мотивированный* феномен, как можно предположить из объяснения Фрейда, сколько догадки сидящего в мозге орфографического корректора. Более вероятным представляется то, что это явление может быть вызвано одновременно мотивационными и немотивационными механизмами.

Я подозреваю, что такая активность мозга, как заполнение и наделение значением, наблюдается не только у тех, кого изучают ученые, но и у самих ученых. Я не утверждаю, что между теориями заговора ХАМАС, с одной стороны, и постмодернистской литературной критикой, функциональным объяснением или психоанализом — с другой, нет различия. Я только хочу заметить, что все они вдохновляются потребностью отыскать смысл и систему, выходящие за рамки имеющихся у нас на руках фактов. В случае научных гипотез короткие пути и экстраполяции не могут быть оправданы и объяснены потребностью в немедлен-

ном действии, как это бывает в повседневной жизни. Они отражают скорее глубокое беспокойство, которые мы испытываем, когда оказываемся не в состоянии приписать миру смысл, независимо от того, требуется от нас совершить какое-то действие или нет. Потребность в когнитивной завершенности и неспособность признаться в собственном невежестве (глава VII) также отражают это беспокойство, равно как и гиперрациональность, которую демонстрируют те, кто больше вкладывает в рациональное решение, чем в то, что это решение ставит на карту (глава XII). Еще одним примером является тенденция винить жертву. Предполагая, что мир в основе своей рационален, так что люди получают по заслугам, мы склонны обесценивать и преуменьшать число жертв совершенно случайных событий, даже при отборе на военную службу по жребия. Такая же склонность наблюдается у самих жертв.

Библиографические примечания

Обсуждение страха основано на работе Ж. Леду «Эмоциональный мозг» (LeDoux J. The Emotional Brain. New York: Simon & Schuster, 1996). Воздействие окситоцина на доверие рассматривается в статье М. Костфельда и др. «Окситоцин увеличивает в людях доверие» (Kosfeld M. et al. Oxytocin increases trust in humans // Nature. 2005. No. 435. P. 673–676). Идея неприятия предательства подтверждается в работе А. Боне и Р. Зекхаузера «Доверие, риск и предательство» (Bohnet L., Zeckhauser R. Trust, risk and betrayal // Journal of Economic Behavior and Organization. 2004. No. 55. P. 467–484). Исследование доверия и мести проведено Ж. Ф. де Кервеном и др. в работе «Нервная основа альтруистического наказания» (Quervain J. F. de et al. The neural basis of altruistic punishment // Science. 2004. No. 305. P. 1254–1258). Исследование доверия и злорадства (Schadenfreude) приводится в работе «Эмпатические нервные реакции модулируются восприятием честности других» Т. Сингера и др. (Singer T. et al. Empathic neural responses are modulated by the perceived fairness of others // Nature. 2006. No. 439. P. 466–469). Стимулирующим мысль руководством по явлению заполнения являются «Фантомы в мозге» В. С. Рамачандрана и С. Блейксли (Ramachandran V. S., Blakeslee S. Phantoms in the Brain. New York: Quill, 1998) — книга, в которой объясняется, как на смену первоначальному неприятию Рамачандраном фрейдистского объяснения анозогнозии пришла вера в то, что в ней должен присутствовать элемент отрицания. Кроме того, можно обратиться к работе Н. Леви «Самообман без мысленных экспериментов» (Levy N. Self-deception without thought experiments), которую можно найти в Интернете по адресу au.geocities.com/neil_levy/Documents/articles/Self-deception.pdf. По поводу обвинения жертв см. работу С. Рубина и А. Пепая «Вера в справедливый мир и реакция на жребий другого» (Rubin S., Pepau A. Belief in a just world and reaction to another's lot // Journal of Social Issues. 1973. No. 29. P. 73–93).

XVI. Объяснение с точки зрения последствий и естественного отбора

ПОДКРЕПЛЕНИЕ

Любое объяснение каузально. Мы объясняем событие, указывая на его причину. Причины предшествуют своим следствиям по времени. Отсюда вытекает, что мы не можем объяснить событие (например, действие) через его последствия. Однако если экспланандум является моделью повторяющегося поведения, его следствия могут в какой-то момент войти в число причин, увеличивающих вероятность наступления этого события в следующий раз. Это может произойти двумя способами: путем *подкрепления* и путем *отбора*. Я сосредоточусь на втором способе, более важном для целей этой работы, но начну с того, что скажу несколько слов о первом.

Если то или иное поведение приносит приятные или выгодные следствия, мы чаще ведем себя таким образом. Если его следствия неприятны или влекут наказание, такое поведение встречается реже. Лежащий за этим механизм может быть просто рациональным выбором, если мы замечаем приятные или неприятные последствия и *решаем* действовать в будущем так, чтобы повторить или избежать повторения данного опыта¹. Однако часто подкрепление не является следствием осознанного выбора. Когда младенцы начинают плакать, зная, что таким способом заставляют родителей взять себя на руки, нет оснований считать, что они сознательно отмечают выгоды от плача и позднее делают это по собственной воле, чтобы их получить. Когда дети постарше устраивают истерики, чтобы добиться своего, родители обычно могут понять, что это симуляция.

Подкрепляющее обучение хорошо изучено в лабораторных экспериментах над животными. Обычно животным предлагают нажать на рычаг или на один из нескольких рычагов, после чего им дают вознаграждение в зависимости либо от числа нажатий на рычаг с момента последней на-

¹ Однако не будем забывать, что не всегда легко понять, какой из двух опытов является менее болезненным (головоломка «холодной воды» из главы XII).

грады, либо от времени, которое прошло с момента получения последней награды. В обоих случаях функция может быть детерминистской или вероятностной. При *режиме с постоянным соотношением количества реакций и подкреплений* животное получает награду после того, как нажало на рычаг установленное количество раз, тогда как при *режиме с вариативным числом подкреплений* количество нажатий, необходимых для получения награды, произвольно варьирует. В обоих случаях каждое нажатие приносит поощрительное очко, которое добавляется к предыдущим. При *режиме с постоянным интервалом между подкреплениями* нажатие даст награду через определенное время после получения последней награды, тогда как при *режиме с вариативным интервалом* временной промежуток, по прошествии которого выдается награда, произвольно меняется. В обоих случаях время выдачи наград не зависит от количества нажатий. Каждый режим подкрепления через некоторое время производит специфический стабильный паттерн поведения, который к тому же исчезает, как только положительный стимул (награду) убирают. Так, реакции, усвоенные, когда награда следует за каждым нажатием на рычаг (особый случай режима с постоянным соотношением, известный как постоянное подкрепление), исчезают быстрее, чем те, которые были усвоены при режиме с вариативным соотношением. Интуиция подсказывает противоположную идею, поскольку кажется, что постоянное подкрепление порождает более сильную привычку, но, как это иногда бывает, интуиция заблуждается.

Значение этих открытий за пределами лаборатории зависит от цели.

Если цель в том, чтобы *повлиять* на действие, например, в школьном классе, в казино или на рабочем месте, организатор может (более или менее свободно) установить режим вознаграждений, чтобы сформировать желаемое поведение. Так, режим с вариативным интервалом часто используют для воздействия на поведение, например, когда учитель использует политику произвольных викторин. При режиме с вариативным соотношением, который работает для многих азартных игр, легче повлиять на поведение, если первая награда дается рано². Поскольку менеджеры казино и ипподромов не располагают технологией, которая позволила бы им втягивать новых игроков, предлагая им более высокий выигрыш, им приходится полагаться на так называемое везение новичков³. Действия мошенников, однако, часто полагаются на выдачу быстрых на-

² Это также легче установить, когда игровая технология предусматривает возможность промежуточных выигрышей. Хотя каждый из таких выигрышей обладает меньшей подкрепляющей способностью, чем собственно выигрыш, таких промежуточных выигрышей больше.

³ Их удача в данном случае становится неудачей для них и удачей для казино.

град. В школьном классе и в казино режимы вознаграждений работают «за спиной» у учеников или игроков, то есть они воздействуют на поведение не эксплицитными стимулами, но скорее как в случае с плачущим ребенком, то есть посредством бессознательного процесса. Наоборот, когда менеджеры платят работникам, если те достигли установленной цели (режим с постоянным соотношением) или ежемесячно (режим с постоянным интервалом), они просто устанавливают систему стимулов. Поскольку поведение работников может быть адекватно объяснено ожиданием вознаграждения, нет нужды обращаться к *собственно* вознаграждению.

Если цель в том, чтобы *объяснить* модели поведения их реальными последствиями, режимы вознаграждения имеют смысл, только если возникают естественным образом и, более того, настолько непрозрачны, что не создают открытых стимулов. С двумя постоянными режимами такое, похоже, случается нечасто. На поведение моих друзей влияет не то, сколько я им улыбаюсь, а то, насколько последовательны и уместны мои улыбки. В естественных условиях редко встречаются награды, выдаваемые на регулярной основе (как зарплата). Важнее два вариативных режима. Человек, играющий в «горячо-холодно» (режим с вариативным соотношением) с представителем противоположного (или того же) пола, может вызвать более сильное влечение, чем тот, кто неизменно демонстрирует дружественное поведение. Режим с вариативным интервалом возникает, когда вы хотите до кого-то дозвониться, а линия занята. Вы знаете, что рано или поздно дозвонитесь, но не знаете, когда. В этой ситуации возникает стабильный паттерн повторяемого набора номера, но теория рационального выбора не может дать точного прогноза. Эта теория могла бы предсказать любое число паттернов в зависимости от представлений звонящего о том, как долго может продлиться разговор. Однако представляется маловероятным, чтобы у людей были стабильные убеждения на сей счет.

Паттерн реакций, порождаемый подкреплением, как правило, отличается от паттерна, порождаемого сознательным, рациональным выбором. Предположим, животное должно нажать рычаг, выдающий награды в режиме либо с вариативным соотношением, либо с вариативным интервалом. Рациональный паттерн, максимально увеличивающий общее вознаграждение, состоит в том, чтобы большую часть времени нажимать рычаг с вариативным соотношением, набирая поощрительные очки, и при этом время от времени нажимать на рычаг с вариативным интервалом, чтобы посмотреть, не вышло ли новое вознаграждение. Подкрепляющее обучение, однако, не производит такой паттерн. Вместо этого животные гораздо чаще, чем надо, нажимают на рычаг с вариативным интервалом. Они уравнивают *средние* награды при нажатии одного или другого рычага, вместо того чтобы, как диктует рациональ-

ность, уравнивать *предельные* (marginal) вознаграждения. При других комбинациях режимов подкрепляющее обучение иногда подражает рациональному выбору, но делает это непоследовательно. Если и есть какой-то неинтенциональный механизм, способный надежно симулировать рациональность, его нужно искать в другом месте.

ДИФФЕРЕНЦИРОВАННОЕ ПОЛОВОЕ СООТВЕТСТВИЕ

Чаще всего упоминают механизм естественного, или социального, *отбора*. В третьей части книги я указываю на то, что объяснять поведение можно, используя предположение о более или менее рациональном *приспособлении агентов к своей среде обитания*. В радикально иной перспективе мы можем допустить, что *агенты отбираются средой*. Хотя отбор может быть плодом деятельности интенционального агента, когда, например, выводится порода послушных домашних собак или когда в лаборатории отбирают более умных крыс, многие механизмы отбора покоятся на каузальных процессах, не предполагающих интенционального агента.

В частности, *дифференцированное выживание* организмов, основанное на паттернах их поведения, может привести к оптимальному (для репродукции) поведению в популяции даже в отсутствие любого оптимизирующего отбора или намерения. Предположим, что 10% организмов в популяции из 100 особей так эффективно добывают корм, что оставляют потомство численностью в среднем 10 особей, которое доживает до взрослого состояния, тогда как оставшиеся 90% оставляют только 5 особей. Если поведение родителей (через какой-то механизм) передается потомству, следующее поколение взрослых организмов будет включать группу 100/550 ~ 18%, которая будет демонстрировать более эффективное поведение. Еще несколько поколений — и практически все организмы будут демонстрировать такое поведение. Если мы спросим, *почему* оно повсеместно распространено, ответ будет: его последствия лучше⁴. Этот механизм работает через все поколения. В отличие от подкрепляющего обучения, он модифицирует поведение не отдельной особи, а следующих друг за другом поколений особей.

ЕСТЕСТВЕННЫЙ ОТБОР

Теория естественного отбора подробно описывает эту историю. Здесь я приведу упрощенную, или классическую, версию теории, которая удовлетворяет нашим целям. Я хочу продемонстрировать, чем естественный отбор отличается от намеренного и тем более от рационального

⁴ Лучше чего? Вскоре мы затронем этот вопрос.

выбора. Естественный отбор является оптимизирующим механизмом, но только в слабом смысле. Также я хочу заложить фундамент для разбора в следующей главе того, что социальный отбор вряд ли способен провести даже столь слабую форму оптимизации.

Приспособляемость организма, измеряемая численностью его (способного к репродукции) потомства⁵, определяется одновременно его окружением и физиологическими и поведенческими особенностями, или фенотипом. Его генотип — это набор инструкций, который, действуя согласно со средой, определяет фенотип. Эти инструкции записаны в длинных молекулах ДНК, которые имеют много общего с письменным языком. Буквы ДНК — это четыре молекулы (нуклеотиды), называемые Т, А, Г и Ц. Слова, или кодоны, ДНК — это триплеты нуклеотидов, каждый из которых содержит инструкции для сборки одной из 20 аминокислот, которые являются кирпичиками протеинов в организме. Поскольку триплетов 64 (61, если мы исключим 3 из них, выполняющих другие функции), одна и та же аминокислота может кодироваться более чем одним триплетом. Ген — это сегмент ДНК, кодирующий один отдельный белок.

Организм, который более эффективно добывает пищу, сначала предстал как результат случайной мутации генома. В классической картине эволюции предполагалось, что мутации возникают в результате случайных небольших ошибок в дупликации генетического материала, которые происходят в процессе репродукции.

То, как мутации вносят свой вклад в эволюцию, можно схематически изобразить при помощи аналогии с трехуровневой структурой печатного предложения. Рисунок XVI.1 показывает параллели между двумя этими случаями.

Когда делается типографский набор книги для нового издания, невнимательность наборщика может породить отступления от оригинального текста. Мы можем представить их как замену одной буквы на другую, превращавшую, например, «руку» в «щуку» или «муку». Ошибки являются произвольными в том смысле, что они не связаны с содержанием книги. Сходным образом мутации являются произвольными, поскольку вероятность их возникновения не связана с характером изменения фенотипа, которое они вызывают. Некоторые буквы заменялись чаще других, если у наборщика были проблемы с различением «п» и «л». Второе издание некоторых книг может содержать больше ошибок, неже-

⁵ Приспособляемость в репродуктивном смысле отличается от экологической приспособляемости, измеряемой продолжительностью жизни. Хотя экологическая приспособляемость обычно является средством для репродуктивной, они, тем не менее, могут расходиться (например, когда родители рискуют жизнью, чтобы защитить потомство).



РИС. XVI.1

ли прочие, если, допустим, наборщик был пьян. Точно так же коэффициенты мутаций могут быть подвержены действию мутагенных агентов⁶.

Опечатки в книге или в любом письменном сообщении обычно негативно воздействует на содержание, иногда (в буквальном смысле) фатально. Перенос запятой, например, может быть вопросом жизни и смерти⁷. Подобно этому большинство мутаций имеют вредные последствия для организма, в котором они происходят, снижая тем самым его репродуктивную приспособляемость. Произошедшая случайная мутация, однако, может усилить приспособляемость организма так же, как случайная ошибка в воспроизведении текста первого издания может дать более точную или обновленную информацию. Но дальше эта аналогия уже не работает. Благоприятная мутация будет гораздо сильнее представлена в следующем поколении, поскольку организм, в котором она произошла, будет иметь больше потомства, которое унаследует эти мутации. В противоположность этому не существует механизма, благодаря которому книга со случайно исправленными ошибками первого издания продавалась бы лучше.

⁶ Мутации может вызвать, например, радиация; вот почему рентген для определения рака сам может вызвать рак, спровоцировав мутации.

⁷ Поскольку я не нашел эквивалента в английском, я проиллюстрирую это утверждение двумя норвежскими предложениями: «Vent, ikke heng ham» и «Vent ikke, heng ham», означающими соответственно «Подождите, не вешайте его» и «Не ждите, вешайте его».

В типографских ошибках любые буквы могут быть заменены любыми другими. Новое слово может и не иметь смысла: «рука» может с одинаковым успехом быть заменена на «муку» или на «щуку». При мутации любой из нуклеотидов в триплете также может быть заменен на любой другой. В свою очередь не любая аминокислота может возникнуть из любой другой посредством одной-единственной мутации нуклеотида. Некоторые изменения протеинов могут произойти таким образом, другие — нет. Однако всегда возможно перейти от одной аминокислоты к другой через несколько последовательных мутаций. Этот процесс до некоторой степени напоминает игру, в которой данное слово при помощи однобуквенных замен трансформируется в другое при условии, что все промежуточные замены тоже являются словами, которые можно найти в словаре. От HAIR к HAIL можно перейти за один шаг, а от HAIR к DEAN — за четыре (HAIR — HEIR — HEAR — DEAR — DEAN). Но я совершенно уверен в том, что никакое количество шагов не способно привести вас от HAIR к LYNX, если только не выйти за пределы словаря.

В соответствующей «игре мутаций» все 64 кодона есть «в словаре», за исключением трех, которые служат знаками «абзаца» («терминирующие», завершающие протеины). Хотя три «кодона-точки» исключают некоторые одношаговые изменения триплетов, это ограничение слишком слабо, чтобы блокировать многшаговые изменения. Таким образом, можно перейти от аминокислоты метионин к триптофану за два шага, или от АТГ к ТТГ (леуцин) к ТТГ или от АТГ к АГГ (глицин) к ТТГ⁸. Кроме того, есть маршруты, включающие более одного шага, такие как АТГ — АЦГ — ТЦГ — ТТГ.

Представим теперь более сложную игру и предположим, что каждое слово включено в предложение и что слово, взятое для замены, не только должно быть в словаре, но и давать предложение, обладающее смыслом⁹. Теперь есть два ограничения для допустимых буквенных замен. Предположим, изначальное предложение: «I tend my looks». Заменяв его на «I mend my looks», мы соблюдаем оба ограничения — осмысленности и наличия в словаре. Заменяв его на «I send my tooks», мы нарушим оба условия. Если далее заменить «I send my looks» на «I send my books», оба

⁸ Я остановился на этом примере потому, что эти две аминокислоты — единственные, закодированные в одном триплете. Каждая из остальных кислот имеет два или четыре кодона. Таким образом, фраза «путь от аминокислоты X к аминокислоте Y» должна в этих целях читаться как «путь от данного кодона для X к одному из колонов для Y».

⁹ Такого рода языковую игру практиковала группа французских писателей под названием УЛИПО. Самый известный из них, Жорж Перек, написал роман «Исчезновение», в котором вообще не использовалась буква «е».

условия будут удовлетворены. Чтобы получить это (осмысленное) предложение из изначального (осмысленного) предложения путем замены одной буквы, нам пришлось бы выйти за пределы множества имеющих смысл предложений.

ЕСТЕСТВЕННЫЙ ОТБОР ПОРОЖДАЕТ ЛОКАЛЬНЫЕ МАКСИМУМЫ

Биологическая аналогия ясна. Аминокислота, появившаяся в результате замены нуклеотида, заложена в протеине. Поскольку протеины имеют жизненно важное значение для организма, они должны быть биологически жизнеспособным (аналог наличия смысла у предложений). Если любой промежуточный протеин на пути от начальной к финальной форме нежизнеспособен, он не может закрепиться в популяции, потому что организм, в котором произошла такая мутация, не оставит потомства, которое могло бы перейти на следующий уровень. Даже если финальная форма протеина увеличит приспособляемость по сравнению с первой стадией, этот факт может не создать никакого эволюционного давления на промежуточных стадиях. Если бы все осуществимые пути, от первой до финальной стадии, в какой-то момент потребовали от организма применить обходную стратегию «один шаг назад, два шага вперед», окончательная форма никогда не была бы достигнута. Поскольку большинство — и, по всей видимости, значительное — мутаций являются вредоносными, такая ситуация вполне достоверна. В упрощенной форме (не принимающей во внимание множественность путей) эта структура представлена на рис. XVI.2. Три разных протеина дают разные степени приспособляемости (измеряемой вертикально). Структура генетического кода допускает мутации с одним нуклеотидом от А в В и от В в С, но не от А в С.

Естественный отбор (в классическом виде) ограничивается небольшими улучшениями. Организм взбирается по склону приспособляемо-

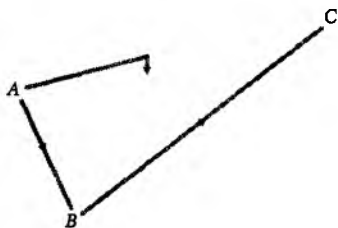


РИС. XVI.2

сти до тех пор, пока не достигнет *локального максимума*, определяемого как состояние, в котором дальнейшие одношаговые изменения только уменьшат приспособляемость. Хотя в ландшафте приспособляемости могут присутствовать пики повыше, там может не быть достижимых одношаговых изменений. Этот процесс отличается от намеренного выбора тремя аспектами. В главе 6 мы отмечали, что благодаря своей интенциональности человеческие существа способны (1) использовать обходные стратегии, (2) ждать и (3) целиться с опережением в движущуюся цель. Мы только что видели, что естественный отбор не способен на (1). Что касается (2), рассмотрим рис. XVI.3.

Структура генетического кода допускает возникновение мутаций с одним нуклеотидом от А в В и от А в С, но не от В в С. Если происходит мутация от А в В, популяция застревает в локальном максимуме (нижнего уровня) В, поскольку одношаговая мутация в глобальный максимум С блокируется. Нет механизма, который благоприятствовал бы мутации в С, а не в В. В отличие от интенциональных агентов, натуральный отбор *оппортунистичен*, хватается за любую возможность улучшения, будучи неспособным *дождаться*, пока представится возможность оптимального изменения.

Что касается пункта (3), то популяции адаптируются к постоянно меняющимся условиям. Если изменения регулярные (например, сезонные или дневные) они приспособляются к изменениям. Если какое-то событие происходит только один раз (как например, внезапная смена климата), поведение, которое до того было в локальном максимуме приспособляемости, может стать субоптимальным, то есть мутации, которые прежде оказались бы вредоносными, становятся благоприятными. Если изменения сохраняются (например, потепление или похолодание климата в течение долгого периода), этот процесс может никогда не достичь локального максимума. Популяция будет отслеживать изменения в окружающей среде с эффективностью, зависящей от относительной скорости

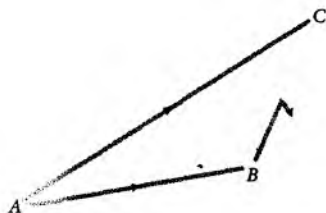


РИС. XVI.3

двух процессов. На удивление тонко отрегулированная приспособляемость, наблюдаемая у животных и растений, указывает на то, что животные адаптируются к внешней среде гораздо быстрее, чем та меняется. И все же организмы всегда будут отставать, потому что не могут *предвосхищать* изменений внешних условий. Наоборот, человеческие существа могут осознавать будущие изменения (например, глобальное потепление) и принимать меры предосторожности, прежде чем те произойдут, или, если такие перемены вызваны поведением человека, препятствовать их возникновению.

Среда обитания среди прочего включает другие виды, с которыми данная популяция может находиться в отношениях «добыча — хищник». Как «добыча» она может развивать стратегии уклонения, как «хищник» — стратегии охоты. Подобно тому как конкретная лиса охотится в полях за конкретным зайцем, так и «лиса» как вид охотится за «зайцем» как видом на протяжении поколений. Но хотя логика естественного отбора не позволяет «лисе» предсказать, где будет «заяц» несколько столетий спустя, некоторые хищники способны перехватывать добычу на пути бегства. Сходным образом процесс локальной максимизации естественного отбора произвел в людях способность к глобальной максимизации.

Мутации являются одновременно условиями и продуктом естественного отбора, поскольку скорость, с которой они происходят, может регулироваться эффектами приспособляемости. Здесь задействовано несколько механизмов. В популяциях бактерий, которые постоянно подвергаются воздействию новых антибиотиков, виды с высокой скоростью мутаций могут оказаться в более благоприятных условиях. Это частный случай общей идеи, что в меняющейся среде немутулирующие популяции вымирают. В то же время слишком высокая скорость мутаций может привести к потере ценной генетической информации и к ее исчезновению. Между двумя этими крайностями есть оптимальная скорость. В постоянном окружении существует компромисс между преимуществами низкой скорости мутаций и издержками «на вычитывание гранков» и восстановительными механизмами, необходимыми для поддержания этой низкой скорости.

Такая классическая картина выглядит слишком простой по ряду параметров, которые следовало бы описать, если бы это был учебник по биологии. Поскольку наша книга им не является, я лишь отмечу, что более сложная картина модифицирует (но не отменяет) утверждение о том, что естественный отбор склонен застревать в ловушке локального максимума. Да, крупные мутации происходят, некоторые из них могут вызывать изменения, которые никогда бы не произошли посредством мелких мутаций. Кроме того, низшие формы не сразу уничтожаются отбором. На рис. XVI.2 мутация в *B* необязательно производит организм, являю-

щийся нежизнеспособным в строгом смысле, то есть неспособным выживать или производить потомство. Некоторые организмы в состоянии В могут выжить и произвести организмы в состоянии С. На рис. XVI.3 некоторые организмы в состоянии А могут уцелеть в борьбе с более эффективными организмами в состоянии В довольно долго, чтобы произошла мутация в состояние С. Возникнет ли при этом глобальный максимум? Это вопрос относительной скорости двух процессов — вымирания более слабых видов и скорости, с которой происходит благоприятная мутация. Однако нет механизма, который мог бы *систематически* имитировать способность интенциональных существ предвосхищать события, ждать или использовать обходные стратегии.

Основываясь на тенденции естественного отбора порождать оптимизирующее поведение или оптимальные структуры через постепенное восхождение, мы не должны делать вывод о том, что все черты, наблюдаемые у организмов, являются оптимальными. С одной стороны, как уже было сказано, может существовать определенный лаг, если популяция не может угнаться за изменениями в окружающей среде. То, что мы наблюдаем, может быть оптимально в какой-то момент времени в прошлом, а потом перестать быть таковым. С другой стороны, феномен *плейотропии* говорит о том, что естественный отбор может поощрять черты, которые сами по себе могут выглядеть субоптимальными, но при этом являются частью оптимального решения. Такое возможно, потому что ген воздействует на фенотип неоднозначно. Если положительные следствия перевешивают отрицательные, то он будет поощряться естественным отбором. В качестве примера можно привести существование мужских черт у насекомых, которые причиняют физический вред своим партнерам при совокуплении. Считалось, что этот вред вызван либо негативным побочным плейотропическим эффектом приспособления, которое дало мужским особям репродуктивное преимущество в ином контексте, либо самим приспособлением мужских особей. Эксперименты показывают, что первое объяснение более правдоподобно.

ЕДИНИЦЫ ОТБОРА

Естественный отбор не только оппортунистичен и близорук, он еще (за двумя исключениями, о которых я вскоре скажу) жестко *индивидуален*. Он благоприятствует не виду или популяции, а индивидуальному организму. Если качество, возникшее в результате мутации, увеличивает относительную приспособляемость организма, оно закрепится в популяции, даже если приведет к снижению абсолютной приспособляемости. Представим себе популяцию рыб, сталкивающуюся с хищниками и плавающую рассеянной стаей. Если мутация заставит рыбу, в которой

она произошла, переместиться в центр стаи, она будет лучше защищена от хищников и, как следствие, оставит больше потомства. По мере того как такое поведение будет распространяться в популяции, стая сделается компактнее, тем самым становясь более легкой добычей для хищников. На любом этапе процесса лучше быть в середине, чем на периферии стаи. Но с точки зрения абсолютной приспособляемости, такой исход для всех хуже, чем начальная ситуация, а с точки зрения относительной приспособляемости ничего не меняется. Точно так же ничем не сдерживаемый сексуальный отбор — правдоподобное объяснение огромных и дисфункциональных рогов, которые встречаются у некоторых видов оленей.

Одним исключением из индивидуализма является *семейный отбор* (форма субиндивидуализма), в котором единицей отбора является скорее ген, чем индивидуальный организм. Выбор единицы не имеет значения, когда действие гена одновременно и в равной пропорции увеличивает его присутствие в популяции и количество потомства, демонстрирующего соответствующее поведение. Так происходит при развитии более эффективных способов добычи пищи. Но в некоторых случаях ген может получить преимущества, даже если организм, в котором он иницирует то или иное поведение, таких преимуществ не получает, а именно когда организм жертвует собой ради близких родственников, у которых предположительно есть такой же ген. Когда животное замечает хищника и подает сигнал тревоги, его шансы на выживание часто снижаются, тогда как шансы его родственников растут. Поскольку эти родственники или хотя бы некоторые из них также наделены этим «геном предупреждения», более высокая вероятность их выживания может привести к распространению данного гена в популяции, если это выживание компенсируют низкие шансы выживания животного, подававшего сигнал об опасности. Это неединственное возможное объяснение распространенной формы поведения, хотя в некоторых случаях оно является наиболее достоверным. В каких-то случаях сигналы тревоги служат для того, чтобы отвлечь хищника или предупредить сородичей (независимо от того, принадлежат ли они к одной семье). В других случаях оно предназначено для того, чтобы отвлечь сородичей, например, чтобы заставить их уйти от скудных ресурсов. Таким образом, в сигналах тревоги нет ничего изначально альтруистического или жертвенного, хотя они и могут в некоторых обстоятельствах имитировать настоящий альтруизм.

Еще одно исключение — *групповой отбор* (форма сверхиндивидуализма). Рассмотрим две популяции рыб, в одной из которых произошла мутация, стимулирующая передвижение ближе к центру стаи, а в другой — нет. Со временем первая популяция будет оставлять меньше по-

томства, чем вторая, и будет постепенно вытеснена. Может показаться, что естественный отбор действует на уровне группы, а не особи. Однако пока две группы сосуществуют, во вторую могут проникать представители первой. Независимо от того, вызвано поведение, направленное на перемещение к центру, мутацией или просачиванием, результат один и тот же — вытеснение тех, кто ведет себя иначе. Точно так же если организмы в популяции имеют ген, который мешает им заниматься выбиванием пастбища, они могут быть вытеснены размножением более раскрепощенных организмов, у которых такой ген отсутствует. По этой причине групповой отбор не рассматривался как правдоподобный механизм порождения кооперации или самоограничения. Однако это возражение может быть опровергнуто в свете теории альтруистического наказания, приведенной в главе XV. Если организмы в популяции наделены геном, который заставляет их наказывать тех, кто не проявляет солидарности, последние не приобретут репродуктивных преимуществ от своего «безбилетничества». Этот механизм может обеспечивать кооперацию там, где можно с уверенностью вычислить несотрудничающих. Маловероятно, что он может применяться к рыбам, стремящимся попасть в центр стаи, но его вполне можно применить к животным, отказывающимся делиться пищей с сородичами.

Семейный и групповой отбор обеспечивают два механизма, порождающих кооперативное поведение; первый базируется на общем гене, а второй — на альтруистическом наказании. Третий механизм — механизм *взаимного альтруизма*, или «услуга за услугу» в повторяющихся интеракциях, таких как «Ты почешешь спину мне, а я — тебе» (для некоторых животных это надо понимать буквально) или «Я угощу тебя, когда у меня будет излишек, а когда лишняя еда окажется у тебя, ты угостишь меня». Другая сторона медали — наказание (или по крайней мере воздержание от кооперации), если другая сторона не сумела ответить взаимностью. Для работы этого механизма необходимо частое взаимодействие особей, чтобы они могли себя сдерживать на протяжении какого-то времени, помнили о том, что другие делали раньше, и узнавали их при новой встрече.

Этот механизм объясняет кооперацию во многих случаях двухчастного взаимодействия. Вот некоторые примеры и контрпримеры. Когда двум обезьянам-капуцинам дали возможность несколько раз поделиться пищей, чем больше еды отдавала первая обезьяна, тем больше получала от второй. Шимпанзе с большей готовностью делятся пищей с теми, кто ухаживал за ними в течение дня. Когда каждый из двух скворцов получал пищу тогда, и только тогда, когда второй давил на рычаг, они давили на рычаг по очереди. Самки летучих мышей-вампиров отгрыгивают кровавую пищу для тех, кто не смог добыть ее в недавнем прошлом.

Вампиры способны узнавать друг друга и склонны давать кровь тем, кто давал им ее в прошлом, в особенности родственникам. В другом эксперименте каждая из двух голубых соек получала больше, если нажимала на «эгоистический рычаг» А, чем когда нажимала на «рычаг кооперации» В, независимо от того, что делала вторая. И обе получали больше, если обе нажимали на В, когда на него не нажимала ни одна. Здесь модели «услуга за услугу» не возникло: после короткого сотрудничества в самом начале обе сойки стали постоянно нажимать на рычаг А.

Взаимный альтруизм, однако, непригоден в качестве механизма, порождающего кооперацию в более крупных группах. Единственной работающей в данном случае стратегией наказания могла бы быть четкая реакция, в результате которой некооперативное поведение одного из членов группы мгновенно заставляло бы остальных прекращать сотрудничество как с нарушителем, так и друг с другом. Все же интуитивно такая реакция кажется неправдоподобной. Похоже, на практике она не встречается.

Библиографические примечания

В работе «Отбор по последствиям» (Selection by consequences // Science. 1981. No. 213. P. 501–504) Б. Ф. Скиннер (Skinner B. F.) отмечал важность трех способов объяснения поведения через его последствия: естественный отбор, работающий с особями, подкрепление и (хотя он не использует этот термин) групповой отбор. Полезное введение в теорию подкрепления можно найти в книге Дж. Э. Р. Стэддона «Адаптивное поведение и обучение» (Staddon J. E. R. Adaptive Behavior and Learning. Cambridge University Press, 1983). Исследование того, как теория подкрепления может использоваться для воздействия (а не объяснения) на поведение, содержится в статье Д. Ли и Р. Белфлора «Повышение школьной успеваемости: обзор режимов подкрепления» (Lee D., Belflore P. Enhancing classroom performance: A review of reinforcement schedules // Journal of Behavioral Education. 1997. No. 7. P. 205–217). Классическое изложение теории естественного отбора с характерным акцентом на его индивидуалистской природе приводится в книге Дж. Уильямса «Приспособление и естественный отбор» (Williams G. Adaptation and Natural Selection. Princeton, NJ: Princeton University Press, 1966). Обсуждение постепенного восхождения и метафоры ландшафтов приспособляемости можно найти в главе 2.4 С. Гаврильца «Ландшафты приспособляемости и происхождение видов» (Gavrilets S. Fitness Landscapes and the Origin of Species. Princeton, NJ: Princeton University Press, 2004). Взгляд на ген как на единицу отбора представлен в книге Ричарда Докинза «Эгоистичный ген» (Докинз Р. Эгоистичный ген. М.: Мир, 1993). Прекрасное введение в изучение сигналов у животных содержится в книге У. А. Сёрси и С. Новицки «Эволюция коммуникации у животных» (Seag-

су W.A., Nowicki S. *The Evolution of Animal Communication*. Princeton, NJ: Princeton University Press, 2005). Обсуждению того, как альтруистическое наказание делает возможным групповой отбор, посвящена статья Е. Фера и У. Фишбахера «Социальные нормы и человеческая кооперация» (Fehr E., Fischbacher U. Social norms and human cooperation // *Trends in Cognitive Sciences*. 2004. No. 8. P. 185–190). Основополагающее исследование кооперации «услуга за услугу» у животных, не связанных отношениями родства, приводится в статье Р. Аксельрода и У. Гамильтона «Эволюция кооперации» (Axelrod R., Hamilton W. The evolution of cooperation // *Science*. 1981. No. 211. P. 1390–1396).

XVII. Отбор и человеческое поведение

ИЗМЕНЧИВОСТЬ И ОТБОР

В какой мере человеческое поведение выбирается или отбирается механизмами, отличными от интенционального выбора? Чтобы ответить на этот вопрос, для начала отметим, что любой механизм отбора нуждается в сырье, вводных данных, с которыми он будет работать. Как и сам процесс отбора, источник вариаций может быть либо интенциональным, либо неинтенциональным.

| Отбор | Источник вариации | |
|-------------------|--|--|
| | Интенциональный | Неинтенциональный |
| Интенциональный | Искусственный отбор при разведении растений и животных | Постепенное улучшение лодок Евгеника Селективный отбор и инфантицид |
| Неинтенциональный | Отбор компаний в рыночной конкуренции | Естественный отбор (глава XVI) |

НЕИНТЕНЦИОНАЛЬНАЯ ИЗМЕНЧИВОСТЬ, НЕИНТЕНЦИОНАЛЬНЫЙ ОТБОР

Очевидно, что естественный отбор повлиял на физическую структуру человека, дав ему возможности как для действия, так и для их ограничения. Те, кто пытается объяснить человеческое поведение с точки зрения естественного отбора, выдвигают более серьезные требования. Они хотят объяснить сами поведенческие модели, а не только строение, делающее их возможными.

Наиболее вероятный механизм заключается в том, что эволюция производила эмоции вместе с характерными тенденциями к действию. Поскольку самец никогда не может быть совершенно уверен в том, что

является отцом потомства, а у самки не возникает сомнений в ее материнстве, мы можем ожидать, что естественный отбор породит более сильную склонность к сексуальной ревности в мужчинах, чем в женщинах. Это подтверждает статистика убийств. Так, из 1060 убийств супругов, совершенных в Канаде с 1974 по 1983 год, 812 были совершены мужчинами и 248 — женщинами. Среди них 195 убийств из ревности было совершено мужчинами и только 19 — женщинами. Теория естественного отбора также предсказывает, что родители будут испытывать к своим биологическим детям большую эмоциональную привязанность, чем к приемным. Эту догадку тоже подтверждают факты. Так, в 1976 году американский ребенок, живущий в семье с одним неродным родителем, подвергался риску насилия примерно в 100 раз чаще, чем ребенок, у которого оба родителя родные. Естественный отбор может также поощрять *отсутствие* эмоций. Опасности инбридинга (родственного скрещивания) сдерживаются у людей и других видов приматов отсутствием взаимного сексуального влечения у двух росших вместе молодых особей, независимо от того, являются они родственниками или нет¹.

Естественный отбор, работающий с группами, а не с особями, мог благоприятствовать формированию эмоций гнева и негодования по отношению к тем, кто нарушает нормы взаимопомощи, побуждая к наказанию зачастую даже в ущерб интересам «палача» (глава XVI). Более сложным является вопрос: действительно ли отбор поощряет чувство презрения по отношению к нарушителям скорее социальных, нежели моральных норм? Если да, то почему он это делает? Поскольку многие социальные нормы являются произвольными и даже дисфункциональными (глава XX), трудно понять, как этот процесс может поддерживаться групповым отбором. Учитывая тенденцию подвергать остракизму тех, кто попирает социальные нормы, их соблюдение в большей степени способствовало бы репродуктивной приспособляемости, если бы издержки от наказания превосходили выгоды от несоблюдения норм. Загадка в том, почему вообще возникает такая тенденция. Например, почему люди осуждают адюльтер? Социальные нормы, направленные против адюльтера, затрагивают реакции третьих лиц, отличные от ревности, которую испытывает вторая сторона. Хотя А может извлечь выгоду из того, что С осуждает авансы В в отношении жены А, эта выгода не оказывает на С побуждающего селективного давления. Если групповой отбор может поощрять гены, побуждающие к наказанию халаящи-

¹ Таким образом, табу на инцест может быть обращено на соблазн, который возникает гораздо реже, чем было принято считать. Фрейд, наоборот, считал, что табу на инцест возникло, чтобы противодействовать бессознательному желанию иметь сексуальные отношения с близкими родственниками.

ков третьими лицами, то выгоды, получаемые группой от преследования адюльтера третьими лицами, не столь очевидны. Хотя тенденция к более строгому порицанию женского адюльтера, нежели мужского, наводит на мысль об эволюционном объяснении, трудно себе представить, какой за этим может стоять механизм².

Другие утверждения (например, мысль о том, что *самообман* развился у людей вследствие эволюционных преимуществ) носят более спекулятивный характер. Рассуждение строится следующим образом. Обман других часто приносит пользу. Однако обманывать сознательно или лицемерно трудно. Следовательно, самообман развился для того, чтобы люди могли успешнее обманывать других. Слабость этого аргумента в том, что самообман может стать источником ложных заблуждений, которые могут иметь катастрофические последствия, если ими пользоваться как руководством к действию. Никто еще убедительно не доказал, что *результатирующий* эффект этих противоположных друг другу следствий является положительной величиной, как в случае, если бы самообман действительно усиливал эволюционную приспособляемость.

Еще более спекулятивным является утверждение, что *однополярная депрессия* могла развиться как инструмент давления на переговорах (нечто вроде забастовки). Например, предполагаемая функция постродовой депрессии — заставить других принимать участие в воспитании ребенка, подобно тому как рабочие объявляют забастовку, чтобы заставить работодателей поделить прибыль. Вызванные такой депрессией самоубийства, по этой логике, являются ценой, которую приходится платить, чтобы угроза самоубийства выглядела правдоподобной. То есть это попытки самоубийства, которые не сумели провалиться. Бессонница истолковывается как выделение когнитивных ресурсов на преодоление кризиса, реакцией на который является депрессия, тогда как патологическая сонливость определяется как способ сокращения производительности и повышения таким способом эффективности депрессии как инструмента в переговорах. Хотя этот аргумент согласуется с некоторыми фактами, известными о депрессии, он игнорирует целый ряд других, например, что депрессия и самоубийство повторяются в семье на протяжении нескольких поколений, что те, кто находится в разводе (без партнера по переговорам), в большей степени подвержены депрессии, чем те, кто женат или кто всегда был холост; что жизнь, полная стрессов, не является необходимым или достаточным условием для депрессии.

Рассмотрение депрессии в качестве переговорного инструмента — это еще одна иллюстрация повсеместных поисков *смысла*, или *предна-*

² Но любой стоящий социобиолог, возможно, сможет придумать этому объяснение в течение получаса.

значения, внешне лишнего его поведения. До определенного момента поиски смысла служат хорошей исследовательской стратегией, но в дальнейшем она приобретает все более надуманный или, как явствует из приведенных примеров, даже абсурдный характер. Существует столько способов сохранения вредоносных черт внутри популяции, что не следует принимать на веру тезис, что часто встречающееся поведение обязательно увеличивает репродуктивную приспособляемость агента³. Естественный отбор, безусловно, благоприятствовал развитию способности испытывать физическую боль. Нет априорной причины, по которой аналогичной поддержки не получила бы боль психическая. Но для того чтобы установить, какую функцию выполняет депрессия, недостаточно предложить объяснение в духе «просто так», которое учитывало бы лишь некоторые из известных черт болезни. Самое главное — гипотеза должна объяснять факты, выходящие за рамки того предмета, который она была призвана истолковать изначально (глава I). Желательно, чтобы она проясняла также *новые факты*, которые стали известны после того, как были предсказаны этой гипотезой.

ИНТЕНЦИОНАЛЬНАЯ ИЗМЕНЧИВОСТЬ, ИНТЕНЦИОНАЛЬНЫЙ ОТБОР

Хотя этот двойной интенциональный механизм был и останется (на что мы лишь можем надеяться) несущественным для объяснения моделей *человеческого* поведения, стоит изучить его структуру.

В «Происхождении видов» Дарвин писал, что природа дает последовательные вариации, а человек адаптирует их в определенных, выгодных ему направлениях. Но это не просто ситуация «человек предполагает, природа располагает», поскольку, как он отмечал, вводные данные могут видоизменяться человеческим поведением.

Высокая степень изменчивости, очевидно, благоприятна, так как составляет обильный материал для деятельности отбора, и при крайней тщательности отбора было бы вполне достаточно простых индивидуальных различий для кумуляции модификации большего размера в почти любом желательном направлении. Поскольку вариации, явно полезные или приятные для человека, возникают только изредка, вероятность их появления будет возрастать при большем числе содержащихся особей. Отсюда, численность имеет величайшее значение для успеха. Исходя из этого принципа, Маршалл (Marshall) заметил ко-

³ Можно утверждать, что в дополнение к субоптимальному поведению, вызванному временными лагами и плейотропией (глава XVI), существует разнородность других генетических механизмов, связанная с наличием у организмов с половым размножением двух различных форм (аллелей) каждого гена.

гда-то об овцах в некоторых частях Йоркшира, что «они никогда не будут улучшаться, потому что обычно принадлежат бедному населению и содержатся маленькими партиями». С другой стороны, владельцы питомников, разводящие в большом количестве одно и то же растение, обычно гораздо успешнее любителей выводят новые и ценные разновидности.

Сегодня мы можем добавить, что искусственный отбор усиливается посредством введения мутаций. Кроме того, поддержание «генетических библиотек» помогает предотвратить сокращение генетических разновидностей, являющееся неизбежным результатом отбора ради отдельных черт.

Что касается самого естественного отбора, Дарвин различал два уровня интенциональности.

В настоящее время выдающиеся животноводы пытаются путем методического отбора, преследующего определенную цель, произвести новую расу или подпороду, превосходящую все прочие в стране. Но для нашей цели важнее та форма отбора, которую можно назвать бессознательным отбором и которая является следствием того, что всякий пытается обладать и поддерживать породу от самых лучших особей. Так, человек, который намеревается держать пойнтеров, естественно, старается достать лучших собак, каких может, и затем получает потомство от них, хотя он не руководился желанием или надеждой на то, что порода будет все время улучшаться⁴.

НЕИНТЕНЦИОНАЛЬНАЯ ИЗМЕНЧИВОСТЬ, ИНТЕНЦИОНАЛЬНЫЙ ОТБОР

Есть множество случаев, когда новый организм или новая форма возникают случайно и принимаются либо отбраковываются на основании интенционального выбора. Если естественный отбор тяготеет к производству одинакового количества мужских и женских особей, инфантицид⁵ по половому признаку или недавно появившиеся аборты, ориентированные на пол будущего ребенка, могут создавать серьезный половой дисбаланс среди населения. Только в Индии и Китае по этой причине «недосчитались» 80 миллионов женщин. Евгеника широко использовалась для запрета обзаведения потомством для психически больных и отсталых лиц. В нацистской Германии на этом основании были стерили-

⁴ Цит. по: Дарвин Ч. Сочинения: в 3 т. М.: Изд-во АН СССР, 1939. Т. 3. С. 19, 16. — *Примеч. ред.*

⁵ Склонность к истреблению своего потомства (у некоторых видов животных). — *Примеч. ред.*

зованы до 400 тысяч человек. По мере совершенствования технологий определения пола ребенка в утробе матери селективный аборт может стать важным определяющим фактором в формировании человеческих популяций. Если в дальнейшем станет возможно определять пол ребенка при зачатии, отбор будет вытеснен интенциональным выбором.

Случайная изменчивость в сочетании с интенциональным отбором может также повлиять на развитие артефактов. Посетивший в 1862 году Англию норвежский священник и социолог Эйлерт Зундт (Eilert Sundt), узнав о дарвиновской теории естественного отбора (ставшей достоянием публики в 1859 году), вознамерился применить ее вариацию к судостроению:

Создатель лодки может быть очень умелым, и все же ему никогда не удастся сделать два совершенно одинаковых судна, даже если он будет стремиться к этому. Возникающие в этом случае вариации можно назвать *случайными*. Но даже самая мелкая из них становится заметна во время плавания, и тогда моряки *неслучайно* замечают, что лодка *улучшилась* или стала более удобной для их целей и они должны предложить *выбрать* ее в качестве модели для *воспроизведения*... Можно посчитать, что каждая из этих лодок безупречна в своем роде, поскольку достигла совершенства путем одностороннего развития в отдельном направлении. Каждый вид улучшения достиг той степени, когда дальнейшее его развитие привело бы к возникновению недостатков, заслоняющих преимущества... И я понимаю этот процесс следующим образом: когда впервые появилась идея новых и улучшенных форм, *длинная серия осторожных экспериментов*, каждый из которых был направлен на небольшое изменение, могла привести к *счастливому* результату, когда со стапеля сходит такая лодка, какую все желали.

В этом тексте Зундт превзошел Дарвина в ключевом аспекте⁶. Если Дарвин признавался, что ему неведом источник изменчивости, Зундту пришлось в голову, что этот источник лежит в *ошибках саморепликации*, напоминающих типографские ошибки и (насколько нам известно) мутации в ДНК. Несовершенство изготовителя лодок, его неспособность создавать точные копии — условие совершенства конечного результата. Зундт

⁶ Это усовершенствование, конечно же, стало возможно только потому, что он обратился к иной проблеме, поскольку в 1862 году ни у кого бы не хватило концептуальной смелости вообразить, что источником вариаций в организме могут быть случайные ошибки в саморепликации. Этот скачок оказался возможным только после того, как Мендель показал дискретную природу единиц наследственности (генов), а Уотсон и Крик доказали, что процесс наследования включает саморепликацию. Интересно, что ответил бы Дарвин, если бы Зундт спросил его: не является ли источником биологического разнообразия несовершенство репродуктивного механизма?

осторожно замечает, что результатом процесса является локальный максимум, никакое дальнейшее улучшение которого путем постепенных изменений невозможно. В последнем предложении он замечает, что процесс может превратиться в искусственный отбор, если люди займутся сознательными экспериментами, не ожидая возникновения случайных вариаций. Как и Дарвин, он утверждал, что интеллект или рациональность могут возникнуть на двух уровнях: во-первых, когда люди *замечают*, что одна модель более пригодна для плавания, чем другая, а потом, когда они *понимают*, что улучшения можно ускорить, если на смену случайным вариациям придут систематические эксперименты⁷.

ИНТЕНЦИОНАЛЬНАЯ ИЗМЕНЧИВОСТЬ, НЕИНТЕНЦИОНАЛЬНЫЙ ОТБОР

Функционирование рынков и естественный отбор имеют некоторые общие черты. Возможны две версии этой аналогии — более близкая к естественному отбору и более отдаленная. Их объединяет общая предпосылка: учитывая разнообразные ограничения человеческой рациональности, фирмы или менеджеры *неэффективны*, поскольку не способны просчитать производственные и маркетинговые решения для максимизации прибыли. Несмотря на это рыночный механизм выбраковывает неэффективных участников, так что в любой момент времени на рынке присутствуют успешные компании. Все происходит так, как будто менеджеры эффективны.

В соответствии с первой и самой простой версией все компании *стремятся* к постоянному наращиванию прибыли посредством копирования и инноваций. Хотя первое само по себе не генерирует новых данных, то есть материал для работы механизма отбора, эту функцию, как уже отмечалось, может выполнять *несовершенное* воспроизведение. Инновация, по определению, также является источником новых данных. Когда в силу абсолютной случайности неточное копирование, или инновация, дают компании возможность производить товар по более низкой цене, она может прибегнуть к демпингу в отношении своих конкурентов и вытеснить их из бизнеса, если только они тоже не применяют более эффективные стратегии. Благодаря обоим механизмам, банкротству и подражанию, эти эффективные технологии получают распространение среди всей совокупности компаний. Если мы предположим, что и копирование, и инновации продвигаются мелкими шажками и что конкуренция происходит в преимущественно стабильной среде, это приведет к достижению локального максимума равновесия прибылей.

⁷ Тем самым можно предотвратить ситуацию, в которой изготовители лодок станут так хороши, что не будут никогда допускать ошибок.

Вторая версия отрицает стремление фирм к максимизации прибыли. Вместо этого они используют стандартные *практики* или правила, которые соблюдаются до тех пор, пока обеспечивается удовлетворительный уровень дохода. Они не столько максимизируют, сколько *удовлетворяют*, если воспользоваться таким неологизмом. Значение этого выражения может зависеть от многих факторов, но для простоты мы можем предположить, что компания, прибыль которой постоянно оказывается ниже удовлетворительного уровня, или обанкротится, или станет жертвой недружественного поглощения. Самая простая стратегия — делать все как раньше, пока прибыли на удовлетворительном уровне. Более сложные практики могут включать установление цен со стабильной наценкой на издержки или инвестирование определенного процента прибыли в новое производство. Идея удовлетворительности отражена в принципах «нельзя резать курицу, несущую золотые яйца», или «работает — не улучшай». В одном отношении *удовлетворизация* могла бы даже быть оптимальной. Как уже было сказано выше: «Величайшая монопольная выгода — тихая жизнь».

Теперь предположим, что прибыль падает ниже удовлетворительного уровня. Компания, из года в год делающая одно и то же, может стать жертвой некоего организационного аналога коррозии или склероза. Внешние потрясения (рост цен на нефть или изменение важного обменного курса) могут привести к росту издержек или сокращению дохода. Может измениться потребительский спрос; конкуренты могут усовершенствовать методы или представить новые продукты; рабочие могут устроить дорогостоящую забастовку. Какова бы ни была причина (зачастую даже неизвестная компании), неудовлетворительный уровень прибыли подтолкнет к поискам новых стратегий путем сочетания инновации и копирования. Обе процедуры скорее всего будут локальными, то есть ограниченными альтернативами, близкими к уже существующим практикам. Любого рода крупные изменения могут оказаться слишком дорогостоящими для переживающей трудные времена компании (глава IX), а резкие скачкообразные инновации требуют слишком больших концептуальных затрат.

Процесс копирования ориентирован на следование поведению успешных соперников. Является инновация случайной или целенаправленной, зависит от восприятия причин кризиса, вызвавшего ее возникновение. Если падение прибыли стало результатом роста цен на нефть, фирма может начать поиски методов, которые позволили бы сэкономить на этом ресурсе⁸. Если же оно спровоцировано изменением курса доллара и евро, компания, скорее всего, обратится к случайным поискам.

⁸ Если не ожидается дальнейшего роста цен, рациональность не требует от компании применения *нефтесберегающих технологий*, так как никто

В любом случае в ее поведении будет присутствовать мощная интенциональная составляющая. Решение изменить текущие практики — интенционально, так же как решение об объеме необходимых инвестиций в инновации или копирование. Выбор объекта для подражания носит сознательный характер, и компания может намеренно склониться в пользу практик определенного направления.

Новые практики, возникающие из этих процессов, становятся полем слепой игры рыночных сил. Если они дают компании возможность достичь удовлетворительного уровня прибыли, она прекращает поиски до нового кризиса. Если нет, то компания может предпринять новую попытку или будет вынуждена объявить о банкротстве. Рано или поздно все неудовлетворительные компании выбывают. Сам по себе этот процесс не способствует возникновению фирм, ориентированных на максимизацию прибыли. Чтобы понять, как это получается, нужно вписать в эту картину конкуренцию. Если мы предположим, что одной из практик фирм является инвестирование определенного процента прибыли в новое производство, то из них, по чистой случайности открывшие более продуктивные в сравнении с конкурентами методы, будут осуществлять экспансию, так что со временем распространенность этих практик среди компаний значительно возрастет⁹.

МОДЕЛИ ОТБОРА В СОЦИАЛЬНЫХ НАУКАХ

Полезность таких моделей зависит от эмпирического вопроса: какова скорость устранения неэффективных фирм по сравнению со скоростью изменения среды? В предыдущей главе я поднимал этот вопрос по отношению к естественному отбору и утверждал, что слишком хорошо отрегулированная адаптация организмов к условиям их обитания говорит о том, что они менялись сравнительно медленно. В случае экономической среды мы можем дать прямую оценку. В современном мире компании сталкиваются с беспрецедентной скоростью изменений. Если бы они ограничивались только регулярными наблюдениями за обстановкой, они оказались бы хронически не приспособленными к ней. Вероятность успеха больше для фирм, способных *предугадывать* изменения, стреляя на опережение. Большую часть времени такая стратегия дает

не может знать, каков набор возможных инноваций (см. главу XIV). При этом рост цен на нефть делает эти инновации более заметными.

⁹ Согласно этой версии, компании не пытаются сознательно вытеснить конкурентов (например, пользуясь высокими прибылями, продавать товар по заниженной цене до тех пор, пока другие не покинут рынок), поскольку они не озабочены получением «более-чем-просто-удовлетворительной» прибыли.

сбои, но по крайней мере не всегда. Более того, благодаря своему политическому влиянию крупные корпорации способны *воздействовать* на среду, в которой присутствуют. Важными или неважными были описанные мною механизмы на более ранней стадии «дикого» капитализма среди мелких фирм, мы не знаем. Сегодня они мало что объясняют в наблюдаемых явлениях.

На кону стоит общий вопрос. В ответ на обвинения в недостаточном реализме теоретики рационального выбора обычно возражают, что они всего лишь пытаются объяснить поведение на основании допущения, что люди как будто стремятся максимизировать своими действиями полезность (или прибыли, или любую другую цель). Часто они добавляют, что такое допущение может быть обосновано механизмом отбора. Предполагается, что в экономической сфере эту функцию выполняет конкуренция. Я утверждал, что это невозможно¹⁰. В самом общем виде это возражение можно сформулировать так: даже если можно было бы показать, что рыночная конкуренция действительно повышает эффективность путем устранения неэффективных компаний, между «повышением эффективности» и сверхсложной как будто максимизацией, приписываемой компаниям в экономических моделях, все еще остается гигантский шаг.

В политической сфере предполагается, что благодаря выборной конкуренции единственные наблюдаемые нами политики — те, кто были избраны или переизбраны, таким образом, можно заключить, что все политики ведут себя так, как будто они озабочены исключительно своими избирательными перспективами. Однако скачок от электоральной озабоченности к *исключительной* озабоченности не оправдан. Методологически беспристрастный взгляд на политику показывает, что существует три вида политических акторов: оппортунисты (обеспокоенные только тем, чтобы быть избранными), реформаторы (заботящиеся о том, чтобы проводить в жизнь определенную политику) и активисты (которые больше озабочены тем, чтобы сделать заявление)¹¹. Взгляд на поли-

¹⁰ В дополнение к упомянутому мною общему аргументу экономика командного спорта может опровергнуть идею максимизации прибыли посредством отбора. Если бы максимизирующие прибыль бейсбольные и футбольные команды использовали свои прибыли для скупки всех лучших игроков в лиге, их превосходство в какой-то момент стало бы так велико, что сами матчи потеряли бы часть своей неопределенности и, следовательно, часть удовольствия от игры, то есть в конечном счете способность приносить прибыль.

¹¹ Более формально эти три группы можно различить следующим образом. Оппортунисты предпочитают предлагать политику А, а не политику В, когда А имеет более высокие шансы на победу, чем В при условии, что оппозиционная партия предлагает некоторую фиксированную политику С. Активисты

тику как на взаимодействие между этими тремя группами внутри одной и среди разных партий гораздо реалистичнее, чем представление о политике как о «палатках с мороженым» (глава XIX), согласно которой партии, максимизирующие голоса, будут тяготеть к центру. Ярким опровержением тезиса об исключительной озабоченности политических деятелей вопросом избрания на новый срок, может служить линия французских политиков от Жана Жореса через Леона Блюма и Пьера Мендес-Франс к Мишелю Рокару, которые были мотивированы желанием беспристрастно продвигать ценности социальной справедливости и экономической эффективности. Необходимо заметить, что в случае Рокара его неприязнь к электоральной политике отрицательно сказалась на его политической эффективности.

Вне зон конкуренции как будто рациональность еще менее оправдана. Действие механизмов отбора, имитирующих рациональность, не распространяется на потребительские предпочтения, поведение избирателей, посещение церкви, выбор профессии и прочие виды поведения. Разумеется, они подпадают под действие *ограничений*, которые могут снизить значение выбора в целом и рационального выбора в частности (глава IX). Ограничения действуют еще до самого факта, делая некоторые варианты выбора неосуществимыми. Отбор действует после факта, устраняя тех, кто сделал определенный выбор. Хотя оба механизма вносят свой вклад в объяснение поведения, ни вместе, ни по отдельности они не могут объяснить его полностью. Выбор остается ключевой концепцией в социальных науках.

Библиографические примечания

Статистические данные по убийствам и насилию над детьми взяты из книги М. Дейли и М. Уилсона «Убийство» (Daly M., Wilson M. Homicide. New York: Aldine de Gruyter, 1988). Возражения против предложенного этими авторами объяснения содержатся в главе 7 книги Д. Баллера «Приспосабливающийся мозг» (Buller D. Adapting Minds. Cambridge, MA: MIT Press, 2005). По поводу двух сторон аргумента о самообмане см. Р. Триверс «Социальная эволюция» (Trivers R. Social Evolution. Menlo Park, CA: Benjamin-Cummings, 1985) (где отдается предпочтение эволю-

сты предпочитают предлагать А, а не В, когда рядовой член партии извлечет из А большую полезность, чем из В (независимо от того, что представляет собой политика С). Реформаторы предпочитают предлагать А, а не В с учетом того, что противоположная сторона предлагает С, когда ожидаемая полезность для рядового члена партии от А выше, чем от В. Таким образом, оппортунистов интересуют только вероятности, активистов — только полезность, а реформаторов — и то и другое.

ционному объяснению) и В.С. Рамачандран и С. Блейкли «Фантомы в мозге» (*Ramachandran V.S., Blakeslee S. Phantoms in the Brain. New York: Quill, 1998*) (с противоположными взглядами). По поводу двух сторон адаптивной природы депрессии см. Э.Х. Хагген «Модель депрессии как инструмента переговоров» (*Haggen E.H. The bargaining model of depression // Hammerstein P. (ed.). Genetic and Cultural Evolution of Cooperation. Cambridge, MA: MIT Press, 2003*) (где выдвигается эволюционное объяснение) и П. Крамер «Против депрессии» (*Kramer P. Against Depression. New York: Viking, 2005*) (где оспаривается этот взгляд). Анализ рынков в категориях естественного отбора происходит из работы А. Алхияна «Неопределенность, эволюция и экономическая теория» (*Alchian A. Uncertainty, evolution, and economic theory // Journal of Political Economy. 1950. No. 58. P. 211–221*). Его наиболее изощренная версия (которая не поддерживает как будто максимизацию) представлена в работе Р. Нельсона и С. Уинтера «Эволюционная теория экономических изменений» (*Нельсон Р., Уинтер С. Эволюционная теория экономических изменений. М.: Дело, 2002*). Теория «удовлетворизации» позаимствована из «Поведенческой теории рационального выбора» Г. Саймона (*Simon H. A behavioral theory of rational choice // Quarterly Journal of Economics. 1954. No. 69. P. 99–118*). Экономика командных видов спорта рассматривается в книге Д. Берри, М. Шмидта и С. Брук «Выплата выигрышей» (*Berri D., Schmidt M., Brook S. The Wages of Wins. Stanford, CA: Stanford University Press, 2006*). Различие между оппортунистами, реформаторами и активистами взято из книги Дж. Ремера «Политическая конкуренция» (*Roemer J. Political Competition. Cambridge, MA: Harvard University Press, 2001*). Касательно аргумента «как будто», см. также библиографические примечания к главе I.

ВЗАИМОДЕЙСТВИЕ

Ч А С Т Ь П Я Т А Я

Социальное взаимодействие может принимать разные формы. (1) Исход для каждого агента зависит от исходов для других агентов. Взаимозависимость исходов может возникнуть, если материальное или психическое благополучие других влияет на мое собственное (глава V). (2) Исход для каждого может зависеть от действий всех. Эта взаимозависимость отражает общую социальную взаимообусловленность (глава 18), иллюстрируемую глобальным потеплением (являющимся плодом рук человеческих). (3) Действие каждого зависит от (предвосхищаемых) действий всех. Эта взаимозависимость является специфическим предметом изучения *теории игр* (главы XIX и XX), в рамках которой рассматриваются также (1) и (2). (4) Убеждения каждого зависят от действий всех. Эта взаимозависимость может возникнуть посредством нескольких механизмов (например, плюралистическое неведение или информационные каскады) (глава XXIII). (5) Предпочтения каждого зависят от действий всех. Эта взаимозависимость является хуже всего понятым аспектом социальной жизни. Хотя я уже затрагивал некоторые аспекты этого вопроса, общее его изложение дано в главе XXII.

Эти взаимозависимости могут возникнуть через децентрализованные действия индивидов, между которыми нет организованных отношений (глава XXIV). Однако социальная жизнь зачастую более структурирована. Многие результаты возникают в процессе согласования позиций — спора, голосования, торга, посредством которого группы индивидов достигают решений, обязательных для каждого из них (глава XV). Наконец, организации действуют по правилам, предназначенным для того, чтобы приводить в соответствие индивидуальные интенции и цели организации (глава XXVI).

XVIII. Непреднамеренные последствия

НЕПРЕДНАМЕРЕННЫЕ ПОСЛЕДСТВИЯ ИНДИВИДУАЛЬНОГО ПОВЕДЕНИЯ

Не всегда случается так, как мы планировали. Многие события происходят неумышленно. Иногда причины тривиальны (например, когда мы жмем вместо тормоза на газ или по ошибке бьем по клавише «удалить»). Однако некоторые механизмы более целенаправленные. Хотя создание «общей теории непреднамеренных последствий» вряд ли реально, можно по крайней мере начать их систематизировать. Я рассматриваю случаи, в которых события являются не только непреднамеренными, но и непредсказуемыми. Предсказуемые побочные эффекты действия не являются преднамеренными, особенно если они отрицательны, но я не буду включать их в число непреднамеренных последствий действия.

Непреднамеренные последствия могут возникнуть как из индивидуального поведения, так и при социальном взаимодействии. Начав с первого, мы можем распространить схему желание—возможности, намеченную в главе IX (рис. XVIII.1).

Если действия возникают на основе желаний (или предпочтений), они сами могут формировать желания. Соответственно, помимо предусмотренного исхода действия есть непреднамеренный: изменение желания. Хороший пример — наркотическая зависимость. Под действием наркотических веществ люди начинают активнее дисконтировать будущее, тем самым ослабляя сдерживающий эффект утраченного, связанный с долгосрочным вредом от зависимости. Если бы этот эффект можно было предвидеть, агент, вероятно, постарался бы избежать пагубного пристрастия, но, как правило, этого не происходит. Подобные явления наблюдаются и в повседневных ситуациях. Я иду на вечеринку, намериваясь выпить всего два бокала, чтобы потом иметь возможность уехать домой на машине, но после второго бокала моя решимость растворяется в алкоголе, и я выпиваю третий. Если бы я знал об этом заранее, то выпил бы только один бокал.

Бессознательное предпочтение нового или изменения также может влиять на желания («У соседа всегда трава зеленее»). В сказке Ганса Христиана Андерсена «Уж что муженек сделает, то и ладно» крестьянин

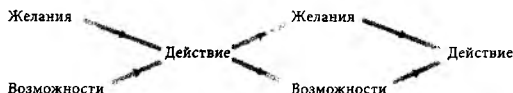


РИС. XVIII.1

с утра отправляется на рынок продать или обменять лошадь. Сначала ему попадает человек с коровой, которая так ему нравится, что он меняет ее на лошадь. В последующих транзакциях корова обменивается на овцу, овца — на гуся, гусь — на курицу и, в конце концов, курицу меняют на мешок гнилых яблок. Путь к разорению крестьянина был вымощен пошаговыми улучшениями. Всякий раз крестьянин полагает, что обмен повышает его благосостояние, но суммарный результат всех обменов ужасающ¹.

Эффект вклада, следствие неприятия убытков (глава XII), тоже иллюстрирует вызванное выбором, но непреднамеренное изменение предпочтения². Многие блага приобретают большую субъективную ценность для владельца после приобретения, чем до него, что подтверждается тем, что минимальная цена продажи обычно превышает максимальную цену покупки с коэффициентом, варьирующим от 2 до 4. Когда ситуация отказа от большинства благ воспринимается в качестве потери, а получение — в качестве выигрыша, поскольку потери переживаются тяжелее, чем такие же приобретения, в этом и заключается эффект, предсказанный механизмом неприятия потерь. К тому же эксперименты доказывают, что будущие покупатели недооценивают минимальную перепродажную цену, лишний раз подтверждая, что изменение предпочтений дей-

¹ При более формальном подходе представим человека, который регулярно (хотя и бессознательно) корректирует свои желания так, чтобы все время больше желать того, чего у него сейчас меньше всего. Предположим, он сталкивается с последовательностями благ, сгруппированных по два: $(\frac{1}{2}, \frac{3}{2})$, $(\frac{3}{4}, \frac{1}{2})$, $(\frac{1}{4}, \frac{3}{4})$, $(\frac{3}{8}, \frac{1}{4})$... Тогда, если в данный момент он потребляет благо n в последовательности, а в дальнейшем ему предлагается выбрать между набором n и набором $n+1$, он всегда будет выбирать последний, потому что тот дает больше блага, которого у него сейчас меньше всего. Но поскольку последовательность стремится к нулю, эти локальные улучшения ведут к полному краху. Этот эффект напоминает ситуацию агента, который может умереть от потери крови, перебирая предпочтения по кругу (глава XI), но механизм в этом случае другой.

² Так случилось, что термин «эффект вклада» стал использоваться для обозначения как тенденции к переоценке благ, которыми владеет агент, так и полезности, которую человек может извлечь в настоящем из полезности в прошлом (глава II). Оба эти значения не связаны между собой.

ствительно непредсказуемо³. Другой механизм, который может породить этот искажающий эффект укрепления, то есть тенденцию воспринимать свой выбор в положительном свете, раз он уже сделан, предлагает теория когнитивного диссонанса (глава I).

В качестве примера того, как действие может влиять на *возможности* непреднамеренным и непредвиденным образом, рассмотрим хулигана, который добивается своего в транзакциях с другими людьми, потому что они предпочитают уступить ему, а не оказывать сопротивление. Непреднамеренным следствием его поведения может стать то, что другие начнут его сторониться, таким образом, у него окажется меньше возможностей для транзакций с ними. Он добивается своего в каждом таком взаимодействии, но взаимодействий становится все меньше. Последнее следствие может оказаться не только непреднамеренным и непредвиденным, но и незаметным. Ему кажется, что тактика хулиганства действует⁴. Если он все же замечает негативный эффект своего поведения, он может продолжить упорствовать, если позитивные эффекты перевешивают. В таком случае негативные последствия можно предвидеть, но они не могут быть преднамеренными.

Часто выбор одного варианта сегодня может устранить некоторые из множества допустимых возможностей в будущем. Этот эффект можно предвидеть: ограничения моего бюджета позволят мне купить один автомобиль, но не два. Но иногда агент может не знать, что выбор имеет необратимые последствия. Крестьянин может владеть куском земли, на котором есть деревья и поля. Вырубка леса вызывает эрозию почвы, оставляя ему меньше земли для возделывания. В примерах, которые я собираюсь кратко упомянуть, эрозия может быть следствием *коллективного* поведения. Например, если эрозия может начаться тогда, и только тогда, когда деревья вырубят крестьянин и два его соседа. Но очень часто бывает так, что индивид в силу своего неведения подрывает возможности для дальнейшего действия. В этом виноват *когнитивный дефицит*: агент не может предсказать будущие последствия настоящего поведения. В других случаях это может произойти из-за *мотивационного дефицита*: агент придает слишком маленькое значение (известным и определенным) будущим последствиям в сравнении с выигрышем в настоящем (глава VI).

³ Действительно, трудно понять, как это можно спрогнозировать. Человеку придется одновременно учитывать два момента: состояние до покупки, относительно которого благо является приобретением, и состояние после нее, относительно которого оно является потерей.

⁴ Эту ограниченную перспективу разделяют некоторые социальные теоретики, которые утверждают, что такие эмоции, как гнев, могут быть рациональными или по крайней мере адаптивными, потому что дают агентам возможность добиться своего во взаимодействии с другими.

ЭКСТЕРНАЛИИ

Обратимся теперь к непредвиденным последствиям *интеракции* — ключевой теме социальных наук. Согласно памятной фразе Адама Фергюсона, история есть «результат усилий человека, но не осуществление его замысла». Его современник Адам Смит ссылаясь на «невидимую руку», которая влияет на поступки человека. Полстолетия спустя Гегель упоминал «хитрость разума» для того, чтобы объяснить прогресс свободы и истории. Примерно в то же время Токвиль сделал похожее заявление, что в прогрессе демократии «каждый играет свою роль: те, кто стремился обеспечить успех демократии, равно как и те, кто и не думал ей служить; те, кто за нее боролся, точно так же, как и те, кто объявил себя ее врагом». Несколько лет спустя Маркс писал об отчуждении людей от их собственных действий, утверждая, что «это закрепление социальной деятельности, это консолидирование нашего собственного продукта в какую-то вещную силу, господствующую над нами, вышедшую из-под нашего контроля, идущую вразрез с нашими ожиданиями и сводящую на нет наши расчеты, является одним из главных моментов во всем предшествующем историческом развитии».

Среди этих авторов только Адам Смит и Маркс указали специфические механизмы для производства непреднамеренных последствий. Говоря современным языком, они указали, как *внешние факторы, экстерналии* (externalities) поведения, могут накапливаться и порождать не предусмотренные и не предполагавшиеся агентами исходы. Представим себе, что каждый из многочисленных идентичных агентов предпринимает некое действие в собственных интересах. В качестве побочного продукта этого действия он приносит небольшие издержки или небольшую выгоду (негативная или позитивная экстерналия) каждому из агентов, в том числе себе. В дальнейшем каждый агент становится объектом множества подобных действий. Суммировав эффект, а затем сложив сумму с частной прибылью агента мы получаем итоговый результат, который агенты произвели своими действиями. Поскольку мы допускаем, что все агенты идентичны, их первоначальное состояние, состояния, которых они намериваются добиться, и состояния, которых они коллективно добиваются, могут быть представлены простыми числами x , y и z соответственно⁵.

⁵ Многие из явлений, которые я перечисляю в этой книге, экономисты не считают экстерналиями. Они скорее всего включают загрязнение окружающей среды, но не эффекты, порождаемые рынком, такие как кейнсианская безработица. Но для моей цели важнее то, что их объединяет: в погоне за собственной выгодой каждый индивид приносит небольшие издержки или небольшую прибыль всем остальным *и себе самому*. Компания, проводящая сокращение персонала или урезающая зарплаты, вызовет небольшое снижение спроса на товары собственного производства.

Сначала рассмотрим модель $z > y > x$, образующую положительную экстерналию. Она стала основным предметом интереса Адама Смита: когда агент ставит «свое дело так, что то, что он производит, становится величайшей ценностью, притом что он думал только о обственной прибыли, в этом, как и во многом другом, им руководит „невидимая рука“, содействующая движению к цели, не входившей в его намерения. И это отсутствие намерения не всегда так уж плохо для общества. Преследование своего интереса зачастую гораздо эффективнее способствует интересам общества, чем когда ему пытаются помочь умышленно». В рыночной конкуренции цель каждой компании — получить прибыль, произведя более дешевые товары, чем у конкурентов, но при этом выигрывают и покупатели. Покупатели, будучи рабочими или менеджерами, своими усилиями, направленными на победу в конкуренции, тоже могут приносить пользу другим. Результатом этого становится заметный долгосрочный рост. Этот эффект можно было предвидеть или нет, но он, конечно же, не являлся «частью» намерения.

Теперь возьмем $y > z > x$, то есть слабую отрицательную экстерналию. Благополучие агентов благодаря их усилиям повышается, но в меньшей степени, чем они ожидали, из-за тех издержек, которые они заставляют друг друга нести. Возможно, людям приятнее ездить на работу на машине, чем на общественном транспорте, если он плохо развит, но пробки и загрязнение окружающей среды приводят к тому, что они получают от этого меньше выгоды, чем предполагали. Если внешний фактор связан с пробками, он не останется незамеченным ими. Если же он результат загрязнения, может потребоваться какое-то время, чтобы понять, что люди сами наносят друг другу вред, а не являются жертвами (допустим) промышленного загрязнения.

Наконец, обратимся к сильной отрицательной экстерналии, где $y > x > z$. В результате общего стремления преуспеть каждому становится только хуже. Это одно из основных обвинений, выдвигавшихся Марксом против децентрализованной капиталистической экономики. Именно такую общую структуру имеет его основное объяснение кризисов капитализма, «теория падения нормы прибыли». Для поддержания роста прибыли, по его утверждениям, у каждого капиталиста есть стимул заменять физический труд машинным. Но когда все капиталисты делают это одновременно, они рубят сук, на котором сидят, поскольку конечным источником прибыли является прибавочная стоимость, производимая трудом. Этот аргумент соблазнителен, но при ближайшем рассмотрении он оказывается неправильным по целому ряду причин. Гораздо интереснее другое замечание, сделанное Марксом походя и легшее позднее в основу теории безработицы, разработанной Джоном Мейнардом Кейнсом. Как отмечает Маркс, капиталист находится в неоднозначных

отношениях с рабочими. С одной стороны, он хочет платить *своим* рабочим низкую зарплату, чтобы получать более высокую прибыль. С другой стороны, он желает, чтобы у всех *других* рабочих были высокие зарплаты, чтобы обеспечить высокий спрос на производимые им товары. Оба этих требования вполне удовлетворимы для одного капиталиста, но, очевидно, не для всех капиталистов одновременно. Это противоречие капитализма Кейнс описывал следующим образом: каждый капиталист отвечает на падение прибыли сокращением рабочих мест, тем самым экономя на заработной плате. Но поскольку спрос со стороны работников и есть то, что прямо или косвенно поддерживает компании, вследствие одновременного сокращения штатов всеми капиталистами прибыль еще больше сократится, а это в свою очередь приведет к новым сокращениям или банкротствам.

Вот некоторые частные случаи сказанного выше. Чрезмерный вылов рыбы, вырубка леса и чрезмерное выбивание пастбищ («трагедия общин») могут быть индивидуально рациональными, но коллективно субоптимальными и даже пагубными. Если в каждой семье в развивающейся стране родится как можно больше детей в качестве защиты от бедности в старости, перенаселенность приведет к еще большей бедности. При нехватке воды каждый индивид, использующий воду для второстепенных нужд, увеличивает вероятность отключения воды на несколько часов в день, что отразится и на ее использовании для базовых целей. Эти последствия можно предвидеть или нет. Ключевой чертой этой категории непреднамеренных последствий является то, что даже если их можно предвидеть, поведение останется прежним. В следующей главе показано, что это *доминирующая стратегия*: ее рационально выбирать независимо от того, что делают другие.

ИНТЕРНАЛИИ

Отчасти сходное рассуждение применимо к внутренним факторам, интерналиям, определяемым как выгода или вред, который выбор человека может принести благосостоянию, получаемому им от последующих выборов. Говоря метафорически, интерналии — это экстерналии, которые человек налагает на свои будущие я. Обсуждая опеку над детьми (рис. XI.3), я утверждал, что время, проведенное с ребенком, создает положительную интерналию для родителя. Пагубное пристрастие служит хорошим примером отрицательной интерналии. Чем больше человек потреблял вызывающее зависимость вещество, тем меньше удовольствия он получает от его нынешнего потребления. Такой же эффект привыкания может возникнуть от благ, не вызывающих зависимости. Даже если вам нравится мороженое с орехами пекан, оно вам надоеет,



РИС. XVIII.2

если вы будете есть его пять раз в день. Кроме того, потребление в прошлом имеет еще один эффект. Хотя оно и делает настоящее потребление менее приятным, оно еще усиливает разницу в благополучии между потреблением и отсутствием потребления в настоящем («уход»). Схематически это представлено на рис. XVIII.2.

Таким образом, независимо от воздержания или потребления в прошлом, в настоящем агенту лучше от потребления блага, чем от его отсутствия. Потребление — доминирующая стратегия. В то же время от повторяющегося потребления агенту становится хуже в любое время (за исключением первоначальных случаев), чем от повторяющегося воздержания, подобно тому как от стратегии иметь много детей хуже становится всем. Конечно, есть очевидные различия между интерналиями и экстерналиями. Одно из них — временная асимметрия: если все индивиды могут причинять друг другу вред, будущие я не могут навредить прошлым. Другое различие заключается в том, что последовательные я — лишь временные слои *одного* человека, принимающего решения, тогда как разные индивиды не являются разнесенными в пространстве частями одного суперорганизма. Как только человек (один и тот же) понимает, что его выбор в настоящем может оказать негативное воздействие на благополучие, которое он может получить от последующих выборов, у него появляется стимул изменить свое поведение. Насколько

силен этот стимул, зависит от того, насколько сильны симптомы ухода, и от того, в какой степени агент дисконтирует будущее благополучие. Некоторые агенты, которые никогда бы не сделали первый шаг, если бы знали о последствиях, оказавшись на крючке, решают не выходить из игры.

СИНДРОМ МЛАДШЕГО БРАТА

Непреднамеренные последствия социального действия могут порождаться *механизмом младшего брата*. Я проиллюстрирую его популярным примером из экономической теории — «паутинной моделью» рынка, называемой еще «свинным циклом», так как впервые ее выдвинули в качестве объяснения циклических колебаний в производстве свиней. Очень часто такие паттерны выделялись в колебаниях в судостроительной промышленности, когда за рыночной конъюнктурой, выгодной для продавцов, следовало переинвестирование и насыщение. Выбирая профессию на основании текущего спроса на выпускников, студенты могут коллективно разрушить фундамент своих решений.

Когда свиноводы решают, сколько свиней собираются поставить на рынок в следующем году, это решение определяется предполагаемой ценой за свинину и затратами на выращивание животных. Увеличение ожидаемой цены заставляет фермеров произвести больше (см. восходящую кривую поставок на рис. XVIII.3).

Предположим, что в год 1 цена на свинину составляет p_1 . Предполагая, что цены останутся такими же в год 2, фермеры в следующем году выбрасывают на рынок q_2 . Однако при таких объемах цена установления рыночного равновесия между спросом и предложением составляет p_2 , а не p_1 . Предполагая, что цены останутся на прежнем уровне и в год 3, фермеры производят на этот год объем продукции q_3 . Цена равновесия будет составлять p_3 , заставляя фермеров производить q_4 в год 4, и так далее. В этом случае цены и объем продукции образуют направленную вонне спираль, или «паутинную модель», выделенную на диаграмме жирными линиями. Приятные сюрпризы перемежаются неприятными, но предполагаемый доход так и остается недостижимым. Если относительный наклон кривых предложения и спроса будет модифицирован, может быть получена направленная вовнутрь спираль, сходящаяся к цене равновесия p^* и к объему равновесия q^* .

В поведении фермеров есть нечто иррациональное. Каждый из них считает, что может свободно варьировать объем продукции, чтобы максимизировать прибыль, молчаливо предполагая, что все остальные автоматически будут делать то же, что и в прошлом году. Хотя это поведение, возможно, иррационально, оно, безусловно, интеллигильно.

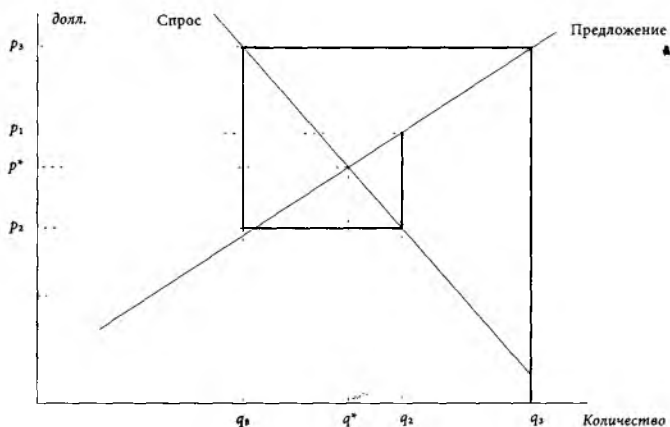


РИС. XVIII.3

Французский философ Морис Мерло-Понти говорил, что у нас есть спонтанная склонность считать других людей «младшими братьями». Нам трудно наделить других такими же способностями к взвешиванию и обдумыванию, которыми, как подсказывает нам интроспекция, владеем мы сами, равно как и приписать им те же смятение, сомнение и тревогу (глава XXIII). Кажется, мысль о том, что другие так же расчетливы и могут мыслить стратегически, как и мы сами, не возникает естественным путем.

Три примера, касающиеся поведения избирателей, иллюстрируют эту идею. Предположим, я — член левого крыла социалистической партии своей страны. Я бы, конечно, предпочел, чтобы у власти находились социалисты, а не коммунисты, но поскольку, судя по опросам, социалисты должны получить твердое большинство, я голосую за коммунистов, чтобы моя партия сдвинулась влево. Однако я не задаю вопроса: а не будут ли так же рассуждать и другие члены левого крыла. Если многие из них рассудили как я, коммунисты могут победить на выборах. Намерение произвести результат высокого уровня (победа социалистов с сильным присутствием коммунистов) может породить результат низкого уровня (победа коммунистов). Еще более распространенный сценарий: если многие избиратели останутся дома, будучи уверенными, что их партия выиграет, партия может проиграть. Наконец, вспомним пример из главы XII. Возможным объяснением приведше-

го к катастрофе объявления досрочных выборов может стать неспособность Ширака предугадать, что избиратели сделают из его решения вывод о его убеждениях, а не поведут себя в соответствии с данными опросов.

Неспособность видеть в других интенциональных и максимизирующих агентов наблюдается, когда законодатели или администраторы предлагают политику, саморазрушающуюся по мере привыкания к ней агентов. Согласно римскому праву, кража одной лошади или быка делала человека вором скота, тогда как кража менее четырех свиней или менее десяти овец не считалась преступлением. Комментатор отмечает: «При таком положении вещей можно ожидать, что кража трех свиней или восьми овец должна широко распространиться». Чтобы обеспечить стабильную занятость, многие страны приняли законы, по которым увольнение работников, проработавших два и более года становится незаконным. На это работодатели рационально ответили привлечением сторонних ресурсов или заключением временных контрактов с работниками, тем самым понизив степень защищенности работников. Города могут строить скоростные дороги, чтобы сократить количество пробок, а потом выясняется, что поскольку количество людей, добирающихся на работу на машинах увеличилось, дороги забиты, как и раньше, а загрязнение воздуха выросло. Правительство может ограничить иммиграцию состоящими в браке с человеком, который уже находится в стране на легальных основаниях, что приведет к увеличению числа людей, вступающих в брак лишь с этой целью. Освобождение от призыва для студентов создает стимул поступать в колледж. Даже создание открытой системы поощрения может выйти боком. Так, когда Чикагское управление государственных школ создало в 1996 году систему, в которой школы, показавшие плохие результаты, подвергались наказанию, ее изобретателям не пришло в голову, что, помимо создания стимула для учителей, чтобы те лучше работали, система поощряла завышение оценок, что и произошло в действительности.

Синдром младшего брата имеет важные социальные последствия. Токвиль отмечал, что в годы, предшествовавшие Французской революции, высшие слои общества открыто разоблачали пороки режима и его опустошающее воздействие на народ, как будто последний был глух к тому, что они говорили: «Это напоминает мне о чувствах мадам Дюшатель, которая, по словам секретаря Вольтера, раздевалась в присутствии слуг-мужчин, не считая своих лакеев мужчинами». Секретарь в своих мемуарах писал, что «родовитые дамы считали своих лакеев всего лишь автоматами». Одновременная демонстрация презрения по отношению к низшим слоям общества и разоблачение их нищеты подготовило умы к революции. Из более свежих примеров, аргумент,

стоящий за «кривой Филлипса», согласно которому правительство может обеспечить низкую безработицу ценой высокой инфляции, предполагает, что социальные акторы ничего не знают об этой политике. Когда правительства пытались достичь своей цели, стратегическое поведение рациональных профсоюзов и других игроков подрывало их усилия, вызывая стагфляцию — высокую инфляцию в сочетании с высокой безработицей.

В отличие от вызванных экстерналиями непреднамеренные последствия, порожденные синдромом младшего брата, могут прекратить свое существование, как только агенты его осознают. В описанных мною случаях нет доминирующих стратегий, только оптимальные при (имплицитном) допущении, что другие менее рациональны, чем я. Как только все агенты начнут считать других рациональными, их поведение будет стремиться к предсказуемому результату. Все фермеры, выращивающие свиней, будут ожидать, что возобладает цена равновесия. Действуя в соответствии с этим предположением, они произведут объем, соответствующий точке равновесия. Такое разделяемое всеми убеждение перформативно. Эта идея раскрывается в следующей главе.

Библиографические примечания

Свидетельства воздействия пагубной зависимости на дисконтирование во времени приводятся у Л. Джордано и др. в «Лишение легких опиатов увеличивает степень дисконтирования отложенного получения героина и денег у пациентов с зависимостью от опиатов» (Giordano L. et al. Mild opioid deprivation increases the degree that opioid-dependent outpatients discount delayed heroin and money // *Psychopharmacology*. 2002. No. 63. P. 174–182). Источником замечания, опирающегося на модель сказки Андерсена, стала работа С. С. фон Вайцекера «Замечания об эндогенном изменении вкусов» (Weiszacker C. C. von. Notes on endogenous change of tastes // *Journal of Economic Theory*. 1971. No. 3. P. 345–372). Относительно обсуждения взглядов Маркса на непреднамеренные последствия см. мою книгу «Осмысляя Маркса» (Making Sense of Marx. Cambridge University Press, 1985), глава 1.3.2 и далее по тексту. Модель зависимости взята из статьи Г. Беккера и К. Мерфи «Теория рационального привыкания» (Беккер Г., Мерфи К. Теория рационального привыкания // *Экономическая теория преступлений и наказаний*. 2001. № 3). Превосходное концептуальное исследование непреднамеренных последствий содержится в книге Т. Шеллинга «Микромотивы и макроповедение» (Schelling T. Micromotives and Macrobehavior. New York: Norton, 1978). Что касается идеи внутренних факторов, см. Р. Хернштейн и др. «Максимизация полезности и улучшение: внутренние факторы в индивидуальном выборе» (Herrnstein R. et al. Utility maximization and melio-

ration: Internalities in individual choice // Journal of Behavioral Decision Making. 1993. No. 6. P. 149–185). Пример с чикагскими государственными школами взят из работы Б.Джейкоба и С. Левитта «„Гнилые яблоки“: исследование распространения и прогностических факторов обмана среди учителей» (Jacob B., Leavitt S. Rotten apples: An investigation of the prevalence and predictors of teacher cheating // Quarterly Journal of Economics. 2003. No. 118. P. 843–877).

XIX. Стратегическое взаимодействие

СТРАТЕГИЧЕСКОЕ ВЗАИМОДЕЙСТВИЕ С ОДНОВРЕМЕННЫМИ ВЫБОРАМИ

Изобретение *теории игр* можно считать наиболее важным единичным достижением в социальных науках в XX веке. В некоторых случаях она позволяет нам объяснять поведение, которое до сих пор представлялось загадочным. Но что еще более важно, она проясняет структуру социального взаимодействия. Когда ты смотришь на мир через призму теории игр, или теории взаимосвязанных решений, как ее еще называют, все выглядит иначе.

Для начала рассмотрим игры, в которых агенты принимают решения одновременно. Цель в том, чтобы понять, могут ли, и если да, то каким образом, n агентов или *игроков* самостоятельно добиться координации своих *стратегий*. Особое внимание мы уделим специальному случаю — $n = 2$. Игроки могут общаться друг с другом, но не могут вступать в соглашения, накладывающие обязательства. Каждой из n -ого количества стратегий, по одной на каждого агента, соответствует свой *исход*. Агент ранжирует возможные исходы в соответствии со своим *порядком предпочтений*. При необходимости будем считать, что условия для представления предпочтений как количественных полезностей соблюдаются (глава XI). *Структура награды* — это функция, которая любому множеству стратегий приписывает множество полезностей. Хотя само слово «награда» может ассоциироваться с денежным вознаграждением, оно будет использоваться для обозначения психологических исходов (полезностей и, в конечном счете, предпочтений). Когда структуры денежных или материальных и психологических наград расходятся, релевантна только последняя.

Как было кратко упомянуто в предыдущей главе, у агента может быть *доминирующая* стратегия, дающая результат лучший, чем тот, который мог быть получен при выборе любой другой стратегии независимо от поведения остальных. *Исход*, но не сам *выбор* агента может зависеть от действий других. В иных случаях есть взаимозависимость выборов. Если все едут по левой стороне дороги, мне тоже лучше выбрать левую, если они едут по правой, мне тоже лучше так сделать.

Равновесие — это n -ое множество стратегий с тем свойством, что ни один игрок не может, отступив от своей равновесной стратегии, в одностороннем порядке получить исход, который он наверняка предпочтет равновесному. Иначе говоря, при равновесии стратегия, выбранная каждым игроком, является наилучшим ответом на стратегии, выбранные другими игроками, в том смысле, что ему *ничего не остается*, как выбрать стратегию равновесия, если другие уже выбрали свои стратегии. Однако эта стратегия необязательно должна быть оптимальной в том смысле, что ему будет хуже, если он отступит от нее в одностороннем порядке. В общем случае в игре может быть несколько равновесий. Кратко остановимся на нескольких примерах. Допустим, есть только одно равновесие. Кроме того, предположим, что структура наград и рациональность игроков — общеизвестный факт¹. При таком допущении мы можем предсказать, что все игроки выберут стратегию равновесия, так как это единственная стратегия, основанная на рациональных представлениях о том, что делают остальные.

Некоторые игры с уникальным равновесием вращаются вокруг существования доминирующих стратегий. Это может означать один из двух случаев, представленных на рис. XIX.1². В аварии с участием двух автомобилей повреждены обе машины. В происшествии с участием пешехода и автомобиля пострадал только пешеход. Авария машина—машина происходит только, если по крайней мере один из водителей был неосторожен. Если оба были неосторожны, исход хуже. Происшествия с автомобилем и пешеходом возникают только при обоюдной неосторожности. Должная осторожность обходится дорого. Из этих предпосылок следует,

¹ Факт является общеизвестным, если все его знают, все знают, что все остальные тоже его знают, все знают, что все остальные знают, что все его знают и так далее. Чтобы не слишком полагаться на фразу «и так далее», которая указывает на бесконечную последовательность убеждений, эта идея может быть сформулирована следующим образом: нет такого n , что факт является общеизвестным до уровня n в последовательности, но не до уровня $n + 1$. Приведем простой пример: общеизвестное знание может присутствовать в классе. Когда учитель что-то рассказывает ученикам, они все это знают, все знают, что все остальные это знают, и так далее.

² Согласно устоявшимся правилам, первое число в каждой клетке представляет выигрыш для «игрока в строке», который выбирает между нижней и верхней стратегиями, а второе — выигрыш «игрока в столбце», который выбирает между левой и правой стратегиями. В зависимости от контекста выигрыши могут быть количественными полезностями, порядковыми полезностями, деньгами или чем-нибудь еще, что могут максимизировать игроки. На рис. XIX.1 выигрыши могут рассматриваться в качестве замены порядковых полезностей, отражающих предпочтения, а не исходы. Здесь и далее точки равновесия обведены кружком.

| | | А) | | В) | |
|--------|----------------------|--------|-----|---------|-----|
| | | МАШИНА | | ПЕШЕХОД | |
| МАШИНА | Должная осторожность | 5:5 | 2:3 | 0:2 | 0:3 |
| | Неосторожность | 3:2 | 1:1 | 1:2 | 1:1 |

РИС. XIX.1

что в авариях двух машин должная осторожность является доминирующей стратегией для обоих водителей. В аварии с машиной и пешеходом неосторожность — доминирующая стратегия для водителя.

У пешехода нет доминирующей стратегии, поскольку должная осторожность — лучший ответ на неосторожность, а неосторожность — на должную осторожность. Поскольку он знает, что у водителя доминирующая стратегия — неосторожность и что, будучи рациональным, он ее выберет, пешеход тем не менее выберет должную осторожность³.

Игры, в которых у всех игроков есть доминирующие стратегии очень распространены. В теории они несколько тривиальны, за исключением периодически повторяющихся случаев. Игры, в которых только у некоторых игроков есть доминирующие стратегии, заставляющие других делать четкий выбор, распространены меньше, но тоже важны. Для них, однако, существуют более строгие информационные требования, поскольку пешеходы должны знать возможные исходы не только для себя, но и для водителя, тогда как двум водителям достаточно знать только свой исход. Часто мы приписываем доминирующие стратегии другим. Так, мы обычно не смотрим по сторонам, переходя дорогу с односторон-

³ Согласно некоторым юридическим исследованиям, важной функцией деликтного (т.е. вытекающего из обязательств при причинении имущественного вреда. — *Примеч. ред.*) права является использование штрафов и возмещения убытков в целях преобразования системы поощрений таким образом, чтобы возникшее равновесие обладало некоторыми желательными свойствами (эффективность или справедливость).

ним движением, потому что предполагаем, что боясь аварии, водители соблюдают правило односторонности.

Особый класс игр обладает координационным *равновесием*, часто называемым конвенцией, при котором у каждого игрока не только нет стимула отступать в одностороннем порядке, он предпочитает, чтобы этого не делали другие. При равновесии, когда все едут по правой стороне дороги, может произойти авария, если я отступлю от правила или это сделает кто-то еще. В этом случае равновесие не является уникальным, поскольку езда по левой стороне имеет те же особенности⁴. Часто не так важно, что мы делаем, пока мы все делаем одно и то же. Значения слов произвольны, но как только они зафиксированы, они становятся конвенцией. В других случаях неважно, что мы делаем, но важнее, чтобы все делали одно и то же. Ниже я вернусь к этим примерам.

ПРИМЕРЫ ДУОПОЛИИ

Некоторые игры обладают уникальным равновесием, не вращающимся вокруг доминирующих стратегий. Примером может служить поведение дуополии (рис. ХІХ.2). Когда на рынке доминируют две фирмы, снижение выпуска продукции одной компанией вызовет рост цен и рост выпуска продукции другой. Иными словами, у каждого игрока есть график наилучшей реакции, который диктует ему объемы выпуска продукции в зависимости от таковых у конкурента. При равновесии объем продукции одной фирмы — лучшая реакция на этот же показатель другой. Из этого положения не следует, что это единственное возможное для них решение. Если бы они образовали картель и ограничили выпуск продукции уровнем ниже равновесного, обе получили бы более высокую прибыль, хотя эти обоюдно оптимальные уровни производства не являются наилучшей взаимной реакцией. На самом деле компании сталкиваются с дилеммой заключенного (определенной на рис. ХІХ.3).

В качестве еще одного случая дуополии рассмотрим двух продавцов мороженого на пляже, которые пытаются найти лучшее место для своих прилавков, исходя из допущения, что покупатели (которые, как предполагается, равномерно распределены по берегу) подойдут к ближайшей из них. Доминирующая стратегия отсутствует. Если один поставит палатку на некотором расстоянии слева от центра пляжа, лучший ответ со стороны другого продавца будет встать справа, на что лучшим ответом со стороны первого продавца будет тоже сдвинуться вправо и так далее, пока палатки не окажутся рядом в центре пляжа. Это уникальное равнове-

⁴ Хотя отсутствие уникальности не вытекает из формального определения, это общая черта координационных игр в реальной жизни.

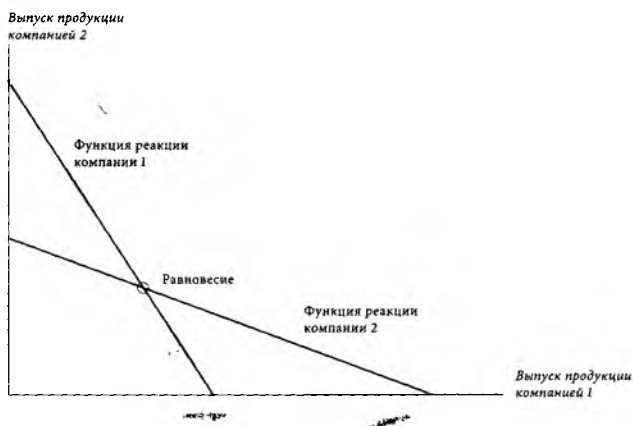


РИС. XIX.2

сие, очевидно, не самое лучшее для покупателей. Для них наилучшим решением было бы расположение палаток между центром и одним из концов пляжа. Хотя для продавцов такой исход ничуть не хуже равновесного, это расположение не является наилучшим ответом друг другу. Данная модель также применялась для объяснения тенденции политических партий (в двухпартийной системе) сдвигаться к середине политического спектра.

Предположим, что, когда палатки стоят в центре, покупатели на краях пляжа отказываются от покупки мороженого, потому что к тому моменту, как они вернутся назад, оно уже растает. Если покупатели готовы пройти пешком не более половины пляжа — четверть расстояния, чтобы дойти до палатки, и еще четверть, чтобы вернуться, оптимальным для них исходом будет уникальное равновесие, поскольку ни у кого нет стимула перемещаться. Предположим, что длина пляжа составляет 1000 метров. Если продавец на отметке 750 метров передвинется на отметку 700, он потеряет 50 покупателей между отметками 950 и 1000, которые не готовы идти пешком более 500 метров, и приобретет 25 покупателей между отметками 475 и 500, к которым его палатка теперь стала ближе, чем конкурирующая, что станет чистым убытком. Похожее рассуждение может объяснить, почему политические партии никогда не передвигаются в центр: экстремисты с обоих концов предпочитают воздерживаться и не голосовать за центристскую партию. Кроме того, как я отмечал в конце главы XVII, маловероятно, чтобы максимизация голосов была единственной целью политических партий.

НЕКОТОРЫЕ РАСПРОСТРАНЕННЫЕ ИГРЫ

Несколько простых структур взаимодействия с выигрышами, представленными на рис. XIX.3, очень часто встречаются в самых разных контекстах⁵. *C* и *D* означают кооперацию и отказ от кооперации. В игре «Телефон» игрок в колонке — тот, кто звонил первым. В игре «Координатор» *A* и *B* могут быть любой парой действий, которые оба игрока предпочитают скорее координировать, чем не координировать, но при этом им безразлично, каким из двух способов они будут это делать.

Игры иллюстрируют структуру двух центральных проблем социального взаимодействия — *кооперацию* и *координацию*. В обществе без взаимовыгодной кооперации жизнь будет «одинокой, бедной, неприятной, грубой и короткой» (Гоббс). То, что она будет *предсказуемо* плохой, — слабое утешение. Жизнь в обществе, в котором люди неспособны координировать свое поведение, будет полна непреднамеренных последствий, как «сказка, рассказанная идиотом, полная шума и ярости, но лишенная всякого смысла» («Макбет»). И кооперация, и координация иногда достигают успеха, но часто терпят неудачу. Теория игр может прояснить как их успех, так и неудачу.

«Дилемма заключенного», «Охота на оленя» и «Цыпленок» так или иначе предполагают выбрать кооперацию или отказаться от нее. «Дилемма заключенного» называется так, потому что первоначально для ее иллюстрации использовался следующий сюжет. Каждому из двоих преступников, замешанных в одном и том же преступлении, но находящихся в разных камерах, говорят, что если он донесет на другого, а тот не сделает того же, его выпустят на свободу, а его подельник сядет в тюрьму на десять лет. Если они оба не донесут друг на друга, оба отправятся в тюрьму на пять лет⁶. В таких обстоятельствах донос является доминирующей стратегией, хотя обоим преступникам будет лучше, если никто доносить не станет. Исход определяется комбинацией «соблазна безбилетника» (выйти на свободу) и «страхом, что тебя подставят» (десять лет тюрьмы).

Негативные экстерналии, обсуждавшиеся в предыдущей главе, могут рассматриваться как «Дилемма заключенного» со множеством игроков.

⁵ Хотя здесь они описываются как игры с двумя участниками, эти игры можно экстраполировать на ситуацию с множеством агентов. Например, «Дилемма заключенного» для $n+1$ игроков представлена на рис. XXIV.2.

⁶ Выигрыш в «Дилемме заключенного», представленный на рис. XIX.3, может показаться искусственным. Для нашей цели имеет значение только (порядковое) ранжирование исходов. Позднее выигрыши будут интерпретированы как денежные награды.

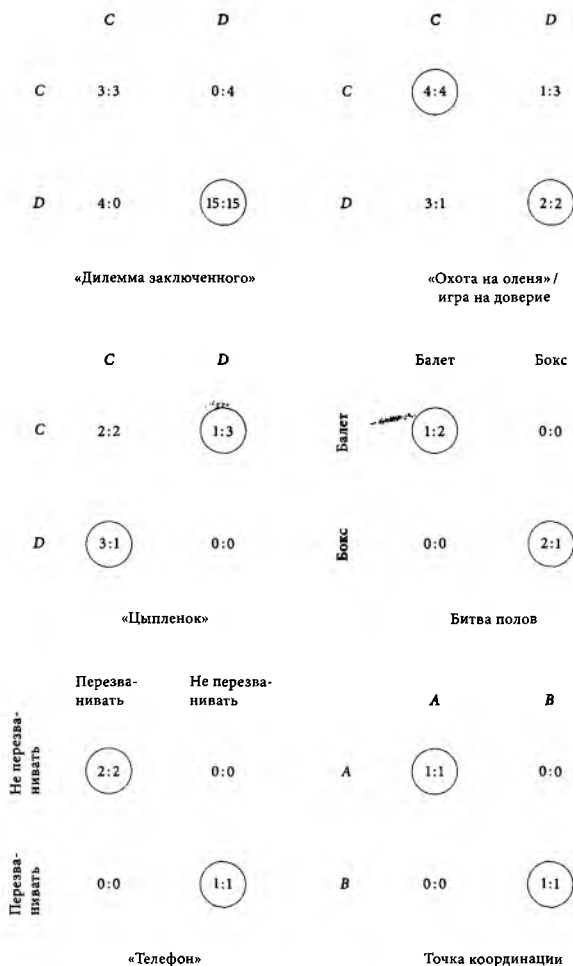


РИС. XIX.3

Будут и другие примеры. Для каждого работника (с учетом его эгоистических мотиваций) лучше не быть членом профсоюза, даже если для всех лучше, чтобы все в него вступили и получили более высокую заработную плату. Для каждой компании в картеле лучше выйти из него

и выпускать больше продукции, чтобы воспользоваться высокими ценами, вызванными ограничениями на выпуск продукции для других компаний, но когда все это сделают, цены упадут до конкурентного уровня; максимизация прибыли каждой компанией подрывает максимизацию общей прибыли. Организация стран — экспортеров нефти (ОПЕК) также уязвима в этом отношении. Ситуации, в которых нужно бежать как можно быстрее, чтобы удержаться на том же месте, как например, во время гонки вооружений между США и Советским Союзом, в политическом консультировании или в случае студентов, которые ищут работу для преподавателя, который «ставит оценки по восходящей».

Идею игры «Охота на оленя» часто приписывают Жану-Жаку Руссо, хотя его язык не всегда прозрачен⁷. В более условной форме она включает в себя двух игроков, которые могут выбирать — охотиться за оленем (С) или за зайцем (D). Зайца каждый может поймать сам, для того чтобы добыть оленя, необходимо (и достаточно) совместное усилие обоих. Пол-оленя лучше, чем целый заяц. Если к охоте на зайца подключатся оба охотника, на это уйдет больше времени и сил, потому что шум, который поднимают охотники, распушивает зайцев. Как и в «Дилемме заключенного» у охотника, который будет охотиться на оленя, есть риск потерпеть фиаско, пока другой добудет зайца. Однако соблазна «безбилетничества» здесь нет. У игры две точки равновесия, в верхней левой и нижней правой клетках.

Хотя первое равновесие лучше, оно может быть неосуществимо. Чтобы понять почему, мы можем выдвинуть предположение, что структура выигрышей общеизвестна и она позволяет агентам получить ошибочные представления о структуре выигрышей других агентов. Действия, основанные на такой вере, образуют *равновесие в слабом смысле*, если для каждого агента действия, предпринятые другими, подтверждают его убеждения. Предположим, что в «Охоте на оленя» каждый агент ошибочно полагает, что у других агентов предпочтения «Дилеммы заключенного». Учитывая это, рациональный агент должен уклониться от сотрудничества, тем самым подтвердив убеждения других о том, что у него предпочтения «Дилеммы заключенного». Такое общество может столк-

⁷ «Если охотились на оленя, каждый понимал, что для этого он обязан оставаться на своем посту; но если вблизи какого-либо охотника пробегал заяц, приходилось сомневаться, что этот охотник без зазрения совести пустится за ним вдогонку и, настигнув добычу, весьма мало будет сокрушаться о том, что таким образом лишил добычи своих товарищей» (цит. по: Руссо Ж.-Ж. Рассуждение о происхождении и основаниях неравенства между людьми // Тракаты / пер. с франц. А. Хяютина. М.: Наука, 1969. С. 75. — Примеч. ред.). Можно сказать, что преследование зайцев утверждается в качестве доминирующей стратегии.

нуться с высоким уровнем уклонения от налогов и коррупции. Я вернусь к таким случаям плюралистического неведения в главе XXIII. В другом обществе, в котором люди справедливо полагают, что другие предпочтут «Охоту на оленя», возникнет хорошее равновесие, в котором люди платят налоги, не дают и не берут взятки. Культуры коррупции могут зависеть от убеждений, а не от мотиваций.

Международный контроль над инфекционными заболеваниями может иметь структуру «Охоты на оленя». Если хотя бы одна страна не сможет принять адекватных мер, остальные будут не в состоянии себя защитить⁸. В качестве другого примера рассмотрим контртеррористические меры. Если хотя бы одна из двух стран их финансирует, она приносит благо и другой стране, и себе. Если издержки превысят ее личную выгоду, она не будет заниматься односторонним финансированием. Если вкладывают обе страны, возможность разделить информацию может обеспечить каждой из них более высокий уровень безопасности, чем тот, которого она могла бы достичь, эксплуатируя ресурсы другой.

В этих примерах структура выигрышей связана с каузальной природой ситуации. В «Охоте на оленя» и в случае контроля над заболеваниями «пороговая технология» предполагает, что индивидуальные усилия бессмысленны. В случае борьбы с терроризмом в основании лежит своего рода экономика масштаба: десять единиц усилий приносят в два раза больший эффект, чем пять. В других случаях структура выигрыша связана с тем фактом, что агенты заботятся не только о своей материальной выгоде, но и о других вещах. Такие ситуации принято называть игрой на доверие (Assurance Game). Даже если структура материального выигрыша представляет собой «Дилемму заключенного», каждый индивид может согласиться сотрудничать, если *удостоверится*, что другие тоже сотрудничают. Желание быть честным, или нежелание быть «безбилетником», может помочь преодолеть соблазн наживаться на кооперации других. В противном случае «Дилемма заключенного» может превратиться в игру на доверие.

Рассмотрим выигрыши в «Дилемме заключенного» (рис. XIX.3) как материальные награды. Предположим, что полезность для каждого человека равна его материальной награде плюс половина материальной награды другого. В этом случае выигрыш в полезности будет соответствовать игре на доверие на рис. XIX.4. «Дилемма заключенного» может трансформироваться в игру на доверие при помощи третьего механизма, если третья сторона введет штраф за выбор некооперативной стратегии *D*. Если

⁸ Это крайнее упрощение, которое приведено здесь исключительно для примера.

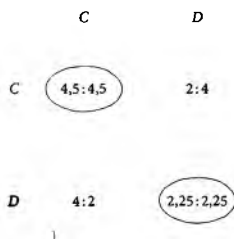


РИС. ХІХ.4

мы снова будем рассматривать выигрыши в «Дилемме заключенного» на рис. ХІХ.3 как денежные награды и предположим, что агентов волнуют только они, вычитание 1,25 из награды за отсутствие сотрудничества превратит ее в игру на доверие. Например, профсоюз может ввести формальные или неформальные санкции для работников, не оформивших свое членство. Наконец, можно превратить «Дилемму заключенного» в игру на доверие, награждая за сотрудничество, например, предложив бонус или взятку в размере 1,25 за сотрудничество. Обещание награды, однако, должно исполняться, тогда как санкции необязательно приводить в жизнь, если они действенны. Если выигрыш «безбилетника» очень высок, выгоды от сотрудничества могут быть недостаточны, чтобы вкладываться во взятки⁹. В некоторых случаях используются награды. Рабочие, вступающие в профсоюз, могут выиграть не только благодаря более высокой заработной плате, которая обычно в равной степени достается и нечленам, но и от пенсионных планов и дешевого отпуска, доступного только участникам.

Игра «Цыпленок» названа так в честь подросткового ритуала из фильма 1955 года «Бунтарь без причины». Лос-анджелесские подростки на краденых машинах едут на крутой обрыв. В этой игре двое из них одновременно направляют машины на край обрыва, останавливаясь в последнюю минуту. Тот, кто первым остановится, «Цыпленок», проиграл. В другом варианте две машины едут навстречу и проигрывает первый, свернувший в сторону. В каждой из двух точек равновесия каждый агент

⁹ Независимо от того, что используется — награды или наказания, затраты на наладку системы и мониторинг агентов также должны финансироваться за счет прибыли от сотрудничества. На практике это может сделать такого рода мероприятия невозможными или слишком затратными.

делает противоположное тому, что делает другой. Даже имея общее знание о структуре вознаграждения и рациональности агента, мы не можем предсказать, какое равновесие будет выбрано (если оно будет выбрано). С точки зрения рационального выбора ситуация является *неопределенной*. Во второй версии игры («Поворот») участник может попытаться разрушить неопределенность, надев себе повязку на глаза (которая видна другому), тем самым заставив другого игрока свернуть. Но это создает такое же затруднение с двумя вариантами — «надевание» и «ненадевание повязки», а не со «свернет» и «не свернет»¹⁰. Это глубоко фрустрирующая ситуация.

В каком-то отношении гонка вооружений представляет собой игру в «Цыпленка». Карибский кризис часто приводят в качестве примера, в котором две супердержавы оказались в ловушке конфронтации, и СССР «моргнул первым». Еще один пример. Двое фермеров используют одну и ту же систему ирригации для своих полей. Уход за системой может адекватно осуществляться и одним фермером, но она в равной степени выгодна обоим. Если один из двух не выполнит свою часть работы по уходу, второй заинтересован в том, чтобы выполнить свою. Случай Китти Дженовезе также может рассматриваться в этой перспективе, если мы предположим, что каждый сосед предпочел бы вмешаться тогда, и только тогда, когда больше никто не вмешался.

Обратившись теперь к вопросам *координации*, рассмотрим «Войну полов». В основе истории лежит следующий стереотип. Мужчина и его жена собираются сходить куда-нибудь вечером. Они решили после работы пойти или на балет, или на боксерский поединок и окончательно договориться по телефону. Его телефон сломался, поэтому им приходится координироваться молча. Они оба заинтересованы в том, чтобы побыть вместе, но в том, что касается места посещения, их интересы расходятся. Как и в игре «Цыпленок», в этой игре две точки равновесия — координация на балете или на матче по боксу. И так же, как в игре, общее представление о структуре выигрышей и рациональности не подскажут паре, где встретиться. Ситуация снова будет неопределенной.

Игры такого рода возникают, когда координация может принимать разные формы; любая из них лучше для всех агентов, но каждая для некоторых агентов предпочтительнее других¹¹. В социальной и политиче-

¹⁰ Точно так же «решение» «Дилеммы заключенного», состоящее в том, что каждый человек обещает сотрудничать, заново задает еще одну «Дилемму заключенного» с выбором между тем, чтобы «сдерживать обещание» или «отступить».

¹¹ Как мы увидим позднее (глава XXV), вопрос разделения прибыли от кооперации может рассматриваться в рамках теории сделки, специального раздела теории игр.

ской жизни это скорее исключение, чем правило. Все граждане могут предпочесть любую конституцию (внутри определенного набора режимов) ее полному отсутствию, потому что долгосрочная стабильность нужна им, чтобы иметь возможность планировать будущее. Когда закон зафиксирован и не меняется, можно регулировать свое поведение в соответствии с ним. При этом каждая заинтересованная группа может выбрать специфическую конституцию: лобби кредиторов предпочтет запрет на бумажные деньги, политическая партия предпочтет ту выборную систему, которая ей выгодна, те, у кого есть сильный кандидат в президенты, захотят, чтобы был предусмотрен сильный президентский пост и так далее.

Множественные точки равновесия при координации возникают и в том случае, когда в странах складываются разные стандарты веса, длины или объема, а позднее они открывают потенциальные выгоды общего решения. Континентальная Европа и англосаксонский мир сохранили в этой области разные стандарты. В отличие от множественных конституционных решений препятствием к достижению договоренности является не постоянное расхождение интересов, но краткосрочные издержки на переход. Кроме того, выбор стандарта может оказаться игрой в «Цыпленка». Сделаем маловероятное предположение, что стандарт записан в конституции как статья, не подлежащая изменению. У каждой страны тогда будут стимулы принять на себя обязательства прежде, чем это сделают другие.

Игра «Телефон» определяется потребностью в правиле, которое подсказало бы сторонам, что делать, когда телефонный разговор прервался. Есть две точки равновесия координации: номер должен набрать человек, который звонил, или же тот, кто принял вызов. Любое правило лучше ситуации, в которой они оба будут набирать номер или его не будет набирать никто. Но в этом случае, в отличие от «Битвы полов», одно равновесие для обоих лучше, чем другое. Разумнее, если номер перенаберет первый, потому что он лучше знает, по какому номеру звонил. Рациональные, владеющие всей информацией агенты сойдутся на равновесии более высокой координации. Однако такое утверждение игнорирует стоимость нового вызова. Если стоимость высока, игра становится «Битвой полов».

Наконец игра «Точка координации», которая может восприниматься как вариант «Битвы полов». Супруги договорились посмотреть фильм, который идет в кинотеатрах А и В, но не договорились, куда именно пойти. Мы предполагаем, что ни один из кинотеатров не находится ближе, чем другой, и не обладает никакими другими дополнительными преимуществами. Как и в «Битве полов» ни общее знание, ни рациональность не подскажут им, куда идти. Однако в ситуации может при-

существовать психологический сигнал, который послужит как бы фокусной точкой для координации. Если у пары в кинотеатре А было первое свидание, это может заставить их прийти именно туда. В этом случае сигнал является частным событием. В других случаях сигналы могут разделяться значительной частью населения. Так, нью-йоркский анекдот утверждает, что если вы с вашим спутником потеряли друг друга, нужно встречаться в полдень под главными часами на Большом центральном вокзале. И даже не зная этот анекдот, многие люди встречаются на вокзалах, поскольку во многих городах это самые важные здания, существующие в единственном экземпляре¹². Уникальность делает их привлекательными в качестве точки координации. Полдень обладает тем же свойством¹³.

Эффект точки координации легко продемонстрировать в экспериментах. Если попросить всех членов группы записать положительное целое число на клочке бумаги и сказать всем, что они получают награду, если все запишут одну и ту же цифру, они неизменно сойдутся на 1. Это уникальное наименьшее целое число, но не уникальное наибольшее. В других контекстах в качестве точки координации может выступать 0. В период холодной войны во время дебатов о том, могут ли США использовать тактическое ядерное оружие и не перевести при этом эскалацию в полномасштабную ядерную войну, предлагались разные идеи «яркой отметки», которая позволит ограничить его использование. В конце концов было решено, что единственной точкой координации может быть *отсутствие* его использования.

Сходное замечание о важности обычая сделал Паскаль: «Почему мы следуем старым законам и старым мнениям? Потому что они лучше? Нет, но они уникальные и устраняют источник различия». В другом месте он писал:

Самые неразумные вещи в мире становятся самыми разумными из-за расстройства ума. Что может быть более неразумного, чем выбирать для управления государством первого сына королевы? Мы же не выбираем для управления судом самого родовитого из его пассажиров. Подобный закон был бы смешон и несправедлив. Но поскольку люди таковы и всегда будут таковыми, каковы они есть, это становится разумным и справедливым, ибо кого еще избрать? Самого добродетельного и самого способного? И вот уже все хватаются за оружие, утверждая, что они самые добродетельные и самые способные. Тогда

¹² В Нью-Йорке тот, кто не знает этот анекдот, не пойдет на центральный вокзал, поскольку Пенн-стейшен лишает его уникальности. Они, в свою очередь, могут скоординировать встречу возле Эмпайр-стейт-Билдинг.

¹³ Хотя полночь тоже является точкой координации, она проигрывает полудню из-за неудобства.

давайте наделим этим качеством нечто бесспорное. Это старший сын короля: это четко, здесь не о чем спорить. У разума нет другого выхода, потому что гражданская война — худшее из зол.

Подобное рассуждение действительно может повлиять на выбор короля, если есть несколько претендентов на престол. При выборе короля во время Реставрации во Франции Талейран успешно доказал, что законный наследник последнего короля Франции был единственной точкой координации, которая могла предотвратить распри. Как он писал в своих мемуарах: «*Навязанный* король был бы результатом силы или интриги; и то и другое было бы недостаточно. Для установления долговременной системы, которая будет принята, не встретив сопротивления, нужно действовать согласно принципу». Позднее Маркс утверждал, что Республика 1848 года обязана своим существованием тому факту, что для каждой из двух ветвей королевской семьи это была вторая наилучшая альтернатива. Похожее наблюдение сделал Токвиль, объясняя стабильность правления Наполеона III. Демократию тоже можно рассматривать в качестве точки координации. Когда существует множество конкурирующих между собой оснований для утверждения людьми своего превосходства — мудрость, богатство, добродетель, происхождение, уникальную значимость приобретает количественное решение по принципу большинства. Бывшие колонии, в которых племена говорят на разных языках, могут выбрать в качестве официального язык колонизатора. Конфликтующие стороны легко сходятся на предложении, стоящем у всех на втором месте.

В 1989 году перезахоронение останков Имре Надя стало точкой координации для 250 000 человек, прошедших маршем по улицам Будапешта, чтобы продемонстрировать свое недовольство режимом. Как и в предыдущих примерах, некая точка позволила осуществить координацию через кооперацию. В конфликтных ситуациях точки координации могут приносить совсем иной эффект. Во время Крымской войны французский генерал Пелиссье решил назначить вторую атаку на Севастополь на 18 июня 1855 года, чтобы доставить удовольствие Наполеону III тем, что одержит победу в годовщину битвы при Ватерлоо. Поскольку эта дата и ее значение для французов были прекрасно известны русским, они смогли предвидеть нападение и разгромить французов.

Один из уроков, которые нужно вынести из данной главы, заключается в том, что ситуации в реальном мире могут моделироваться в зависимости от принятых допущений. Гонка вооружений моделировалась как «Дилемма заключенного», как игра в «Цыпленка» и игра на доверие. Вступление в профсоюз может быть «Дилеммой заключенного» или игрой на доверие. Повторный набор телефонного номера рассматривал-

ся как «Битва полов» или как «Телефон». Координация мер и весов может быть игрой в «Цыпленка» или «Битвой полов». Тонкости структуры взаимодействия не всегда сразу заметны. Принуждая нас к откровенности по отношению к природе взаимодействия, теория игр может раскрыть тонкости или отклонения, о которых мы даже не подозревали.

ПОСЛЕДОВАТЕЛЬНЫЕ ИГРЫ

Позвольте коротко остановиться на играх, в которых агенты принимают *последовательные решения* (более подробно я остановлюсь на них в следующей главе). Начну с простого примера, который показывает способность теории игр прояснять структуры взаимодействия, которые раньше находили лишь смутное понимание¹⁴.

На рис. XIX.5 две армии противостоят друг другу на границе между своими странами. Генерал I может либо отступить, сохранив статус-кво (3, 3), либо вторгнуться в чужую страну. Если он перейдет в наступление, генерал II может либо сразаться с исходом (1:1), либо уступить спорную территорию с исходом (4:2). Прежде чем генерал I примет решение, генерал II может сообщить о намерении сразаться в случае нападения, в надежде подвигнуть генерала I выбрать (3:3), а не (1:1). Однако *такая угроза не заслуживает доверия*. Генерал I знает, что как только он перейдет в наступление, в интересах II будет уступить, а не сразаться. Единственный равновесный исход — (4:2). Эта концепция равновесия — не статическая «лучшая реакция», которую мы до сих пор рассматривали. Скорее это динамическая концепция, начинающаяся с поздних стадий игры и постепенно возвращающаяся к более ранним (технический термин — «обратная индукция»). Сначала мы спрашиваем, какое поведение будет рациональным для II, если I нападет на его страну. Ответ «уступить» ведет к исходу (4:2). Таким образом, у I есть выбор между действиями, ведущими к (3:3) и к (4:2). Будучи рациональным, он выбирает последнее.

Как замечал Фукидид в «Пелопонесской войне», чтобы другая сторона основывала на них свое поведение, обещания должны быть правдоподобными:

Клятвы, принесенные в поддержку примирения, недолговечны, поскольку были сделаны каждой из сторон в отсутствие любого иного источника силы, позволяющего выйти из тупика; но тот, кто заставлял противную сторону врасплох в данный момент и хватался за первую возможность нанести смелый удар, будет наслаждаться местью тем более сладкой, потому что он воспользовался чужой доверчивостью, а не вы-

¹⁴ Я придерживаюсь допущения, что рациональность и информация являются общеизвестными.

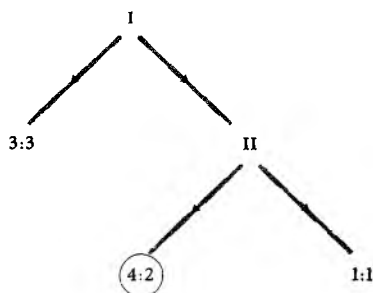


РИС. XIX.5

играл в открытой битве... Ибо нет такого слова, на которое можно целиком положиться, и нет такой клятвы, которая могла бы принести примирение, и все, кто находится в превосходящем положении, поняв, что на безопасность надеяться нельзя, приняли меры во избежание урона, а не позволили себе довериться другим.

Иначе говоря, человек, получивший обещание должен задаться вопросом, рационально ли для того, кто обещает, держать слово. Например, если в игре «Доверительное управление» игрокам позволено общаться, второй игрок может подтолкнуть первого сделать большой трансфер, пообещав ему большой обратный трансфер. Если его ничто не заставляет исполнять обещание, оно не заслуживает доверия. В «Демократии в Америке» Токвиль саркастически замечает о письме военного секретаря индейцам-чероки, в котором тот «утверждает, что они должны оставить надежду удержать занимаемую ими в данный момент территорию, но дает им положительные заверения на сей счет, если они пересекут реку Миссури, как будто тогда у него будут полномочия, которыми он сейчас не обладает». Экономическая реформа в Китае была сопряжена с подобной проблемой. Когда правительство начинало аграрную реформу, оно пообещало фермерам 15-летнюю аренду земли, чтобы стимулировать мелиорацию. Поскольку способа заставить автократическое правительство сдерживать свои обещания не существует, многие фермеры ему не поверили и сразу же распорядились своей прибылью. Автократическое правительство не может лишить себя способности вмешиваться.

Понятие надежности является центральным для теории игр второго поколения, которое появилось приблизительно в 1975 году (первое поколение возникло около 1945 года). Как только мы принимаем эту идею

всерьез, мы должны задать вопрос, как агенты могут *инвестировать в надежность*, чтобы придать действенность своим угрозам и обещаниям. Существует несколько механизмов. Один из них — *зарабатывание репутации*, например, когда приобретают репутацию частично или периодически иррационального человека. Известно, что президент Никсон, поощряемый Генри Киссинджером, сознательно поддерживал образ непредсказуемого, чтобы убедить Советы в способности действовать вопреки интересам Америки, если его спровоцировать. Кроме того, люди могут выполнять угрозы, когда это противоречит их интересам, чтобы приобрести репутацию жестких и крутых, которая заставит остальных поверить в их угрозы позднее.

Еще один механизм — *предварительное связывание себя обязательствами*, обсуждавшееся в главе XIII. В этом случае предварительное связывание себя обязательствами рассматривается как второй лучший рациональный ответ на склонность агента к иррациональному поведению. В стратегическом плане предварительное связывание себя обязательствами может быть полностью рациональным. В игре, изображенной на рис. 19.5, генерал II может построить «машину судного дня», которая автоматически инициирует ядерный удар по противнику, если тот начнет вторжение. Если существование этой машины и то, что страна II не контролирует ее действие, общеизвестны, это должно предотвратить вторжение. Наоборот, генерал II может использовать стратегию «сжигания мостов», то есть отрезать любые пути к отступлению. И снова это должно поспособствовать сдерживанию генерала I, если он знает, что у генерала II нет других альтернатив, кроме как дать отпор в случае нападения.

В некоторых случаях обе стороны могут использовать предварительное связывание себя обязательствами, чтобы получить преимущество перед другими. В трудовых переговорах с руководством угрозы забастовки и бойкота могут выглядеть неубедительными. Менеджмент знает, что, поскольку рабочие должны оплачивать ипотеку и содержать семью, они не могут позволить себе длительную забастовку. Профсоюз знает, что, так как фирма должна выполнять договора на поставку, она не может допустить остановки производства. Чтобы усилить убедительность своих угроз профсоюз может инвестировать в забастовочный фонд (возможно, вместе с другими профсоюзами), а менеджмент может инвестировать в крупную инвентаризацию. Или, наоборот, переговорщики могут публично огласить свои минимальные требования и максимальные предложения, сделав так, что их репутации будет нанесен значительный урон, если они пойдут на уступки. Такая игра предварительного связывания себя обязательствами может представлять «Дилеммой заключенного» или игрой в «Цыпленка» в зависимости от структуры последовательной игры.

Библиографические примечания

Лучшее введение в теорию игр для начинающих дано в книге А. Диксита и С. Скета «Стратегические игры» (Dixit A., Skeath S. Games of Strategy. 2nd ed. New York: Norton, 2004). Для более углубленного изучения я бы предложил «Экономику и теорию игр стратегического взаимодействия» Ф. Вега-Редондо (Vega-Redondo F. Economics and the strategic interaction Theory of Games. Cambridge University Press, 2003). Энциклопедическим исследованием является «Учебник по теории игр и их применению в экономике» Р. Ауманна и С. Харта (Aumann R., Hart S. Handbook of Game Theory with Economic Applications. Vol. 1–3. Amsterdam: North-Holland, 1992, 1994, 2002). Применение этой теории к анализу определенных тем можно найти в книге Дж. Д. Морроу «Теория игр для политологов» (Morrow J. D. Game Theory for Political Scientists. Princeton, NJ: Princeton University Press, 1994) и в книге Д. Бейрда, Р. Гертнера и Р. Пикера «Теория игр и закон» (Baird D., Gertner R., Picker R. Game Theory and the Law. Cambridge, MA: Harvard University Press, 1994). Классическое исследование конвенций — книга Д. Льюиса «Конвенция» (Lewis D. Convention. Cambridge, MA: Harvard University Press, 1969). Эта книга основана на другой классической работе — «Стратегии конфликта» Т. Шеллинга (Шеллинг Т. Стратегия конфликта. М.: ИРИСЭН, 2007), где впервые развивается идея точек координации. Работа Шеллинга стала интуитивным основанием для теории игр второго поколения, получившей формальное развитие в работе Р. Зелтена «Пересмотр концепции идеальности для точек равновесия в экстенсивных играх» (Selten R. Reexamination of the perfectness concept for equilibrium points in extensive games // International Journal of Game Theory. 1975. No. 4. P. 25–55). Касательно различной техники предварительного связывания себя обязательствами в политических играх см. Дж. Фирон «Внутренняя политическая аудитория и эскалация международных споров» (Fearon J. Domestic political audiences and the escalation of international disputes // American Political Science Review. 1994. No. 88. P. 577–592). Относительно его использования в переговорах о заработной плате см. мою книгу «Цемент общества» (The Cement of Society. Cambridge University Press, 1989).

XX. Игры и поведение

НАМЕРЕНИЯ И ПОСЛЕДСТВИЯ

Концептуальная структура теории игр многое проясняет. Помогает ли она *объяснить поведение*? Рассмотрим, как теория игр обосновывает сжигание за собой мостов или кораблей. Такое поведение может быть вызвано стратегическими соображениями, описанными в предыдущей главе, но не только ими. Оксфордский словарь английского языка приводит следующую цитату из книги Э. Р. Берроуза «Тарзан — приемыш обезьяны»: «Поскольку она боялась, что поддастся мольбам этого гиганта, она сожгла за собой мосты». Эту ситуацию нельзя считать примером стратегического рассуждения. Скорее эта женщина боится, что может уступить уговорам, если не лишит себя возможности сделать это. Даже в военной сфере такие нестратегические обоснования могут быть не менее важны, чем стратегические. Командующий может сжечь мосты, опасаясь, что страх перед врагом обратит его солдат в бегство. Он может захотеть отрезать путь к дезертирству *себе самому*, если боится поддаться слабости воли. Командующий А может сжечь мосты или корабли, сигнализируя командующему В, чтобы тот не рассчитывал на возможность его войска обратиться в бегство. Очевидно, именно так рассуждал Кортес, когда, сказав своим морякам (это была правдоподобная ложь), что его флот не в состоянии выйти в море, сжег все корабли, кроме одного (таким образом он мог пополнить пехоту моряками). Чтобы различать эти объяснения, необходимо определить намерения агентов. Реальные выгоды от сжигания мостов никогда не будут ни необходимыми, ни достаточными для того, чтобы дать объяснение с точки зрения ожидаемых выгод (глава III).

Хотя теория игр объясняет поведение, апеллируя к намерениям акторов вызвать определенные последствия, она может описывать ситуации, в которых действующие лица не думают о последствиях. Рассмотрим, например, взаимодействие между Европейским союзом и некоторыми его новыми членами из Восточной Европы. Старые участники могут поддаться соблазну выдвинуть такие условия членства, которые предполагают более низкие дотации сельскому хозяйству для новичков по сравнению со старыми членами такого же размера. В материальном отношении новым странам будет лучше в качестве членов ЕС второго

сорта, чем если их туда не примут вовсе, хотя и хуже, чем если бы они стали полноправными партнерами. В психологическом отношении такое оскорбительное обращение может побудить их отвергнуть эти условия¹. Предчувствуя такую реакцию, старые государства — члены ЕС могут предложить вступление на условиях полного равноправия. Вера в то, что новые страны руководствуются не только материальными интересами, может предоставить самим новичкам лучшие условия.

Поскольку я не был посвящен в ход переговоров о вступлении в Европейский союз, эти заметки являются гипотетическими. Но мы знаем, что аргументы такого рода приводились на Конституционном конвенте в Филадельфии в 1787 году в дебатах по поводу условий присоединения будущих западных штатов. Губернатор Моррис и остальные предлагали, чтобы те были приняты в качестве штатов второго сорта так, чтобы их голоса никогда не смогли отменить голосование первых тринадцати штатов. Оспаривая эти взгляды, Джордж Мейсон активно ратовал за прием новых штатов на равных с остальными условиях. Во-первых, он апеллировал к принципам: принимая западные штаты на равных, они бы сделали то, что, «как нам известно, правильно само по себе». Для тех, кто мог не принять этот аргумент, он добавлял, что новые штаты вряд ли примут такое унижительное предложение:

Если принимать западные штаты в союз, с ними следует обращаться как с равными и не подвергать унижительной дискриминации. У них такая же гордость и другие чувства, как у нас, и они либо не присоединятся к союзу, либо отвернутся от него, если во всех отношениях не будут на равных со своими собратьями.

Мейсон ссылается на «гордость и чувства» новых штатов, а не на их эгоистический интерес. Даже если странам выгоднее войти в союз на неравноправных условиях, чем не вступать в него, они могут из-за обиды остаться вне союза. В то же время он апеллирует к интересу более старых штатов, а не к их чувству справедливости. Если воспользоваться терминологией главы 4, он говорит им, что, поскольку новые штаты могут руководствоваться *чувством*, а не *интересом*, старым штатам выгодно вести себя так, как будто они руководствуются *разумом*, а не *интересом*.

Эта ситуация экспериментально изучалась при помощи игр «Ультиматум» и «Диктатор» (рис. XX.1). В игре «Ультиматум» одно лицо (Пред-

¹ В 2003 году президент Ширак показал пример такого отношения, когда в ответ на выражение восточно-европейскими политиками поддержки американского вторжения в Ирак сказал, что они упустили замечательную возможность промолчать, добавив, что они, очевидно, не получили должного воспитания.



РИС. XX.1

лагающий) может предложить разделить $(x:10-x)$ десять долларов между ним и другим лицом (Респондентом). Предложения принимаются только в целых долларах. Если Респондент принимает предложение, деньги делят. Если нет, никто ничего не получает. Хотя изучали многие варианты этой игры, я сосредоточусь на однократном взаимодействии в условиях анонимности. Поскольку участники взаимодействуют через компьютерные терминалы, они не знают, кто их партнер. Часто им дают понять, что экспериментатор не может определить, каков выбор каждого, тем самым препятствуя тому, чтобы на их решения повлияло желание вызвать его одобрение. Когда участники играют в эту игру многократно, они никогда не встречаются одного и того же партнера, что позволяет обучаться, но не создавать себе репутацию. При таких условиях существуют максимальные возможности для того, чтобы решения отразили ничем не сдерживаемый эгоизм.

Предположив, что оба агента рациональны, эгоистичны и владеют полной информацией о структуре выплат и что эти факты общеизвестны, Предлагающий озвучит $(9:1)$, а Респондент примет предложение. Если бы предложение можно было делать в центах, предлагающий предложил бы $(9,99:0,01)$, что было бы принято, поскольку лучше получить хоть что-то, чем ничего. В экспериментах предложения, как правило, вращаются вокруг $(6:4)$. Обычно Респонденты отвергают предложения, дающие им 2 или меньше². Они готовы назло бабушке отморозить уши. Ясно, что одно из допущений нарушается. Благодаря условиям прове-

² Я говорю «обычно» и «как правило» из-за большой вариативности результатов. Иногда она связана с полом, иногда — с культурными особенностями.

дения эксперимента, мы можем исключить нехватку информации и нехватку общего знания о ней. Однако мы не можем устранить просчеты рациональности или неэгоистические мотивации.

Предлагающий может быть *альтруистом*, предпочитающим примерно равное разделение такому, в котором ему достается все. Хотя альтруизм в отношении совершенно незнакомых людей, которые не испытывают явной нужды, может показаться странным, это по крайней мере не противоречит рациональности. Мы, однако, можем отбросить эту гипотезу, сопоставив поведение в игре «Ультиматум» с поведением в игре «Диктатор». Последняя по сути не является игрой, поскольку в ней Предлагающий в одностороннем порядке делит деньги между собой и Респондентом, не оставляя последнему возможности ответить. Если предложения в игре «Ультиматум» были продиктованы только альтруизмом, распределение ресурсов в «Диктаторе» не должно ничем от них отличаться. Но в экспериментах поведение Предлагающего в игре «Диктатор» оказывается гораздо менее щедрым. Ясно, что поведение Предлагающего в игре «Ультиматум» определяется, по крайней мере частично, ожиданием того, что недостаточно щедрое предложение будет отвергнуто.

Чтобы объяснить этот отказ, мы должны предположить, что Респондентом будет мотивировать *желание* отвергнуть слишком скромные предложения и что эгоистичные Предлагающие, предвосхищая это, будут делать предложения, достаточно щедрые, чтобы их приняли. Если это объяснение верно, следует ожидать, что частота отказа при (8:2) должна быть одинаковой, когда Предлагающий может выдвинуть любое разделение и когда он ограничен выбором (и знает об этом) между (8:2) и (2:8). Как показывают эксперименты, в последнем случае процент отказов меньше. Этот результат свидетельствует о том, что поведение Респондентов определяется соображениями *честности*. Потому что если Предлагающий выставил (8:2), имея возможность предложить (5:5), это рассматривается как менее честное поведение, чем когда его единственная альтернатива была в равной мере для него невыгодна. Именно *намерения*, а не исходы имеют значение.

Такую интерпретацию подтверждает важность взаимности в других играх, например в «Доверительном управлении» (глава XV). Иногда люди готовы наказать других за нечестное поведение, понеся некоторые издержки и не получив никакой выгоды. Такая практика, как представляется, нарушает один из канонов рациональности, перечисленных в главе XII: выбирая между совершением действия и бездействием, рациональный агент не будет действовать, если ожидаемые издержки превышают ожидаемые прибыли. Объяснения с точки зрения альтруизма или зависти не нарушают этот принцип. Для альтруиста исход может быть лучше, если он принесет выгоду другому при некоторых издержках

для самого альтруиста, а для завистливого человека — если он нанесет другому вред при некоторых издержках для него самого. Такое поведение опровергает предположение об эгоистическом интересе, но не о рациональности. Напротив, объяснение с точки зрения честности противоречит допущению о рациональности. Сильная взаимность порождает поведение, подобное тому, как, споткнувшись о камень, мы в гневе пишем его ногой: это не помогает и только усиливает боль³.

ОБРАТНАЯ ИНДУКЦИЯ

В игре «Ультиматум» (рис. XIX.5) и других последовательных играх равновесие находят в обратной индукции. В игре «Ультиматум» Предлагающий предвосхищает реакцию Респондента на его предложение и приспособливает к ней свое поведение. Сопутствующие этим примерам расчеты очень просты. В других экспериментах от участников могут потребоваться более сложные цепочки рассуждений. Например, двое из них проходят три раунда предложений и встречных предложений, чтобы поделить сумму денег, которая в каждом следующем раунде уменьшается на 50%⁴. Каждый раз агент может либо принять предложение и пойти напрямую, либо сделать встречное предложение и снизить ставки. Рациональность, эгоистический интерес и общеизвестность приведут его к следующему рассуждению.

Человек, делающий первое предложение (игрок 1), должен будет учесть, предпочтет ли игрок 2 данное разделение такому, при котором он получит больший кусок от меньшего пирога. В то же время игрок 1 знает, что игрок 2 не сделает предложения, от принятия которого игроку 1 будет хуже, чем если бы он перешел к последнему раунду. На рис. XX.2 игрок 1 может получить по меньшей мере 1,25, забрав все, что осталось в третьем раунде. Таким образом, игрок 2 не может предложить ему менее 1,25 во втором раунде, оставив максимум 1,25 для себя. Зная об этом, игрок 1 предложит (3,75:1,25), и игрок 2 примет предложение.

В экспериментах среднее предложение, сделанное игроком 1, составляет (2,89:2,11), что значительно щедрее равновесного варианта. Очевидно, что нарушается одно или даже несколько допущений. (1) Первый игрок может быть альтруистом. (2) Он может опасаться, что другой игрок отвергнет равновесное предложение, потому что не способен по-

³ На это можно возразить, что, как мы видели в главе XV, люди могут получать удовольствие от наказания других и, более того, наказание может быть мотивировано именно ожиданием удовольствия.

⁴ Уменьшение может рассматриваться как эффект дисконтирования во времени (глава VI).

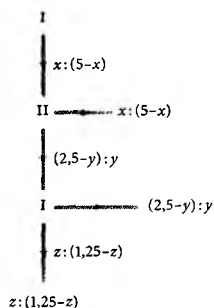


РИС. XX.2

нять логику обратной индукции. (3) Он сам не способен понять эту логику⁵. Первая и вторая гипотезы могут быть устранены, если сказать игрокам, что они играют с компьютером, запрограммированным на оптимальный ответ. В этом случае первое среднее предложение составляет (3,16:1,84), то есть по-прежнему значительно больше равновесного. Поскольку участники, делающие щедрые предложения, едва ли могут испытывать альтруистические чувства к компьютеру или считать его некомпетентным, по-видимому, некомпетентны они сами.

Проблема не в том, что это трудная задача. Как только участникам объясняют логику обратной индукции, в последующих играх они все проделывают безукоризненно. Скорее эксперимент показывает, что такого рода рассуждения людям несвойственны. Даже обычная предусмотрительность не всегда возникает спонтанно, как показывает проклятие победителя (глава XII). Нечто подобное есть и в синдроме младшего брата (глава XVIII). Не то чтобы люди, поразмыслив, не могли понять, что другие так же рациональны и способны к оценке последствий, как и они, просто у них есть стихийная склонность полагать, что другие сильнее укоренены в своих привычках и не способны приспосабливаться к обстоятельствам.

⁵ В отличие от игры «Ультиматум» у игрока I нет причин бояться, что игрок 2 может отвергнуть равновесное предложение как несправедливо низкое, поскольку за игроком I одновременно и первое, и последнее слово.

НЕКОТОРЫЕ СБОИ В ТЕОРИИ ИГР, ОСНОВАННОЙ НА РАЦИОНАЛЬНОМ ВЫБОРЕ

Из множества примеров неудач в предсказаниях, сделанных теорией игр, я хочу остановиться на «Дилемме заключенного» с конечным числом повторений, «парадоксе сетевого магазина», играх «Сороконожка», «Дилемма путешественника» и «Конкурс красоты».

Когда участники многократно играют друг с другом в «Дилемму заключенного» и знают, какой раунд будет последним, мы наблюдаем значительную долю выбора *C*, иногда превышающую 30%. Интуитивно можно предположить, что игрок выберет *C* в одном раунде в надежде на то, что другой ответит ему взаимностью (услуга за услугу). Но если игроки применяют обратную индукцию, то поймут, что в финале оба выберут *D*, поскольку в более поздней игре нет возможности влиять на поведение. В предпоследней игре игроки также выберут *D*, поскольку поведение в финале задается предшествующим рассуждением. Этот аргумент «отыгрывается назад» вплоть до первого раунда, мотивируя отказ от сотрудничества во всех играх.

Сетевой магазин имеет филиалы в 20 городах и в каждом из них сталкивается с потенциальным конкурентом. Конкурент должен решить: открывать ли магазин, чтобы разделить рынок с сетевым игроком, или отказаться от ведения дел в этом городе. Сетевой магазин имеет возможность отреагировать агрессивно, резко снизив цены, тем самым он разорит соперника, но и сам понесет убытки. Или же он может согласиться на раздел рынка. Выигрыши представлены на рис. XX.3; первое число в каждой паре — выигрыш для потенциального конкурента.

Обратная индукция в одной игре дает (5:5) в качестве равновесного исхода: конкурент приходит на рынок, и сетевой магазин соглашаться



РИС. XX.3

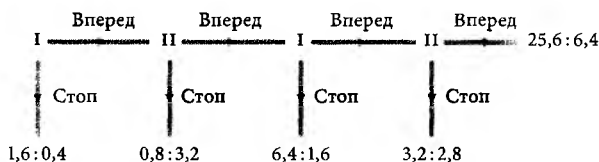


РИС. XX.4

ся на раздел. Но задумавшись о дальнейших проблемах, сетевой магазин может вести более агрессивную политику и уничтожить конкурента с некоторыми издержками для себя, чтобы предотвратить появление потенциальных соперников в других городах. Но если применить обратную индукцию, рассуждая с точки зрения последовательности из 20 игр, такая стратегия окажется нежизнеспособной. В 20-й игре агрессивное поведение уже не несет никаких выгод, так что фирме придется поделить рынок с конкурентом. Но отсюда следует, что нельзя получить выгоды от занижения цен и в 19-й игре и так далее, вплоть до первой. Хотя степень хищнического ценообразования на реальных рынках вызывает споры, на экспериментальных — оно не возникает.

Игра «Сороконожка» (буквально «стоножка». — *Примеч. пер.*)⁶ показана на рис. XX.4 (выплата в долларах). Обратная индукция подсказывает, что игрок I должен выбрать в начале «стоп», оставив каждого из двоих соперников с $\frac{1}{16}$ выигрыша, который они могли бы получить, играя до конца. В одном типичном эксперименте 22% выбрали «стоп» на первом «узле» выбора, 41% из оставшихся выбрали «стоп» на втором узле, после этого половина оставшихся выбрала «стоп» на четвертом узле, а еще половина выбрала «вперед». Отклонение от (кругового) равновесия, предсказываемое обратной индукцией, так же велико, как средний прирост прибыли для игроков.

Чтобы пояснить эти примеры внешне иррациональной кооперации и хищничества, можно предположить существование неопределенности некоторых аспектов игры. В реальной жизни игроки редко сталкиваются с конечным и известным количеством раундов. Часто они полагают, что взаимодействие будет повторяться бесконечное число раз, так что финальный раунд, от которого могла бы оттолкнуться обратная индукция, отсутствует. В таких случаях взаимное использование тактики «услуга

⁶ Это название отсылает к варианту игры со 100 узлами.

за услугу» может стать равновесной в начатой «Дилемме заключенного». (Это равновесие не уникально, поскольку последовательный отказ обоих игроков от сотрудничества тоже представляет собой равновесие. Структурно это напоминает игру на доверие с одним хорошим равновесием и одним плохим.) Если реальная жизнь склоняет к поведению по модели «услуга за услугу», агенты могут применить его и в ситуации эксперимента, для которых оно не является оптимальным.

Вместе с тем агент может быть неуверен в том, с каким типом игрока он имеет дело. Предположим, общеизвестно, что в группе есть некоторые иррациональные индивиды. Известно, что некоторые агенты всегда будут сотрудничать, другие будут использовать стратегию «услуга за услугу» в количественно ограниченных «Дилеммах заключенного», а третьи будут демпинговать, чтобы отпугнуть конкурентов даже в двадцатом городе, и так далее. Однако неизвестно наверняка, кто эти индивиды. Каждый агент с некоторой положительной вероятностью может быть иррациональным. В «Парадоксе сетевого магазина» потенциального конкурента можно отпугнуть, заставив приписать достаточно большую вероятность тому, что менеджер сетевого магазина иррационален. Когда потенциальные конкуренты в других городах увидят подобное поведение, они воспользуются байесовым рассуждением (глава XI), приписав его иррациональности более высокую вероятность. Она может оказаться недостаточно высокой, чтобы их отпугнуть, но если он снова и снова будет вести себя иррационально, то, возможно, достигнет уровня, при котором для них будет рациональнее остаться за пределами рынка. Подобное рассуждение может объяснить сотрудничество в количественно ограниченной «Дилемме заключенного» и игре «Сороконожка».

Еще одна возможность заключается в том, что в многократно повторяемых «Дилемме заключенного» и «Сороконожке» кооперация имеет некоторые свойства точки координации. Хотя *рациональные* индивиды откажутся от сотрудничества в первом же раунде, *разумные* не сделают этого. Несмотря на некоторую неопределенность, это предположение (подробнее я обращусь к нему позднее) более верно, чем рассуждения, основанные на неопределенности типа другого игрока. С одной стороны, такие рассуждения требуют от игроков выполнения невероятно сложных расчетов, занимающих в учебниках многие страницы. С другой — интроспекция и бытовые наблюдения подсказывают, что, принимая решения в повседневной жизни, мы так о других не думаем. Когда я кому-то доверяю небольшую сумму денег, но не доверяю большую, это происходит не потому, что я сомневаюсь в безусловной надежности партнера, но потому что полагаю, что ему можно доверять, только когда ставки не очень высоки.

В «Дилемме путешественника» два игрока одновременно требуют от 80 до 200 долларов за потерянный багаж. Чтобы избежать чрезмерных требований, авиакомпания выплачивает каждому пассажиру минимальную из двух заявленных сумм, доплачивая сумму R человеку, выдвинувшему самое низкое требование, и вычитая такую же у человека с более высоким требованием. Рассмотрим пару требований (100:150), дающую выигрыш $((100 + R) : (100 - R))$. Эта пара не может быть равновесной, поскольку у первого игрока есть стимул требовать 149, тем самым получив выигрыш $149 + R$, на что второй игрок ответит требованием 148 и так далее. Как подсказывает этот пример, такой исход в действительности наблюдается, когда сумма R высока. Когда же она невысока, участники могут запрашивать суммы, близкие к верхней границе 200. И снова я предполагаю, что здесь работает что-то вроде точки координации. Каждый пассажир знает, что, учитывая выгоды координации вокруг высокого требования, глупо применять стратегию равновесия, и предполагает, что другие тоже это знают.

Джон Мейнард Кейнс сравнивал биржевой рынок с «конкурсом красоты». Он имел в виду популярные в те времена в Англии конкурсы, в которых газета печатала 100 фотографий, а люди писали, какие 6 лиц понравились им больше всего. Каждый, кто выбирал самое популярное лицо, автоматически становился участником лотереи, в которой мог выиграть приз. Кейнс писал: «Вопрос не в том, чтобы выбрать те [лица], которые по самому здравому рассуждению являются самыми красивыми, и даже не те, которые среднестатистическое мнение искренне сочтет самыми красивыми. Мы достигли третьей степени, когда используем свой ум, чтобы угадать представления этого среднестатистического мнения о самом среднестатистическом мнении. И, я полагаю, есть те, кто практикует четвертую, пятую и более высокие степени».

В игре, возникшей из этого замечания Кейнса, участников просят выбрать числа между 0 и 100. Игрок, чье число ближе всего к двум третям от среднего из всех выбранных чисел, получает фиксированный приз. Среднее должно быть меньше или равно 100, соответственно две трети от него должны быть меньше или равны 67. Следовательно, какое бы среднее число ни стало результатом выбора других игроков, 67 будет ближе к двум третям от этого среднего, чем любая большая цифра. Но когда коридор ограничен значениями, меньшими или равными 67, две трети от среднего меньше или равны 44, и так далее, пока не будет достигнуто уникальное равновесие — 0. Лишь немногие участники экспериментов выбирали 0; среднее число вращалось вокруг 35. Выбравший это число должен думать, что большинство других выберет большие числа. Здесь проявляется синдром младшего брата. Тот факт, что это число представляет собой примерно две трети от среднего в це-

лом ряду, 50, подсказывает, что типичный участник может считать, что другие выбирают числа произвольно, тогда как он имеет возможность оптимизировать выбор. Или, напротив, типичный участник может полагать, что другие проходят через два раунда исключения, оставляя ему свободу оптимизации, добавляя третий раунд.

Я хочу сказать, что, когда поведение людей расходится с предсказаниями теории игр, это, возможно, происходит, потому что люди *менее чем рациональны* или *более чем рациональны*. Синдром младшего брата, конечно же, является провалом рациональности, равно как неспособность проделать простую обратную индукцию. Быть рациональным — значит преодолевать ловушки рациональности, сосредоточиваться на факте, что оба игрока могут выиграть, игнорируя логику лучшей реакции. Как я отмечал, последняя идея напоминает точки координации, но лишь отчасти. Точки координации являются равновесием, тогда как кооперация в «Дилемме заключенного» с конечным числом повторений, высокое требование в «Дилемме путешественника» или выбор «вперед» в игре «Сороконожка» таковыми не являются. То общее, что есть у этих вариантов с выбором точки координации, с трудом поддается определению и сильно зависит от контекста, свойства очевидности и разумности.

Аргумент такого рода может показаться принадлежащим скорее к магическому мышлению (глава VII), чем к рассуждению, основанному на точке координации. Игнорировать зов сирен рациональности — значит следовать завету Джона Донна, приведенному в «Годовщине»:

Кто в большей безопасности, чем мы? Никто
Предать нас, кроме нас двоих, не может
Уйдем же страхи, ложные иль нет.

Игнорирование *настоящих* страхов, возможно, иррационально или является примером магического мышления (то же относится к игнорированию реальных перспектив получения прибыли). Или же наоборот (и именно так я предпочитаю рассматривать это поведение), оно отражает более высокие стандарты, чем просто рациональность⁷. Это сложные вопросы, и я призываю читателей подумать над ними самостоятельно. Обсуждение некоторых из них будет продолжено в следующей главе.

}

⁷ Я отсылаю не к идее разума (глава IV) в смысле непредвзятой мотивации, а в некоторой степени к более повседневному понятию. Идея, которая есть у меня в голове (в зачаточном состоянии), связана с тем фактом, что люди чаще сотрудничают в «Дилемме заключенного», когда ее описывают как игру «Сообщество», а не как игру «Уолл-стрит».

Библиографические примечания

Источником большинства примеров в этой главе является «Поведенческая теория игр» К. Камерера (*Camerer S. Behavioral Game Theory*. New York: Russell Sage, 2004). Полезный анализ условий, при которых терпят неудачу предсказания стандартной теории игр, есть в работе Дж. К. Гоэри и Ч. А. Холта «Десять маленьких сокровищ теории игр и десять интуитивных противоречий» (*Goeree J. K., Holt C. A. Ten little treasures of game theory and ten intuitive contradictions // American Economic Review*. 2001. No. 91. P. 1402–1422). Как оказалось, внешне простая идея обратной индукции заключает в себе глубокие парадоксы; некоторые из них намечены во введении к моей книге «Цемент общества» (*The Cement of Society*. Cambridge University Press, 1989). Игра, приведенная на рис. XX.2, взята из работы Э. Джонсона и др. «Выявление неудач обратной индукции» (*Detecting failures of backward induction // Journal of Economic Theory*. 2002. No. 104. P. 16–47). «Дилемма путешественника» взята из статьи К. Басю «Дилемма путешественника: парадоксы рациональности в теории игр» (*K. Basu. The traveler's dilemma: Paradoxes of rationality in game theory // American Economic Review: Papers and Proceedings*. 1994. No. 84. P. 391–395).

XXI. Доверие

СНИЖАЯ БДИТЕЛЬНОСТЬ

Токвиль говорил, что эгоизм — «ржавчина общества». По аналогии часто утверждают, что доверие — «смазка общества»¹. Повседневная жизнь стала бы намного труднее, если бы мы не доверяли другим то, что они, по их заверениям, будут делать, по крайней мере в некоторой степени. Хотя ученые по-разному определяли доверие, я буду использовать бихевиоральное определение: доверять кому-то — это снижать бдительность, *воздерживаться от принятия мер предосторожности против партнера по взаимодействию*, даже если другой человек в силу оппортунизма или некомпетентности может действовать так, что эти меры предосторожности могли бы быть оправданными². Под оппортунизмом я понимаю близорукий или грубый эгоистический интерес, не сдерживаемый ни соображениями этики, ни благоразумием. Типичные оппортунистические поступки, которые могут оправдать меры предосторожности со стороны других, — вранье, жульничество на экзамене, нарушение обещания, растрата денег, измена супругу или выбор некооперативной стратегии в «Дилемме заключенного».

Можно доверять или не доверять себе, когда требуется не нарушить условия сделки, воздерживаться от потребления алкоголя или идти твердым курсом, услышав зов сирен. Недоверие к себе обнаруживается в связывании себя обязательствами или в установлении частных правил (глава XIII). Но эти стратегии могут дорого обойтись, потому что они подают сигнал. Если другие увидят пример таких предосторожностей в отношении моих будущих я, они могут ошибочно заключить, как мы видели в главе X, что я себя не контролирую. Соответственно они будут неохотно доверять мне в тех случаях, когда (1) отсутствие самообладания может дорого мне обойтись, (2) отсутствует возможность для предварительного связывания себя обязательствами и (3) частные пра-

¹ В более старой литературе по экономике роль смазки часть приписывалась коррупции.

² Понятое таким образом доверие подразумевает *двойное воздержание*, одна сторона воздерживается от принятия мер предосторожности в надежде на то, что другая воздержится от оппортунистического поведения.

вила могут быть иррелевантны в рамках однократного взаимодействия. Во многих обществах существуют нормы, направленные как против полного воздержания от алкоголя, так и против пьянства (глава XXII).

Недоверие может принимать одну из двух форм. С одной стороны, можно воздерживаться от взаимодействия с потенциальным партнером, когда интеракция ставит человека в положение, уязвимое для некомпетентности или оппортунизма. С другой — человек может участвовать в интеракции, но принять меры против этих рисков. Таким образом, доверие является результатом двух последовательных решений — участия в интеракции и воздерживания от контроля над партнером. Поскольку решение воздержаться от интеракции трудно зафиксировать при наблюдении, степень недоверия в обществе легко недооценить. Может показаться, что недоверия больше в обществе, где все постоянно следят друг за другом, чем там, где люди больше заняты собой. Но при ближайшем рассмотрении выяснится, что последнее — крайне неэффективно в силу большого числа нереализованных взаимовыгодных соглашений.

Монтень приводил пример одной доверительной реакции: «Когда я путешествую, у кого бы ни был мой кошелек, он находится в его полном распоряжении безо всякого надзора». Другие примеры демонстрации доверия предполагают воздержание от совершения следующих действий:

- читать дневник супруга;
- использовать надзирателей для слежки за студентами на экзаменах;
- проверять рекомендательные письма будущего сотрудника;
- требовать залог у арендатора;
- настоятельно требовать оформленные в письменном виде и имеющие юридическую силу контракты;
- требовать от менее состоятельного партнера подписать брачный договор;
- прятать деньги от собственных детей;
- запирасть наружную дверь, уходя из дома;
- связывать себя предварительным обязательством наказывать уклонившихся от сотрудничества в «Дилемме заключенного»;
- обращаться за проверкой диагноза к другому врачу или за повторной диагностикой автомобиля к другому автомеханику.

Как отмечалось, объектом доверия могут быть *способности* других людей или их *мотивация*. Яркий пример такого различия можно найти в истории движения Соппротивления. В странах, оккупированных Гер-

маний во время Второй мировой войны, члены Соппротивления время от времени погибали, потому что их принимали за немецких агентов. Случалось, хотя и реже, что их убивали, потому что не верили, что они будут держать язык за зубами. Человек мог оказаться пьяницей и быть ликвидирован Соппротивлением во избежание раскрытия опасной информации в пьяном виде. Более приземленный пример: я могу поставить под вопрос профессиональные навыки автомеханика или его честность. Я могу обратиться за дополнительной консультацией к другому врачу из-за сомнений в компетентности первого, хотя меня может беспокоить возможность назначения ненужной хирургической операции, чтобы набить себе карман. В дальнейшем я буду затрагивать главным образом вопросы честности.

ОСНОВАНИЯ ДЛЯ ДОВЕРИЯ

Есть несколько причин, по которым люди могут воздерживаться от принятия мер предосторожности³. (1) Затраты на принятие мер предосторожности могут превосходить ожидаемую выгоду как в отдельно взятом случае, так и в течение всей жизни. Если в моей деревне есть автомеханик и мне придется ехать восемьдесят километров на такси, чтобы проконсультироваться с другим, этого делать не стоит. Говоря обобщенно, жизнь слишком коротка, чтобы постоянно бояться, что вас могут использовать. Случайные потери, от того что вы доверились вероломному человеку, невелики в сравнении с душевным покоем, достигаемым благодаря отсутствию тревог. (2) Само принятие мер предосторожности может дать информацию, которой могут воспользоваться оппортунисты. Монтень цитирует латинское изречение «*Furem signata sollicitant Aperta effractarius praeterit*» («Двери на запоре привлекают вора; открытыми взломщик пренебрегает»). (3) Идея принятия мер предосторожности может противоречить эмоциональному отношению агента к другому человеку. Влюбленные люди могут отказаться от холодного расчета, который предполагает брачный контракт. Стихи Донна, процитированные в главе XX, тоже уместны в данном контексте: «Уйдем же страхи, ложные иль нет». (4) У меня могут быть исходные убеждения относи-

³ Я использую слово «воздержание» для указания на его осознанность. В некоторых случаях, которые я затрагиваю, принятие мер предосторожности, например чтение дневника супруга, может никогда не прийти человеку в голову. Но это не то «слепое доверие», как я определяю его позднее, если агент имел возможность принять меры предосторожности. Для другого человека тот факт, что агент имел возможность, но не воспользовался ею, говорит сам за себя, независимо от того, воспринимается ли такое воздержание как сознательное или нет.

тельно надежности этого человека. (5) Я могу пытаться *вызвать* доверие, доверившись сам.

В дальнейшем я остановлюсь на (4) и, вкратце, на (5) пунктах. Хотя многие ученые определяют доверие в категориях (4), я полагаю, что фокус на сознательном воздержании имеет преимущество в подчеркивании *взаимодействия* между доверителем и доверенным лицом. Если доверенное лицо *замечает* отсутствие мер предосторожности, это может подвигнуть его действовать иначе, чем при иных обстоятельствах. В случае (2) это происходит потому, что оно делает вывод, что *случая* для оппортунистического поведения не представится. В других обстоятельствах, к которым я обращаюсь позднее, подобное наблюдение может изменить его *мотивацию* вести себя оппортунистически, тем самым отражая догадку, возникающую до всякого анализа, что доверие обладает перформативностью. То же относится к недоверию. Как заметил Пруст: «Как только ревность раскрыта, ее объект начинает считать ее недостатком доверия, дающим ему право обманывать нас».

ОСНОВАНИЯ ДЛЯ НАДЕЖНОСТИ

Люди могут восприниматься как *заслуживающие доверия* по нескольким основаниям. Я коснусь четырех из них: прошлое поведение, стимулы, знаки и сигналы. Часто мы знаем — или полагаем, что знаем (глава 10) — из наблюдений за другими людьми, что они постоянно выполняют свои обещания, не лгут, осторожно обращаются с чужой собственностью и так далее. Более того, человек, который знает, что он или она (не) достойны доверия, будет руководствоваться убеждением, что другие тоже (не) достойны доверия (так называемый эффект ложного консенсуса), будет склонен им (не) доверять. Как сказал Лабрюер: «Подлецы часто считают других столь же дурными, как они сами; их не обмануть, и они тоже долго не обманывают»⁴. Существует экспериментальное подтверждение того, что этот механизм действительно работает. Напротив, А может доверять С, потому что знает, что В, которому он доверяет, тоже доверяет С. Этот вывод, однако, может оказаться неверным, потому что доверие В по отношению к С может быть вызвано эффектом ложного консенсуса. Как показывают эти примеры, мы часто доверяем или не доверяем людям ошибочно, полагая, что другие будут либо больше похожи на нас, либо «больше похожи на себя» (более последовательны в своем поведении, чем на самом деле).

⁴ Ларошфуко думал иначе: «Люди никогда так легко не обманываются, как когда они сами готовы обмануть других».

В небольшом международном сообществе торговцев алмазами, где соблазн оппортунистического поведения крайне велик, соглашение, достигнутое на словах без свидетелей, обладает не меньшей силой, чем письменный контракт. Торговец, нарушивший соглашение, может положить в карман временную прибыль, но от него *навсегда* отвернутся все остальные торговцы⁵. Кроме того, он не сможет передать бизнес детям, как это часто заведено в сфере торговли алмазами. В случае нью-йоркских торговцев из ультраортодоксальной еврейской среды обманщик будет подвергнут социальному ostracismu. Последний механизм укрепляет надежность, но не является для нее необходимым. Зачастую достаточно стимула поддерживать репутацию честного и надежного человека.

Знаки — это *особенности* индивидов, которые справедливо или ошибочно считаются признаками надежности. Исследование особенностей, заставляющих водителей такси поверить, что пассажиры их не ограбят и не нападут на них, показало, что женщины воспринимаются как заслуживающие большего доверия, чем мужчины, люди постарше — чем более молодые, белые — больше, чем черные, богатые — больше, чем бедные, погруженные в себя — больше, чем излишне любопытные, чистосердечные — более, чем изворотливые. Испаноговорящий водитель в Нью-Йорке сочтет испаноговорящих пассажиров более надежными, чем тех, кто принадлежит к другим этническим группам. Водители-католики в Белфасте сочтут пассажиров-католиков более надежными, чем пассажиры-протестанты, и наоборот. Более общая черта — не слишком близко посаженные глаза и привычка прямо смотреть на собеседника.

Сигналы — это *поведение*, свидетельствующее о надежности. Они могут включать в себя сознательное производство или имитацию знаков. Например, считается, что внимательно смотреть на переносицу собеседника — хороший способ продемонстрировать искренность. В этом случае сигнал будет работать, только если другой человек уверен, что искренний взгляд — надежный индикатор поведения, и не знает, как просто его подделать. Некоторые формы поведения играют роль сигнала, если они слишком дороги, чтобы ненадежные индивиды могли себе их позволить. Для успешной подделки подписи может потребоваться большая практика, тогда как подписать что-то своей рукой очень просто. Бедняк может одеваться как банкир с Уолл-стрит, чтобы вызвать

⁵ Это не просто механизм «услуга за услугу», в котором игрок, который отказывается сотрудничать в одном раунде и наказан в следующем, может быть прощен, если снова будет сотрудничать. Скорее это «стратегия кары», в которой единичный случай отказа от кооперации исключает заглаживание вины при хорошем поведении в дальнейшем.

доверие таксиста, но вероятность такого поведения невелика, поскольку затраты выше, чем прибыль, которую он может получить от ограбления. И наоборот, каждый может размахивать «Уолл-стрит джорнал», чтобы подозвать такси, так что это не может служить основанием для различения надежных и ненадежных клиентов. В той степени, в какой доверие основывается на убежденности в том, что у партнера по интеракции долгий временной горизонт (низкий коэффициент дисконтирования), затратное демонстрирование хорошей физической формы и стройности может служить сигналом, учитывая (ложное) убеждение, что прозорливость как черта характера действует или всегда, или никогда.

Часто мы доверяем людям, потому что ощущаем, что они мотивированы не только своекорыстным интересом. Иногда же мы доверяем людям, только если видим их эгоистические мотивы. В «Мальтийском соколе» мистер Гатмен говорит Хамфри Богарту: «Я не доверяю человеку, который не следит за собой». Наполеон считал, что Талейрану нельзя доверять, потому что он никогда ничего не просил для своей семьи. Говорят, что Франсуа Миттеран, будучи президентом, так же не доверял тем, кто никогда не просил его об одолжениях. Говоря шире, главная проблема мошенников, пользующихся чужим доверием, в том, чтобы убедить своих жертв, что они действуют из личного интереса. Предположим, я подхожу к кому-нибудь и заявляю, что можно сделать целое состояние, инвестировав небольшую сумму денег прямо сейчас. Его первым вопросом будет: «Не слишком ли это хорошо, чтобы быть правдой?» А вторым вопросом будет: «Если это действительно так, почему вы хотите поделиться этим со мной, а не забрать все себе?» Успешный жулик способен вызвать доверие жертвы, рассказав правдоподобную историю о том, почему в его интересах поделиться частью прибыли, без которой претензии на великодушие недостоверны.

Подобно тому, как люди могут быть (восприниматься) более или менее надежными, они могут быть в большей или меньшей степени *доверчивыми*. Иными словами, если *A* и *B* имеют одни и те же представления о *C* (или никаких представлений), *A* может доверять *C*, а *B* — нет. Склонность доверять другим особенно важна при начале совместных предприятий. В повторяющихся интеракциях кооперация может поддерживаться взаимностью, за исключением *первого раунда*, когда еще нет предшествующей истории совместных взаимодействий. Чтобы ее создать, стороны вынуждены сотрудничать в первом раунде без условий. Доверяющий индивид будет следовать стратегии «как аукнется, так и откликнется»: сотрудничай в первом раунде и отвечай взаимностью во всех последующих. Как гласит пословица: «Обманешь меня раз — позор тебе. Обманешь меня дважды — позор мне». Недоверчивый человек

будет следовать стратегии «как откликнется, так и аукнется»: откажись от сотрудничества в первом раунде и отвечай взаимностью в последующих. Для чтобы сотрудничество сдвинулось с мертвой точки, необходимо взаимодействие двух доверяющих друг другу индивидов.

КАК ДОВЕРИЕ МОЖЕТ ПОРОЖДАТЬ НАДЕЖНОСТЬ

Оказываемое доверие может порождать надежность (кредитоспособность), когда другая сторона *знает*, что человек воздержался от принятия мер предосторожности. Монтень писал о слуге, которому доверял свой кошелек: «Он мог бы надуть меня, даже если бы я следил за счетами, так что, если он не сам дьявол, таким безрассудным доверием я обязываю его быть честным». Такая форма взаимности отличается от той, что воплощена в стратегии «услуга за услугу» в «Дилемме заключенного». Услуга за услугу может быть стратегией равновесия, когда игра разыгрывается неограниченное количество раз, и два выбора в любой конкретной игре делаются одновременно. В свою очередь замечание Монтеня относится к однократной игре, в которой одна сторона делает выбор и доводит его до сведения другой прежде, чем та успеет его сделать. Эксперименты по изучению склонности к сотрудничеству в однократных «Дилеммах заключенного» подтверждают эту загадку. В одном исследовании участникам сообщали выбор другого игрока прежде, чем они делали свой выбор. Когда игроки знали, что другой игрок отказался от сотрудничества, сотрудничать соглашались только 3%. Когда они знали, что другой тоже сотрудничал, сотрудничать соглашались 16%.

Можно привести еще один пример, чтобы показать, насколько тонким механизмом является доверие. Это игра «Доверительное управление», разыгрываемая в двух вариантах. В обоих случаях «инвестор» имеет возможность сделать перевод в размере от 0 до 10 из 10 денежных единиц в адрес «доверительного собственника». Затем любая посланная сумма утраивается: если инвестор отправил 10, его партнер получит 30. Доверительный собственник может сделать любой перевод в размере от 0 до всей возросшей суммы (30 в том же примере) обратно инвестору. Наконец, если инвестор решает сделать трансфер, ему нужно указать размер обратного трансфера, который он хочет получить от партнера.

Эти черты определяют не слишком удачно названные (так я утверждаю) «условия доверия». В «условиях стимула» у инвестора также есть возможность заявить в момент совершения трансфера и объявления желательной суммы возврата, что он наложит на доверительного собственника штраф в 4 денежных единицы, если обратный трансфер будет меньше затребованного. Некоторые инвесторы пользуются этой возможностью, другие нет. Если они ее не используют, партнер узнаёт, что

у инвестора была такая возможность, но он ею пренебрег. Результаты эксперимента показывают, что самые крупные обратные трансферы делаются в условиях стимула, когда нет штрафа, а самые маленькие при угрозе штрафа находятся на промежуточном уровне. Такой эффект был предугадан инвесторами, которые вкладывали примерно на 30% больше в ситуации «стимул без штрафа», чем в любом другом случае.

Модель «стимул без штрафа» соответствует моему определению доверия. То, что в экспериментах называлось «условиями доверия», я бы скорее назвал *слепым доверием*. Оно проявляется, когда предосторожности *исключаются*, в отличие от ситуаций, когда их *не выбирают*. Поразительное открытие состоит в том, что (не слепое) доверие порождает большую кооперацию, чем слепое. Различие состоит в снижении бдительности.

Люди могут доверять или не доверять *институтам*. Они могут верить, что банк не обанкротится, или держать свои сбережения под матрасом. Они могут верить в беспристрастность правосудия при решении спора между соседями или предпочесть добиваться справедливости самостоятельно. Строго говоря, эти случаи не соответствуют определению, данному вначале, поскольку граждане не могут принять никаких (или почти никаких) других мер предосторожности, имея дело с институциями, кроме как отказаться иметь с ними дело. Когда я доверяю своему банку — это форма слепого доверия. В незначительной степени мое доверие влияет на вероятность того, что он не обанкротится, но только потому, что мой вклад в его резервы делает его немного менее уязвимым. Напротив, доверие, оказываемое Монтенем слуге, создавало уверенность благодаря стыду или чувству вины, которые тот мог испытать, не оправдав его.

Но, как утверждал Монтень, эти благотворные последствия доверия являются преимущественно побочными продуктами (глава IV). Если вы не станете принимать меры предосторожности с единственной целью заставить других хорошо себя вести, вы едва ли добьетесь успеха. «Отличный способ завоевать сердце и душу другого человека — пойти и довериться ему, отдав себя в его власть — при условии, что это делается добровольно, не вызвано необходимостью и что доверие оказывается с ясностью и чистотой или, по крайней мере, с сознанием, свободным от сомнений».

Библиографические примечания

Свидетельства того, что надежны те, кто доверяет другим, приводятся в Д. Глейзер и др. «Измеряя доверие» (D. Glaeser et al. Measuring trust// Quarterly Journal of Economics. 2000. No. 115. P. 811–846). Сообщество торговцев алмазами в Нью-Йорке анализируется в работе Б. Ричмена «Общинные механизмы принуждения посредством неформальных

обязательств: еврейские торговцы алмазами в Нью-Йорке» (Richman B. Community enforcement of informal contracts: Jewish diamond merchants in New York. Working paper. Harvard University, Olin Center for Law and Economics, 2002). Использование таксистами знаков и сигналов для определения надежности пассажиров — тема работы Д. Гамбетты и Х. Хэмилла «Уличная мудрость» (*Gambetta D., Hamill H. Streetwise. New York: Russell Sage, 2005*). Способы сделать так, чтобы мошенничество показалось заслуживающим доверия, приводятся в забытой книге Н. Леффа «Мошенничество и продажа» (*Leff N. Swindling and Selling. New York: Free Press, 1976*). Эксперимент с сотрудничеством в однократной «Дилемме заключенного» описывается в работе Е. Шафиро и А. Тверски «С точки зрения неопределенности: неконсеквенциалистское рассуждение и выбор» (*Shafir E., Tversky A. Thinking through uncertainty: Nonconsequentialist reasoning and choice // Cognitive Psychology. 1992. No. 24. P. 449–474*). Результаты игры «Доверительное управление», описанные в тексте, приводятся в статье Е. Фера и Б. Рокенбаха «Пазубное влияние санкций на человеческий альтруизм» (*Fehr E., Rockenbach B. Detrimental effects of sanctions on human altruism // Nature. 2003. No. 422. P. 137–140*).

XXII. Социальные нормы

КОЛЛЕКТИВНОЕ СОЗНАНИЕ

Социологи иногда ссылаются на «коллективное сознание» какого-то общества, набор ценностей и верований, разделяемый (фактически или предположительно) его членами. С точки зрения ценностей, коллективное сознание включает моральные и социальные нормы, религию и политические идеологии. С точки зрения убеждений, оно содержит мнения о фактуальных вопросах, а также о каузальных отношениях в диапазоне от слухов о торговле живым товаром до убежденности в пагубных последствиях льгот для безработных. Ниже я рассматриваю социальные нормы и их действие. В следующей главе я обращусь к способам коллективного, или, точнее говоря, интерактивного формирования убеждений. В моем анализе ценностей и убеждений есть двойная асимметрия. С одной стороны, я немного могу сказать о возникновении социальных норм, не потому, что это неинтересный вопрос, а потому что мне сложно об этом судить. С другой — мне почти нечего сказать о сущности народных или коллективных верований. Их содержание значительно варьируется во времени и в пространстве, тогда как механизмы возникновения, распространения, преобразования и распада более постоянны.

ДЕЙСТВИЕ СОЦИАЛЬНЫХ НОРМ

Рассмотрим два высказывания:

Всегда надевайте черную одежду в жаркую погоду.
Всегда надевайте черную одежду на похороны.

Первая рекомендация — вопрос инструментальной рациональности, поскольку воздух между телом и одеждой циркулирует быстрее, если одежда черного цвета. Второе высказывание выражает социальную норму, которая не имеет очевидного инструментального применения. Существование и важность социальных норм невозможно поставить под сомнение. Непосредственные причины их действия вполне доступны разумному пониманию. Но их происхождение и функция (если таковая имеется) остаются дискуссионными.

Социальная норма — это побуждение действовать или воздерживаться от действия. Некоторые нормы являются безусловными: «Делай X, не делай Y»¹. Они включают нормы: не потреблять человеческую плоть, не вступать в сексуальные отношения с братом или сестрой, не нарушать очереди, не носить красного (как некоторые матери говорят своим дочерям), надевать черную одежду на похороны, начинать есть с края и двигаться к центру тарелки, лечить первым самого больного пациента. Другие нормы являются условными: «Если сделаете X, тогда сделайте Y» или «Если другие делают X, делайте X». Во многих группах существует норма, что человек, первым предложивший какое-то дело, становится за него ответственным². В результате многие хорошие предложения не выдвигаются. Бездетная пара может считать для себя нормой, что первым предложивший завести ребенка должен будет впоследствии принимать большее участие в его воспитании, поэтому некоторые пары, которые хотели бы иметь детей, могут так и остаться бездетными³. Может не существовать нормы, заставляющей меня посылать рождественские открытки двоюродным братьям и сестрам, но как только я начинаю это делать, в действие вступает норма, согласно которой я не могу так просто это бросить, и еще одна, побуждающая кузенов ответить тем же. При всей условности эти нормы являются таковыми не для любого исхода, воплощаемого в действии, в отличие от предписания носить черное в жаркую погоду.

Ниже приведены другие примеры. Однако сначала я скажу о том, что придает социальным нормам каузальную эффективность, и о том, чем они отличаются от других норм. Ответ состоит в том, что социальные нормы работают через неформальные санкции, направленные на их нарушителей. Как правило, санкции оказывают воздействие на материальное положение нарушителя либо через механизм прямого наказания, либо через ограничение возможностей посредством социального ostracism. Фермер, нарушающий общинные нормы, может столкнуться с тем, что его амбар сожгут, а овец зарежут. Или же сосед откажется помогать убирать урожай. Механизм *сплетен* может работать в качестве мультипликатора, добавляя санкции со стороны третьих лиц к исходному наказанию второй стороны.

¹ В дальнейшем определения «условный» и «безусловный» будут относиться к содержанию социальной нормы. Как отмечалось в главе V, все социальные нормы условны в том смысле, что их действие зависит от присутствия наблюдателя.

² Эта норма может быть отнесена к рассуждениям по принципу точки координат.

³ То есть норма порождает игру в «Цыпленка».

Рассмотрим, что может сделать владелец животноводческой фермы, когда скот соседа несколько раз нарушил границу его владений. Он может забрать этот скот с выгодой для себя и ущербом для соседа. Он может его уничтожить или уменьшить его ценность (например, кастрировать быка) безо всякой личной выгоды и в ущерб соседу. Он может отогнать скот в какое-то отдаленное место с некоторыми издержками и для себя и для соседа. Или же он может разорвать отношения с соседом (ostracism). Последняя реакция, однако, может оказаться неэффективной, поскольку скот сможет и дальше заходить в его владения. Первая реакция может выглядеть агрессивным захватом, а не наказанием. Второй, а особенно третий ответ более адекватны, поскольку ясно указывают на намерение наказать, даже с некоторыми издержками для наказывающего.

В целом, однако, я полагаю, что ostracism или бойкот — наиболее важная реакция на нарушение норм. Если бы не скот несколько раз нарушал границы собственности, а сам сосед однажды нарушил бы обещание, более естественной реакцией стал бы разрыв отношений. Это утверждение основано на той общей идее, что социальные нормы работают посредством эмоций стыда в нарушителе и презрения в том, кто стал свидетелем нарушения (глава VIII). Поскольку характерная для презрения тенденция к действию заключается в том, чтобы избегать его объект, из-за чего человек, подвергнутый ostracismу, часто несет материальные потери, существует связь между эмоциональной реакцией и применением санкций. Тем не менее санкции гораздо важнее не сами по себе, а как средство трансляции эмоций. Более того, цена, которую платит *применяющий санкцию* может оказаться особенно важной для выражения силы эмоции.

Теория социальных норм как санкций сталкивается с очевидной проблемой: что мотивирует иницилирующую наказание сторону? Зачем ей это? Как правило, применение санкций является дорогостоящим или рискованным делом для тех, кто их применяет. Даже если он не отказывается от взаимовыгодной интеракции, выражение осуждения может вызвать гнев или даже акт насилия со стороны того, на кого оно направлено. Здесь есть важное различие между спонтанным неодобрением и сознательным посрамлением. Последнее может привести к обратному эффекту, вызвав у объекта скорее гнев, чем стыд. Даже если порицание действительно спонтанно, его объект в своих интересах может истолковать его как преднамеренное желание пристыдить и реагировать соответственно. По этой причине санкции — рискованное занятие. Тогда почему же люди к ним прибегают? Один из возможных ответов: не применяющие наказание сами рискуют быть наказанными. Такое случается. В обществе со строгими нормами мести можно ожидать,

что того, кто не будет сторониться человека, не сумевшего отомстить, также начнут избегать. В школе ребенок может быть больше расположен к общению с «ботаником», когда этого не видят другие. Но ребенок, не присоединившийся к толпе, издевающейся на тем, кто дружит с «ботаником», едва ли сам подвергнется нападкам. Таким образом, маловероятно, что участвующие в моббинге третьи лица мотивированы страхом наказания. Для изучения вопроса можно провести эксперимент и посмотреть, станут ли третьи лица наказывать респондентов, которые, приняв очень низкие предложения в игре «Ультиматум», не сумели наказать скупых предлагающих. Я бы удивился, если бы они стали это делать, и удивился бы еще больше, если бы четвертые лица, наблюдатели, стали бы наказывать тех, кто не наказал третьих лиц. На некотором отдалении от изначального нарушения этот механизм перестает быть достоверным. Более экономное и адекватное объяснение санкций основывается на спонтанной стимуляции презрения и связанной с ним тенденции к действию. Здесь также может присутствовать гнев в силу плавающего различия между социальными и моральными нормами. Кроме того, бравирование нарушением социальных норм скорее вызывает гнев, чем презрение, потому что показывает другим, что человека не волнует их реакция. Хотя эти спонтанные тенденции к действию могут сдерживаться издержками и риском применения санкций, они, тем не менее, могут иногда преодолевать последние. Остракизм «ботаника», который в состоянии помочь одноклассникам с домашней работой, может дорого обойтись, так же, как отказ аристократов при *ancien régime* выдавать своих дочерей за богатых буржуа. Когда «вкус» к дискриминации выражается в отказе нанимать или покупать что-либо у представителей презируемого меньшинства или у женщин, это может негативно сказаться на экономической эффективности. Часто такое поведение отражает действие социальных норм, а не индивидуальные идиосинкразические представления, что доказывают такие выражения, как «любитель евреев» или «любитель негров», призванные заклеить тех, кто пошел против социальной нормы.

ЧЕМ НЕ ЯВЛЯЮТСЯ СОЦИАЛЬНЫЕ НОРМЫ

Социальные нормы следует отличать от ряда связанных с ними явлений: моральных, квазиморальных, юридических норм и конвенций. Хотя иногда сложно провести четкое различие, каждая категория ясно очерчивает определенный круг явлений. И моральные, и квазиморальные нормы (глава V) могут влиять на поведение, даже когда агенты полагают, что за ними никто не наблюдает. И наоборот, стыд, поддерживающий социальные нормы, вызывается ощутимым презрением других. Соот-

ветствующая тенденция к действию — бежать от обвиняющих взглядов: спрятаться, скрыться и даже покончить с собой.

Юридические нормы отличаются от социальных, поскольку их соблюдение контролируется специализированными агентами, которые обычно применяют прямые наказания, а не остракизм, несмотря на эксперименты с юридическим «посрамлением». Юридические и социальные нормы взаимодействуют самым разным образом. Например, в 1990 году некоторые законодатели в Луизиане выступали за снижение уголовных санкций за неформальный самосуд над сжигающими флаги. Даже после того как эдикт 1701 года разрешил французской знати заниматься торговлей (только оптовой, но не розничной), прошло более 50 лет прежде, чем та смогла преодолеть социальные нормы, осуждавшие подобную практику. В некоторых сообществах существуют социальные нормы против обращения к юридическим нормам, тогда как в других люди могут подать в суд за сбитую с головы шляпу.

Конвенции или равновесие конвенций может реализовываться посредством чистого эгоистического интереса агента, без каких-либо действий со стороны окружающих. Как уже отмечалось в главе XIX, зачастую они весьма произвольны. В первый день конференции каждый участник может выбрать место более или менее произвольно. На второй день уже складывается конвенция: люди стремятся сесть на выбранные ранее места потому, что с очевидностью (точка координации) включается распределительный механизм. На третий день конвенция окрепла настолько, что дает определенные права. Я рассержусь, если другой участник займет «мое» место. И все же хотя социальная норма цементирует произвольные конвенции и делает более вероятным их соблюдение, без нее можно обойтись. Среди ньюйоркцев существует конвенция праздновать Новый год на Таймс-сквер, но поскольку немногие знают, появится там конкретный человек или нет, возможность санкций невелика. Даже если бы норма езды на машине по правой стороне дороги не была усилена социальными и юридическими нормами, опасность перестраивания на левую оставалась бы сильным сдерживающим фактором.

Более сложная категория — неписанные юридические и политические нормы, такие как конституционные конвенции⁴. Обычно они не могут осуществляться в судебном порядке, хотя суд и может учитывать их при принятии решений. Вместо этого они осуществляются посредством политических санкций или при помощи страха перед такими санкция-

⁴ Фраза «конституционные конвенции» используется по отношению к двум совершенно разным аспектам конституций: неписанные нормы, которые дополняют письменную конституцию и конституционные собрания, которые часто используются для принятия письменных конституций.

ми. До 1940 года американская конституционная конвенция, согласно которой никто не мог оставаться президентом более двух сроков, исполнялась благодаря вере в то, что любой, кто попытается это сделать, проиграет на выборах. Многие из таких норм сходны с социальными, поскольку поддерживаются силами общественного мнения, а не специализированными инстанциями. Другие политические конвенции лучше рассматривать как равновесие в многократно повторяемых играх. Во многих парламентских системах есть, например, конвенция, согласно которой, когда администрация уходит в отставку, ее внутренние документы печатаются и становятся доступны историкам лишь по прошествии нескольких десятилетий. Хотя каждая новая администрация испытывает искушение открыть архивы предшественников и использовать их в качестве политического орудия, знания о том, что они создадут прецедент для тех, кто придет им на смену, достаточно, чтобы воздержаться от таких действий. Это не конвенция в том смысле, в каком это слово используется в главе XIX, поскольку каждая администрация предпочла бы уклониться от нее при соблюдении всеми остальными.

НОРМЫ И ЭКСТЕРНАЛИИ

Существуют нормы, направленные против тех, кто навязывает другим небольшие негативные экстерналии (глава XVII). Когда люди мусорят в парке, плюют на улице, мочатся в озеро или пьют общий кофе в офисе, не опустив монетку в копилку, они, как правило, стараются, чтобы их не заметили. Даже когда они фактически не боятся санкций, мысль о том, что другие могут плохо о них подумать, может удержать от осуществления этих действий у всех на виду. Нормы такого рода являются социально полезными, поскольку от них лучше *всем*. Самым удачным примером может служить норма, запрещающая плевать в общественных местах. До того как люди узнали, как распространяются инфекционные заболевания, это была совершенно приемлемая практика, поощрявшаяся наличием плевательниц. Как только был раскрыт механизм заражения, во многих общественных местах появился знак «не плевать». Сегодня эта норма так закрепилась (по крайней мере, в некоторых странах), что знаки убрали.

В этом примере мы можем наблюдать возникновение нормы и с некоторой уверенностью утверждать, что она возникла, *так как* отвечала общим интересам. Была выявлена угроза, создана юридическая норма, а за ней последовала социальная. Может ли закрепление негативных экстерналий создавать социальные нормы без общественного вмешательства — спорный вопрос. Сам по себе факт осознания потребности в норме необязательно автоматически вызывает ее к жизни. В развивающихся

странах нет социальной нормы, требующей ограничивать численность семьи. Социальные нормы против вытаптывания пастбищ, которые могли бы предотвратить трагедию общин, не возникли спонтанно. Нет норм, регулирующих использование антибиотиков, хотя их чрезмерное потребление отражается на других, способствуя развитию более стойких микроорганизмов. Нормы, направленные против музыки на общественных пляжах и против использования сотовых телефонов в концертных залах, также обязаны своим происхождением (как я предполагаю) действиям соответствующих инстанций. Еще и еще раз подтверждается, для того чтобы помешать людям создавать негативные экстерналии, требуется постороннее вмешательство. В некоторых случаях, как в случае с плеванием в общественных местах, люди могут избегать подобных действий, даже после исчезновения или прекращения действия юридической нормы. В других, как в случае политики «одного ребенка» в Китае, поведение едва ли сохранится, если будут сняты ограничения.

Более мелкие группы могут применять эти нормы без постороннего вмешательства. На рабочих местах часто существуют очень строгие нормы против тех, кто перевыполняет план, потому что считается, что их усилия могут привести к снижению администрацией сдельной оплаты (в этом случае экстерналия принимает форму увеличения вероятности урезания сдельной расценки). Хотя менеджмент может придерживаться политики фиксированных расценок для поощрения повышения производительности, он не в состоянии дать правдоподобные обещания на сей счет. Как правило, штрейкбрехеры подвергаются жестоким санкциям со стороны своих товарищей по работе. Возможно, имеет значение, что оба случая предполагают общее противостояние противнику. В «игре против природы», например, при вытаптывании пастбищ, солидарность рождается с большим трудом, потому что любовь к «безбилетничеству» не рассматривается в качестве *предательства*. С этой точки зрения не следует ожидать спонтанного возникновения нормы, направленной против халтурщиков, в компаниях, предлагающих командные бонусы, поскольку нарушения этой нормы лишь наносит ущерб другим работникам, не принося выгоды «врагу» (но см. также главу XXVI).

Другие социальные нормы направлены на негативные экстерналии, которые одна группа людей навязывает другой. Например, норма против курения, даже в тех местах, где оно пока разрешено⁵. Во многих за-

⁵ Самая важная экстерналия вызывается вдыханием дыма (пассивное курение). Иногда утверждается, что курильщики навязывают негативную экстерналию другим курильщикам, которые собираются бросить, но не могут побороть желание закурить, спровоцированное визуальным сигналом чужого курения.

падных обществах курящие гости сегодня часто воздерживаются от курения, даже не спрашивая разрешения у хозяина. То, что можно назвать «экстерналиями шума» лежит в основе нормы «детей нужно видеть, но не слышать». Этот запрет может стать социальной нормой, а не просто формой родительского наказания двумя способами. Во-первых, дети могут подвергаться остракизму других детей, ее нарушающих. Второе, родители могут подвергаться остракизму других родителей, чьи дети нарушают норму. В купе поезда те, кто хочет навязать другим «экстерналию свежего воздуха», обычно поигрывает в споре с теми, кто навязывает «экстерналию духоты». Причина, возможно, в том, что закрытое окно воспринимается как опция по умолчанию и, следовательно, нормативная базовая линия.

НОРМЫ И КОНФОРМИЗМ

Некоторые социальные нормы — всего лишь предписания не высываться. Жители любого маленького города узнают «закон Янте», записанный (в 1933 году) человеком, который сумел оттуда сбежать:

Не думай, что ты что-то собой представляешь.

Не думай, что ты столь же значителен, как мы.

Не думай, что ты умнее нас.

Не воображай, что ты лучше нас.

Не думай, что ты знаешь больше, чем мы.

Не думай, что ты больше нас.

Не думай, что ты на что-то годишься.

Не смейся над нами.

Не думай, что кому-то есть до тебя дело.

Не думай, что ты можешь чему-то нас научить.

Эти нормы могут иметь очень плохие социальные последствия. Они могут помешать наиболее одаренным воспользоваться своими талантами и повлечь обвинения в колдовстве, если им это удастся. За везение тоже недолго любить. В племени бемба в Северной Родезии говорят, что найти в лесу улей с медом — везение, два улья — очень большое везение, найти три улья — это уже колдовство.

КОДЕКСЫ ЧЕСТИ

Сильные и зачастую изощренные нормы могут регулировать поведение при распрах, вендеттах, на дуэлях и в ситуации мести в целом. Нормы определяют действия, которые требуют воздаяния или вызова, условия и средства осуществления мести и судьбу того, кто не сумел соответствовать основной норме. Начав с последнего, нужно отметить, что неспо-

способность отомстить часто приводит к своего рода гражданской смерти, в которой агент оказывается полностью отрезан от нормальных социальных отношений. Его семья не считается с его мнением; если он решится выйти из дома, его встречают насмешками или чем похуже. Это парадигматическая ситуация презрения, вызывающая непереносимый стыд.

Все, что хотя бы отдаленно может рассматриваться как оскорбление чести агента, может стать поводом для возмездия. В предреволюционном Париже виконт де Сегюр, известный повеса, развлекался написанием стихотворных эпиграмм. Соперник, завидовавший его репутации, написал небольшое стихотворение, в котором тонко высмеивал стихи Сегюра. Чтобы расквитаться, Сегюр соблазнил его любовницу, а когда она призналась, что беременна, сказал ей, что просто использовал ее для мести сопернику и что теперь, когда цель достигнута, она ему больше не нужна (впоследствии она умерла при родах). Сегюр вернулся в Париж и рассказывал историю всем желающим, ни разу не встретив осуждения. Кажется, «Опасные связи» были лишь слабой имитацией реальности.

На Корсике в XIX столетии были четыре обстоятельства, оправдывавшие или требовавшие отмщения: когда обесчещена женщина, когда разорвана помолвка, когда убит близкий родственник и когда лжесвидетельство в суде привело к осуждению члена семьи мстителя. Однажды нотариуса осудили за убийство на основании ложных показаний, и впоследствии он умер в тюрьме. Его брат стал разбойником и за несколько лет убил четырнадцать свидетелей обвинения. Это примеры мести для *сохранения* чести. Однако система чести включала действия, предпринимаемые в целях ее *приобретения*. Монтень ссылается на то, что «итальянцы, когда желают укорить молодых людей за их безрассудную смелость, называют их *bisognosi d'onore*, „нуждающиеся в чести“».

На американском Юге люди острее реагируют на полученные оскорбления, чем северяне. Процент убийств на Юге выше, и люди больше одобряют свирепые реакции на обиды. В хитроумном эксперименте подсадное лицо толкало участника «как бы случайно» и называло его «идиотом». При замере, выяснилось, что уровни кортизола (отражающие реакцию на происшествие) и тестостерона (отражающие готовность к последующей агрессии) у участников-южан выросли радикально выше, чем у участников-северян. В другом эксперименте участники шли по коридору, в котором их «толкнули», и видели, как к ним с решительным видом направляется крупный человек, напоминающий футбольного игрока (подсадное лицо). Коридор был заставлен столами, так что двоим в нем было не разойтись, тем самым создавались условия для игры в «Цыпленка». Прежде чем продемонстрировать «цыплячью реакцию», южане подходили гораздо ближе к другому человеку (3 фута), чем северяне (9 футов).

Выполняют ли кодексы чести социальную функцию? Если да, то объясняет ли эта функция, зачем они существуют? Идея о том, что практика мести — полезная форма контроля численности населения, слишком произвольна, чтобы быть принята всерьез. Альтернативный взгляд, полагающий, что нормы мести обеспечивают функциональный эквивалент организованного контроля исполнения законов в обществах со слабым государством, также малоубедительно. В средиземноморских и ближневосточных обществах, придерживающихся этих норм, уровень насилия и смертности среди молодых мужчин гораздо выше, чем в любых других регионах⁶. Как подсказывает процитированное выше замечание Монтеня, нормы мести и более широкий кодекс чести, в который они встроены, могут разжечь столько же пожаров, сколько погасить. Часто распри усиливают хаос, а не контролируют его.

Некоторые полагают, что нормы чести развиваются в находящихся на удалении друг от друга скотоводческих обществах, в которых репутация человека, всегда готового прибегнуть к насилию, является полезным и даже необходимым средством для сдерживания воровства. Так объяснялась культура чести на американском Юге. Помимо общих проблем функционального объяснения, этот анализ наталкивается на трудность, состоящую в том, что законы чести были в равной степени сильны и при дворе французских королей в XVII–XVIII веках, если взять пример, не относящийся к сельской жизни. Некоторые из тех, кто сосредоточился на законах чести в среде городской аристократии, позднее выдвинули еще одно функциональное объяснение: в отсутствие войны знать «нуждалась» в дуэлях, чтобы поддерживать боевой дух. Если не указать механизм, при помощи которого эта потребность генерирует свое удовлетворение, этот аргумент ничего не стоит. Эти полемические замечания не подразумевают, что я могу предложить лучшее объяснение.

НОРМЫ ЭТИКЕТА

Еще один набор социальных норм — нормы, связанные с правилами хороших манер, или *этикетом*. Коды одежды, языка, поведения за столом и тому подобные нормы часто непоколебимы в деталях, обрекая на ост-

⁶ Могут возразить, что здесь релевантно сравнение с уровнем насилия, который мог бы существовать в «естественном состоянии». Если это состояние характеризуется исключительной сосредоточенностью на эгоистических интересах и отсутствием каких бы то ни было государствовподобных инстанций, в нем не происходило бы насилие, мотивированное завистью, злобой или гневом. Если бы слабые предчувствовали, что произведенные ими блага будут насильно отобраны сильными, это оказало бы губительное действие на производство, предотвратив реальное насилие.

ракизм тех, кто пренебрегает даже нюансами⁷. Во всех обществах есть норма, устанавливающая допустимую дистанцию между людьми при социальном взаимодействии. Если кто-то вторгнется внутрь личного пространства человека (в США это где-то 15 дюймов), от него могут отвернуться как от невоспитанного человека. Однако эта норма необычна, поскольку индивиды, к которым она относится, часто не осознают ее существование и действие. Большинство норм этикета сильно кодифицировано, обычно в буквальном смысле. Они не только (по большей части) бессмысленны, но даже жестоки по своим последствиям, например, когда пятилетняя девочка приходит домой в слезах, потому что ее подруги смеялись над ней из-за *отсутствия тормозов* в ее новой игрушечной коляске. В предреволюционном Париже молодой офицер, богатый, но не знатный, хотел пройти без приглашения на бал в Версале. «К нему отнеслись столь жестоко, что в приступе отчаяния из-за насмешек, которым он подвергся во времена, когда насмешки были худшим злом, по возвращении в Париж он покончил с собой».

Почему же эти, по сути, тривиальные вопросы приобретают такое значение? Непропорциональное осуждение, вызываемое нарушением этикета, может быть связано с неподтвержденным представлением о человеке как о цельной личности, так что нарушитель незначительной нормы в дальнейшем может допустить и более серьезный проступок. Нарушение мелких правил этикета может рассматриваться как открытая демонстрация неуважения к мнению других. Все это, однако, не объясняет, почему несущественные нормы возникают. Загадочно не то, почему та или иная норма существует, а то, почему люди придают вес незначительным вещам. Как только они *начинают рассматриваться* как важные, эти нормы *начинают регулировать* поведение. Если кто-то приходит на собеседование для трудоустройства в банк в розовой кожаной куртке, это означает, что он либо намеренно бравирует нарушением социальной нормы, либо глух к ожиданиям других. В обоих случаях это хорошее основание для того, чтобы взять на работу кого-то другого.

И вновь широкое распространение получили функциональные объяснения. Предполагают, что сложные правила этикета существуют для того, чтобы затруднить посторонним «проход без приглашения» посредством имитации поведения, регламентированного правилами. Без сомнения, эти правила часто имеют *эффект* сдерживания выскокчек, но это не объясняет, почему они существуют. Как выяснили многие студен-

⁷ В аристократических обществах иногда приемлемы *грубые* нарушения, если они рассматриваются как преднамеренные, а не как свидетельство незнания правил. В качестве примера можно упомянуть де Шарлю у Пруста.

ты, решившие добровольно стать пролетариями, тому, кто не принадлежит к рабочему классу по рождению, непросто в него войти. Например, в Норвегии 1970-х годов молодые маоисты узнали, что высмеивание королевской семьи — прямой путь к отчуждению со стороны того класса, в который они так хотели попасть. И все же никто не высказывал предположения, что нормы рабочего класса существуют для того, чтобы посторонним было сложнее выдавать себя за них. Не больше смысла этот аргумент имеет в отношении элит.

НОРМЫ, РЕГУЛИРУЮЩИЕ ИСПОЛЬЗОВАНИЕ ДЕНЕГ

В дополнение к легальным запретам на использование денег для покупки детей, голосов избирателей, школьных оценок, органов для трансплантации и (иногда) секса существует множество социальных норм, которые регулируют их правильное и неправильное расходование. Некоторые нормы действуют среди друзей и соседей. Например, запрет взрослому, живущему в пригороде, косить лужайки за плату (глава XII). Когда соседи кооперируются, чтобы сделать забор, они часто вносят равный вклад — например, один вкладывает свой труд, а другой — материалы. Даже если бы было эффективнее, чтобы один из них выполнил всю работу, а другой выплатил ему компенсацию, норма против финансовых сделок между соседями исключает такое решение. В США преподаватели частных колледжей даже не спрашивают друг друга о том, сколько им платят, если только они не близкие друзья.

Другие нормы, как ни удивительно, контролируют использование денег в отношениях с посторонними людьми. Например, существует норма, согласно которой нельзя подойти к человеку, стоящему в начале очереди на автобус, и предложить ему деньги в обмен на его место⁸. Понятно, что эта норма неэффективна: если человек, которому предложат деньги, согласится перейти в конец очереди, оба выиграют и никто не пострадает. Согласно Токвилю, такие нормы против откровенной демонстрации богатства на публике свойственны демократическим обществам: «Видите этого состоятельного гражданина?.. На нем простая одежда, его манеры скромны. В четырех стенах его дома роскошь». Существуют даже нормы, регулирующие разговоры о деньгах или осознание денежных аспектов транзакций. В некоторых ресторанах парам до сих пор подают два

⁸ Если человек скажет, что ему нужно срочно попасть домой, потому что у него болен ребенок, его могут пропустить в начало очереди, но тогда уступивший ему свое место, станет вторым. Поскольку это может сказаться на всех остальных людях, стоящих в очереди, может возникнуть потребность в их согласии.

меню, одно без цен — для женщин, другое с ценами — для мужчин, что отражает норму, согласно которой галантность должна быть незапятнана заботами о деньгах. Если вы приносите на вечеринку бутылку вина, предполагается, что вы должны снять с нее ценник.

НОРМЫ, СВЯЗАННЫЕ С ПОТРЕБЛЕНИЕМ АЛКОГОЛЯ

Если социальные нормы рассчитаны на благополучие индивида или общества, можно ожидать, что они будут направлены против пьянства, которое, как предполагается, имеет губительные последствия. Действительно, существует много норм такого рода. Некоторые из них, обычно связанные с религией, требуют полного воздержания. В исламе и в некоторых протестантских сектах действует полный запрет на алкоголь. Светские нормы, наоборот, часто предписывают умеренное потребление. Итальянское правило «никогда не пить перед едой» имеет двойной эффект ограничения общего потребления алкоголя и сокращения скорости его поглощения, ослабляющего его краткосрочное воздействие на организм. В Исландии есть нормы против приема алкоголя в присутствии ребенка и против распития алкогольных напитков на рыбалке.

Однако нормы, относящиеся к алкоголю, не всегда способствуют благополучию. Существуют как нормы, осуждающие трезвость, так как и нормы, заставляющие людей напиваться. Среди индейцев мапуче в Чили предосудительным считается и распитие алкоголя в одиночестве, и трезвость; такое поведение рассматривается как проявление недоверия. Традиционная французская культура одинаково осуждает трезвенника и пьяницу. В Италии недоверие к трезвенникам выражается в пословице: «Упаси господь от тех, кто не пьет». В молодежных субкультурах многих стран трезвенники подвергаются сильному давлению и насмешкам. И наоборот, во многих обществах пьянство предписано социально. В Мексике и в Нигерии восхищаются мужскими качествами, проявляющимися в способности много пить. В предреволюционной России чрезмерное пьянство было обязательным в субкультуре молодых офицеров.

Когда трезвость осуждается или когда пьянство социально обязательно, трезвенникам приходится прибегать к хитрости. В Швеции распространен вопрос: «Хочешь шерри или ты за рулем?» Это настолько общепринято, что желающие оставаться трезвыми алкоголики часто говорят, что за рулем, потому что это освобождает их от социального давления, которое в противном случае хозяин оказывал бы на гостя. Норма, связанная с приемом алкоголя, может быть компенсирована только действием другой нормы (направленной против вождения

машины в пьяном виде). Например, обращение в протестантизм дает альтернативу латиноамериканцам, которые хотели бы освободиться от власти сообщества, в котором даже ритуалы зачастую связаны с пьянством и алкогольным опьянением. И снова норма, поощряющая пьянство, может быть преодолена лишь посредством другой, опирающейся на религию.

Существуют случаи стратегического использования норм, но и люди могут вести себя стратегически, чтобы их обойти. Древние китайцы считали алкоголь священным и пили его только на жертвенных церемониях; возможно, иногда они приносили жертву, когда им хотелось выпить. В Испании есть молчаливое культурное предписание не пить на голодный желудок в определенное время суток, так что алкогольные напитки в этом случае подаются только вместе с едой. В обоих случаях мы наблюдаем обратную сторону изначальной каузальной цепочки: вместо того чтобы следовать конвенциональной норме, предписывающей пить только, когда они делают X, люди делают X всякий раз, когда им хочется выпить.

НОРМЫ, КАСАЮЩИЕСЯ ЧАЕВЫХ

Выплата чаевых — существенный феномен. По оценкам, чаевые в американских ресторанах варьируются от 5 до 27 миллиардов в год; чаевые водителям такси, парикмахерам и прочим, возможно, дали бы еще большую цифру. Оценки дохода, который официанты получают от чаевых, варьируются от 8 (по оценкам Службы внутренних доходов) до 58% для официантов, подающих полный обед. В некоторых контекстах наличие чаевых может показаться странным, в других — более уместно. Если вы ходите стричься к одному и тому же парикмахеру, вы платите чаевые, чтобы обеспечить хорошее обслуживание, то же относится к обеду в вашем любимом ресторане. Больше озадачивают чаевые при однократных встречах, как, например, при поездке на такси или обеде в ресторане, в который вы не собираетесь возвращаться. Такое поведение озадачивает вдвойне: оно не может поддерживаться взаимодействием двух сторон во времени или санкциями со стороны третьих лиц в момент взаимодействия. Если вы — единственный пассажир в такси, другие люди едва ли узнают, дали ли вы таксисту подходящие чаевые. Другие посетители в ресторане тоже едва ли заметят, сколько вы дали на чай официанту.

Считается, что чаевые — эффективный способ вознаградить официантов. Очевидно, клиентам легче отслеживать качество услуг, чем владельцу ресторана. Следовательно, децентрализация функции контроля качества и увязывание награды с наблюдаемой работой — способ ре-

шить проблему «принципал—агент» (как предотвратить халтуру?), которая осложняет многие договорные отношения (глава XXVI). Таким образом, они могут быть частью «подразумеваемого контракта» в целях усиления эффективности. Но, как говорил Сэм Голдвин, неписанный контракт не стоит бумаги, на которой он написан. Этот аргумент, равно как и многие другие попытки объяснить социальные нормы, — пример неоправданного функционализма. Идея, что владельцы ресторанов, запрядающих чаевые, вытесняют конкуренты, которые их допускают, чисто гипотетическая; она не может объяснить, почему клиенты их дают. Кроме того, с эмпирической точки зрения, раздача чаевых не выдерживает проверки на эффективность. Например, она не занимает доминирующего положения в профессиях, где легче осуществлять контроль качества. То, что официанты часто складывают чаевые в один фонд, также подрывает довод, связанный с эффективностью.

Я не знаю, почему представителям одних профессий дают чаевые, а другим нет. Но если норма существует, мы можем понять, почему люди ей следуют: им не нравится, что другие, например разочарованный таксист, осудят их, даже если они с ним больше не встретятся. Необязательно быть объектом презрительного взгляда других. Иногда достаточно знать или иметь основание полагать, что другие испытывают презрение. Приведем еще один пример: вера в то, что другие люди могут меня осудить, объясняющая, почему я стараюсь не ковырять в носу на платформе метро, когда мимо без остановки проходит поезд, даже если на обеих платформах больше никого нет.

ПОЧЕМУ НОРМЫ?

Важность социальных норм для регулирования поведения и механизм, при помощи которого они действуют, понятны. Однако я не считаю, что мы хорошо понимаем их происхождение. Есть два вопроса. Во-первых, каково эволюционное происхождение соответствующих эмоций стыда и презрения, поддерживающих социальные нормы? Другими словами, почему вообще это социальные нормы? Во-вторых, почему специфические нормы существуют в специфических обществах? Как и когда они возникают; как и почему исчезают?

Ответ на первый вопрос заключается в том, что нас очень беспокоит мнение других людей. Мы ищем их одобрения и боимся порицания. Однако этот ответ возвращает нас к тому же вопросу: почему мы должны беспокоиться о том, что подумают о нас другие? Конечно, в некоторых случаях репутация полезна и о ней стоит заботиться. Но мы даем чаевые таксисту, чтобы тот не думал о нас плохо, не заботясь о репутации. К тому же, если причина осуждения окружающих кроется в нарушении

социальных норм, объяснять их существование нежеланием быть осужденным другими, — значит завершать логический круг.

Что касается второго вопроса, самый распространенный ответ — нормы возникают, чтобы управлять экстерналиями. В этом что-то есть, если добавить, что социальные нормы против навязывания другим негативных экстерналий обычно вводятся внешней властной инстанцией. Есть общая социальная норма подчиняться закону. Если бы штрафы рассматривались как цены, а пребывание в тюрьме было не более позорным, чем нахождение в больнице, этой нормы не существовало, но такие реакции на нарушение закона не рассматриваются как равные другим, объективно эквивалентным обременениям. Люди стыдятся того, что сидели в тюрьме, и стараются по возможности скрыть этот факт⁹. Когда закон налагает запрет на поведение, навязывающее другим негативные экстерналии, социальная норма подчинения закону может выразиться в осуждении подобного поведения. Норма может сохраниться, даже если давший ей жизнь закон не применяется. Однако такой результат трудно отличить от появления «хорошего равновесия» в игре на доверие (глава XIX). Если государство сначала поощряет сотрудничество, наказывая уклоняющихся от него, а затем упраздняет карательный аппарат, люди могут продолжать сотрудничать, потому что ставят на первое место ситуацию, в которой все сотрудничают (нет соблазна «безбилетничества»).

Гораздо труднее объяснить возникновение и сохранение многих других норм, таких, например, как норма, запрещающая предлагать деньги за место в очереди на автобус, правила этикета и нормы, касающиеся чаевых. Одна линия аргументация, часто предлагаемая экономистами, заключается в том, что сохранение этих норм может быть объяснено с точки зрения поведения в точке равновесия, а их возникновение — вопрос случая и истории, о которых социальная наука мало что может сказать. Поскольку предпосылка этой книги состоит в том, что разделительная линия между социальной наукой и историей носит искусственный и бессмысленный характер, я не могу согласиться с последним утверждением. Что же касается первого — то, социальные нормы, как правило, не демонстрируют логику наилучшего ответа, характеризующую стратегические игры. Когда, оставаясь незамеченным, я наблюдаю, как другой нарушает норму, применение санкций в отношении нарушителя обычно не является наилучшей реакцией.

⁹ В Норвегии в свое время был обязательный трехнедельный тюремный срок за вождение автомобиля в состоянии опьянения. Некоторые люди брали с собой в камеру ультрафиолетовую лампу, чтобы приобрести загар, который подтвердит легенду о проведенном отпуске.

Библиографические примечания

Эта глава построена и (как я надеюсь) совершенствует рассмотрение норм, предложенное мной в «Цементе общества» (*The Cement of Society*. Cambridge University Press, 1989) и кратко в «Социальных нормах и экономической теории» (Эльстер Ю. Социальные нормы и экономическая теория // *THESIS*. 1993. № 3. С. 73–91). Важные рассуждения о социальных нормах можно найти в книгах Дж. Колемана «Основы социальной теории» (*Coleman J. Foundations of Social Theory*. Cambridge, MA: Harvard University Press, 1990), Р.Элликсона «Порядок без закона» (*Ellickson R. Order Without Law*. Cambridge MA: Harvard University Press, 1999) и Э.Познера «Закон и социальные нормы» (*Posner E. Law and Social Norms*. Cambridge, MA: Harvard University Press, 2000). Я кое-что почерпнул из этих книг, но ни одна из них меня не убедила. См. поучительную критику Познера в рецензии Р.Макадамса (*McAdams R. // Yale Law Journal*. 2001. No. 110. P. 625–690). Полезное обсуждение неписаных конституционных норм или конвенций можно найти в статьях Дж. Джаконелли (*Jaconelli J.*) «Природа конституционных собраний» (*The nature of constitutional convention // Legal Studies*. 1999. No. 24. P. 24–46) и «Связывают ли конституционные собрания обязательствами?» (*Do constitutional conventions bind? // Cambridge Law Journal*. 2005. No. 64. P. 149–176). «Закон Янте» взят из романа А.Сандемусе «Беглец пересекает свои следы» (*Sandemose A. A Fugitive Crosses His Trail*. New York: Knopf, 1936). Роль колдовства в поддержании норм, направленных против пытающихся выделиться, исследуется в книге К. Томас «Религия и упадок магии» (*Thomas K. Religion and the Decline of Magic*. Harmondsworth, England: Penguin, 1973). Я рассматриваю кодексы чести и мести в главе 3 «Алхимии ума» (*Alchemies of the Mind*. Cambridge University Press, 1999). История о виконте де Сегюре взята из «Мемуаров графини де Буань» (*Les memoires de la Comtesse de Boigne*. Paris: Mercure de France, 1999. Vol. 1. P. 73–74). Те же мемуары (с. 38) стали источником истории о молодом офицере, покончившем с собой из-за стыда, вызванного насмешками. Нормы этикета являются темой книги П. Бурдьё «Различение» (*Bourdieu P. Distinction*. Cambridge MA: Harvard University Press, 1987; фрагменты работы в русском переводе см.: книги «Западная экономическая социология: Хрестоматия современной классики» (сост. и науч. ред. В.В. Радаев. М.: РОССПЭН, 2004. С. 537–568)), которая носит очевидный функционалистский уклон. Экспериментальные исследования «культуры чести» приводятся в книге Р. Нисбетта и Д. Коэна «Культура чести» (*Nisbett R., Cohen D. The Culture of Honor*. Boulder, CO: Westview Press, 1996). Пример с нормами потребления алкоголя взят из моей книги «Сильные чувства» (*Strong Feelings*. Cambridge, MA: MIT Press, 1999). Злоключения норвежских студентов, решивших стать пролетариями, описываются в очень забавном романе Д. Солстада, к сожалению, не переведенном на английский (*Solstad D. Gymnaslærer Pedersens beretning om den store politiske vekkelsen som har*

hjemstøkt vårt land. Oslo: Gyldendal, 1982). Касательно норм, запрещающих спрашивать другого человека, сколько он зарабатывает, см. статью М. Эдвардса «Законы и социальные нормы сокрытия размеров оплаты» (The law and social norms of pay secrecy // Berkeley Journal of Employment and Labor Law. 2005. No. 26. P. 41–63). Основанное на эффективности объяснение нормы, связанной с чаевыми, предлагается в работе Н. Джейкоба и А. Пейджа «Производство, затраты на информацию и экономическая организация: случай контроля качества покупателем» (Jacob N., Page A. Production, information costs and economic organization: The buyer monitoring case // American Economic Review. 1980. No. 70. P. 476–478). Эта работа подвергается критике в статье М. Конлина, М. Линна и Т. О'Донохью «Норма чаевых в ресторане» (Conlin M., Lynn M., O'Donoghue T. The norm of restaurant tipping // Journal of Economic Behavior & Organization. 2003. No. 5. P. 297–321), которая предлагает объяснение, наиболее близкое к представленному здесь.

XXIII. Коллективное формирование убеждений

ТОКВИЛЬ О КОНФОРМИЗМЕ

Механизмы формирования убеждений, рассмотренные мною в главе VII, действуют по большей части на уровне индивида; иными словами, убеждения одного человека не слишком зависят от тех, которых придерживаются или которые выражают другие. В этой главе я остановлюсь на механизмах формирования коллективных или интерактивных убеждений. Чтобы проиллюстрировать это различие, обратимся к анализу Токвиля американского конформизма. Одно из объяснений, почему американцы склонны разделять одни идеи, заключается в том, что они живут в похожих условиях: поскольку «люди, равные по условиям... видят вещи под одним и тем же углом, их умы естественным образом склоняются к аналогичным идеям; и, хотя каждый из них может отходить от своих современников и формировать свои собственные убеждения, в конце концов оказывается, что все неосознанно и непреднамеренно разделяют целый ряд убеждений». Другое объяснение основывается на давлении, которое заставляет приспособливаться: «В Америке большинство воздвигает громадные барьеры вокруг мысли. Внутри очерченных таким образом границ автор совершенно свободен, но горе тому, кто осмелится выйти за эти пределы. Не то чтобы ему нужно бояться *аутодафе*, но ему придется столкнуться со всякого рода неприятностями и каждодневным преследованием».

Этот последний пассаж свидетельствует, что люди приспособливаются внешне вследствие социального давления, но не обязательно внутренне. Как пишет тот же Токвиль, если вы придерживаетесь других взглядов, «ваши ближние будут вас сторониться как нечистого. И даже те, кто верит в вашу невинность, покинут вас, иначе их тоже будут избегать в ответ». Еще один отрывок указывает, что конформизм проникает в душу, так что у людей может развиваться искренняя вера в мнения большинства. Здесь указывается на два механизма: один — «холодный» (когнитивный), а другой — «горячий» (мотивационный). С одной стороны, «представляется маловероятным... чтобы все были в одинаковой мере просвещенными, истина не должна принадлежать большинству». С дру-

гой стороны, тот факт, что «американские политические законы таковы, что большинство является суверенным... значительно увеличивает свойственное ему влияние на сознание, ибо у человека нет более закоренелой привычки, чем признавать высшую мудрость своего угнетателя».

ЭКСПЕРИМЕНТАЛЬНЫЕ ОТКРЫТИЯ

Я процитировал Токвиля (и еще буду делать это в этой главе) по причине его проницательных догадок по этому вопросу. Проблемы, которые он обозначил (внешний и внутренний конформизм), а также когнитивные и мотивационные механизмы до сих пор остаются для нас крайне актуальными. Прежде чем к ним обратиться, я расскажу о некоторых классических экспериментах, касающихся конформности.

В самом знаменитом эксперименте участников просили указать, какая из трех линий (А, В и С) ближе всего по длине к данной линии D. Было три вида условий: частные, двойные публичные и одиночные публичные. В частных условиях испытуемые давали свой ответ в отсутствие всех остальных и в присутствии одного только экспериментатора. В этом случае 99% участников сказали, что линия D ближе всего к линии В, что говорит об однозначной правильности этого ответа. Но в двух случаях публичных условий существенное меньшинство участников дало разные ответы. В обоих случаях участники отвечали после того, как несколько других (подсадных) лиц единогласно заявили, что линия А ближе по длине. При двойных публичных условиях, где участники давали ответ в присутствии подсадных лиц, около трети согласились, что ближе линия А¹. В одиночно публичных условиях, в которых участники высказывали свое мнение в частном порядке, после того как услышали, что сказали другие, конформизм снизился, но не был полностью устранен.

Избыток конформизма в двойных публичных условиях был, предположительно, связан со *страхом неодобрения*. Остаточный конформизм в одиночных публичных условиях может быть связан с *обучением* («столько людей не может ошибаться») или *редукцией диссонанса*. Последнее объяснение кажется более правдоподобным. Те, кто в частном порядке стремились следовать за большинством, едва ли делали это только на основании рационального обучения, учитывая незначитель-

¹ В главе V я описывал, как наблюдение за другими может вызвать поведение, подобное этому, благодаря квазиморальной норме честности, тогда как тот факт, что за вами наблюдают, может вызвать такое же поведение из-за страха неодобрения. При формировании убеждений наблюдение со стороны других тоже может породить конформность через страх осуждения, тогда как конформность, порождаемая наблюдением за другими, происходит или путем обучения, или через редукцию диссонанса.

ный когнитивный статус мнения большинства. Здесь должны были работать мотивационные факторы.

Еще один эксперимент подтверждает такую интерпретацию. В нем перед участниками стояла более неоднозначная задача: они должны были определить, какое расстояние проходит источник света в темноте. Хотя источник света на самом деле был неподвижен, изолированные участники делали вывод, что он сдвинулся примерно на 10 см («автокинетический эффект»). Услышав, как другие говорят, что он сдвинулся на 38–40 см, испытуемые оценивали расстояние приблизительно в 20 см. Если два подсадных участника оценивали расстояние в 40 см, оценка других была — 35 см. Присутствие одного подсадного участника, таким образом, приводило к увеличению оценки на 10 см, а присутствие второго — еще на 15 см.

В процессе байесова обучения (см. главу XI) я могу полагаться на других наблюдателей, чтобы скорректировать свое восприятие или память. Их оценки какого-либо факта (например, пройденного светом расстояния) могут способствовать изменению моей собственной оценки. Насколько они на меня повлияют, зависит от надежности их восприятия и от количества других наблюдателей. В этом эксперименте участник, предположительно, одинаково оценивает надежность наблюдений каждого из подсадных участников. Какова бы ни была эта надежность, изменение в его оценке, вызванное тем, что один из «сообщников» экспериментатора сказал, что расстояние составляет 40 см, должно быть больше, чем дополнительное изменение, вызванное вторым участником². Что, однако, противоречит результатам, поскольку второй подсадной участник больше повлиял на оценку, чем первый. Скорее всего, это эффект редукции диссонанса, вызванный дискомфортом из-за разногласий с большинством, который не может быть сведен к рациональному обучению.

Второй эксперимент имеет еще одну интересную особенность. Он проводился на нескольких «поколениях», в которых «сообщники» экспериментатора постепенно заменялись наивными участниками. Так, во втором поколении эксперимента с двумя подсадными лицами одно было заменено наивным участником из первого поколения, тогда как в третьем поколении второе подсадное лицо также было заменено наивным участником из более раннего поколения. В последующих поколениях все испытуемые были наивными субъектами, которые ранее имели дело либо с подсадными лицами, либо с другими участниками, имевшими дело с подсадными лицами, и т. д. Устроители эксперимента предпо-

² В числовых примерах, использовавшихся для иллюстрирования байесова обучения в главе XI, каждое новое подтверждающее свидетельство приносит меньшее увеличение вероятности, чем предыдущее. Эта «убывающая предельная ценность новой информации» — самый распространенный феномен.

лагали, что искусственно завышенные оценки будут поддерживаться до бесконечности, но они ошибались. Спустя шесть поколений в группах из трех человек и восемь поколений в группах из четырех — оценки сошлись на 10 см, то есть на цифре, которую давали изолированные участники. Вера в новое платье короля не сохраняется навечно. Если некоторые культурные убеждения, имеющие слабую опору в реальности, и сохраняются по прошествии времени, это может происходить в силу того, что расхождение с реальностью с трудом поддается наблюдению, или потому, что они находят опору в чем-то другом. Использование лотерей для идентификации хороших мест для охоты и рыбалки, как это принято в некоторых обществах, могло сохраниться до нашего времени вследствие своего религиозного значения.

ПЛЮРАЛИСТИЧЕСКОЕ НЕВЕДЕНИЕ

В самом начале этой главы я провел различие между двумя причинами, по которым люди в определенный момент времени могут придерживаться или выражать сходные убеждения: потому что они либо подвержены действию похожих условий (корреляция), либо влияют друг на друга (причинность). Особый случай первого вида составляют многочисленные примеры одновременных открытий — например, изобретение вычислений Ньютоном и Лейбницем приблизительно в одно и то же время. Хотя никто не знает наверняка, что это были за «сходные условия», идея могла «витать в воздухе». В качестве еще одного примера одновременного появления похожих идей рассмотрим идею нового платья короля. Сказка Ганса Христиана Андерсена была опубликована в 1835 году. Во втором томе «Демократии в Америке» (1840) Токвиль предложил похожую идею для объяснения очевидной устойчивости мнения большинства:

Время, события или индивидуальные усилия единичных умов могут в некоторых случаях подрывать или постепенно разрушать веру, причем внешне это может никак не проявляться. Никто не сражается в открытую с обреченной верой. Никто не собирает войско, чтобы пойти на нее войной. Сторонники тихо покидают ее один за другим до тех пор, пока лишь жалкое меньшинство не будет за нее цепляться. В этой ситуации ее царствование сохраняется. Поскольку враги продолжают хранить мир или выражать свои мысли только в тайне, пройдет немало времени прежде, чем они обретут уверенность, что произошла великая революция, а будучи в сомнениях, они бездействуют. Они наблюдают и хранят молчание. Большинство больше не верит, но по-прежнему выглядит так, как будто верит, и этого пустого призрака общественного мнения достаточно для того, чтобы остудить кровь возможным новаторам и заставить их замолчать из уважения.

В отрывке из «Старого порядка» (1856) Токвиля нечто похожее сказано о религии. В ходе Великой французской революции «те, кто сохранил прежнюю веру, побоялись остаться в одиночестве со своей преданностью и, страшась изоляции больше, чем ереси, присоединились к толпе, не разделяя ее убеждений. Так то, что было мнением части нации, стало считаться мнением всех и с тех пор стало казаться неоспоримым даже для тех, кто придал ему эту ложную видимость».

В этом пассаже Токвиль отсылает к убеждениям, которые люди *демонстрируют* (или стараются не отвергать), а не к тем, в которые на самом деле искренне верят. В этом отношении его анализ отличается от поведения в эксперименте с перемещающимся источником света и при одиночных публичных условиях эксперимента с измерением линии³. Тем не менее это не строгое и не мгновенное различие. Как я уже не раз отмечал, не всегда ясно, что означает «верить», что что-то имеет место. Даже при одиночных публичных условиях «вера» участников, сказавших, что подходящая линия А, может оказаться слабой. К примеру, они не захотели бы поставить на нее деньги. Кроме того, высказывание какого-то убеждения может при некоторых обстоятельствах породить тенденцию поддержать его (см. главу VII).

Современная психология переоткрыла догадку Токвиля, дав ей название «плюралистического неведения» (*pluralistic ignorance*). В крайних случаях никто не верит в истину какого-то высказывания, но полагают, что все остальные в него верят. В более реалистических случаях большинство людей в нее не верят, но убеждены, что верит большинство. Обе ситуации отличаются от патологических случаев, когда все публично высказывают определенное убеждение и при этом никто не придерживается его в частном порядке. Такая *культура лицемерия* в крайней степени была свойственна коммунизму, по крайней мере, в его финальной геронтократической стадии. Плюралистическое неведение и культура лицемерия могут опираться на один и тот же механизм, а именно на страх осуждения или наказания за выражение иных взглядов. Различие заключается в том, что при плюралистическом неведении осуждение является горизонтальным — разделяется только согражданами, которые ошибочно полагают, что им следует подвергать ostracism инакомыслящих, иначе они сами окажутся его жертвами. Культура лицемерия, наоборот, действует посредством вертикально налагаемого наказания: те,

³ На самом деле, эксперимент с движущимся светом был вдвойне публичным, оставляя поле для неискренности. Но двусмысленная природа задачи, напротив, упрощала искреннее или квазиискреннее принятие преувеличенных убеждений. Я полагаю, что основанием для такой процедуры было то, что исследование дальнейших поколений было бы трудно проделать в одиночных публичных условиях.

кто не выражает энтузиазма по поводу выполнения плана или ненависти к классовому врагу, могут потерять работу или с ними может случиться что-то похуже. Вертикальное наказание может затем порождать горизонтальные следствия, если люди будут избегать или наказывать инакомыслящих, чтобы самим не быть наказанными за инакомыслие.

Плюралистическое неведение также отличается от механизма, лежащего в основе синдрома стороннего наблюдателя, имевшего место при убийстве Китти Дженовезе. В условной версии последнего случая каждый индивид полагал, что его пассивность оправдана пассивностью других. Причина состояла не в социальном давлении, не в желании соответствовать нормам группы, поскольку 38 посторонних наблюдателей были слишком разобщены, чтобы образовать сообщество. Скорее пассивность оправдывалась следующим *заключением*: поскольку никто больше ничего не делает, ситуация не может быть серьезной. «Сырые данные» (крики девушки) заглушались этим выводом. Вскоре мы рассмотрим этот механизм более подробно. Здесь я только хочу заметить, что данная ситуация не предполагает плюралистического неведения, поскольку нет расхождения между тем, во что человек верит частным образом, и теми убеждениями, которые он приписывает остальным.

В качестве иллюстрации плюралистического неведения использовалась культура потребления алкоголя. Во многих американских кампусах существует культура пьянства среди студентов, особенно мужского пола. Многим из них не нравится такой высокий уровень потребления, но они соглашались пить, потому что ошибочно полагают, что и другие соглашались⁴. Их поведение по отношению к алкоголю соотнобразуется с тем, что они ошибочно считают типичным поведением на кампусе, а не с их частными взглядами. Еще один пример может быть почерпнут из эксперимента, в котором студентам предлагалось прочесть статью, сознательно написанную в таком бестолковом стиле, что она была практически нечитаемой, после чего их спрашивали, насколько хорошо они ее поняли и насколько хорошо, по их мнению, ее поняли другие. В одном случае студенты имели возможность подойти к экспериментатору и попросить у него помощи; в другом им было четко сказано, что они не могут этого сделать. Но даже в первом случае студенты не стали подходить к экспериментатору, потому что процедура была связана с риском поставить себя в неловкое положение. Однако каждый студент полагал, что, если он или она стеснялись обратиться за помощью, другие не делали этого, потому что понимали статью и не нуждались в помощи. Короче

⁴ Однако может быть верно и то, что большинство студентов не пьет и что большинство друзей большинства студентов пьет, если те, кто пьет, имеют больше друзей, чем те, кто не пьет.

говоря, студенты в этом случае были склонны считать, что другие понимают статью лучше, чем они сами. В другом случае это различие исчезло. Гипотетически этот эффект мог быть связан с «синдромом старшего ребенка». Как отмечалось в главе XVIII, мы все осознаем наши тревоги и страхи, но так как у нас нет прямого доступа к внутреннему миру других, мы склонны считать их более зрелыми и лучше владеющими собой.

Исследование потребления алкоголя на кампусе также показало, что со временем частные мнения, представления о мнениях других и поведение пришли во взаимное соответствие, из чего следует вопрос о *стабильности* плюралистического неведения. В действительности, оно может исчезнуть двумя способами: если ложные представления о других становятся истинными и если люди перестают их придерживаться. Если каждый человек принимает убеждения, которые он или она (ошибочно) приписывают другим, такое приписывание действительно может стать правдой. Вероятнее всего, это происходит посредством редукции диссонанса, вызванного либо дискомфортом от несогласия с большинством, либо дискомфортом из-за того, что приходится говорить одно, а верить в другое. Похоже, именно так и происходит с потреблением алкоголя на кампусе.

С другой стороны, ситуация может «растаться»⁵. Предположим, 20% членов группы показывают своим поведением, что они не разделяют общего убеждения и что оставшиеся 80% признают его только на словах, потому что им требуются 20% нонконформистов, чтобы стать таковыми самими. Предположим, в частности, что в группе из 100 человек есть 20 нонконформистов, 10 человек, готовых «открыться», если это сделают, по крайней мере, 25 человек; 15 человек, которые пойдут вслед за (как минимум) 35, и 55, которые присоединятся, если не менее 50 человек раскроют свое истинное отношение. Известно, что культура большинства устойчива. Однако предположим, что 5 самых закоренелых конформистов уходят или умирают и их замещают 5 нонконформистов. В этом случае большинство развалится. 25 нонконформистов создадут условия для того, чтобы к ним присоединились еще 10 человек; полученные таким образом 35 привлекут еще 15, тем самым создав требуемый порог для того, чтобы к ним присоединились оставшиеся 50. Вместо того чтобы считать этот процесс *распадом конформизма*, мы также можем рассматривать его как *лавинообразный рост нонконформизма*. В дальней-

⁵ Исчезновение ошибочных представлений в эксперименте с движущимся светом также является формой раскрытия, вызванной тем фактом, что каждое поколение подопытных использует свои собственные «сырые данные» для корректировки оценки расстояния в сторону уменьшения по сравнению с тем, что они услышали от других.

шем мы обнаружим подобную динамику в коллективных действиях (см. главу XXIV).

Конформизм может распадаться и многими другими путями. Случай маленького ребенка из сказки Андерсена отражен в эксперименте с соотношением линий: стоит только *одному* подсадному лицу высказать верное мнение, что D ближе всего по длине к линии В, конформизм практически исчезает. В качестве другого примера рассмотрим распространенную в Англии и во Франции веру в то, что король своим прикосновением может излечивать золотуху. Реформация подорвала эту веру, поскольку католикам во Франции и англиканцам в Англии приходилось теперь объяснять, почему подобные свидетельства из других стран фальшивы. Но признание возможности грандиозного коллективного заблуждения оказалось слишком опасным, поскольку якобы недостоверные доказательства, использовавшиеся для поддержания этой веры в других странах, не слишком отличались от тех, что применялись в их собственных.

Еще один механизм распада конформизма — публикация опросов общественного мнения. До референдума 1972 года о вступлении Норвегии в «Общий рынок» (как это тогда называлось) правительство, основные политические партии и главные газеты единодушно выступали в поддержку этого решения. Хотя, как показал референдум, существовало народное большинство, настроенное против такой инициативы, если бы не опросы, каждый индивидуальный оппонент считал бы себя членом незначительного меньшинства. Без опросов результаты референдума, скорее всего, были бы иными. Некоторые из тех, кто выступал против, воздержались бы от голосования, поскольку исход считался предрешенным. Кроме того, движение, образованное с целью убедить колеблющихся, было бы маленьким и лишенным влияния. В период между введением всеобщего избирательного права и ростом опросов общественного мнения, по всей видимости, существовало обширное поле для плюралистического неведения по политическим вопросам.

СЛУХИ, СТРАХИ И НАДЕЖДЫ

Еще одна сказка Ганса Христиана Андерсена «В этом нет никаких сомнений» иллюстрирует, как «одно маленькое перышко может легко вырасти в пять куриц» путем последовательных преувеличений. Насколько мне известно, исследование образования и распространения слухов не сильно продвинулось. Основной вклад в него внесли французские (и некоторые англо-американские) историки, которые пошли по пути, намеченном Жоржем Лефевром в его новаторском исследовании «великого страха» 1789 года. В дополнение к этому они изучали слухи, которые практически все были ложными и относились:

- к возвращению Наполеона после двух поражений в 1814 и 1815 годах;
- полной реставрации *старого порядка*;
- уравнительному разделу имущества рабочими накануне Революции 1848 года;
- массированному вторжению в Германию в марте 1848 года обедневших французских рабочих, занимавшихся грабежами, поджогами и убийствами;
- заговору врачей с целью «отравить народ»;
- заговору духовенства и знати с целью «уморить народ голодом»;
- предстоящему снижению налогов;
- предстоящему повышению налогов;
- *франтирерам*, расстреливавшим немецких солдат с крыш во время вторжения Германии в Бельгию в 1914 году;
- десяткам тысяч русских солдат, пополнивших войска союзников в августе 1914 года.

На основании этих исследований, я думаю, можно выявить общие проблемы и, возможно, сделать определенные выводы. Позвольте, однако, начать с эмпирических наблюдений. Во-первых, идея одного пера, превращающегося в пять куриц, не преувеличивает усиливающий эффект слухов. После восстания рабочих в Париже в июне 1848 года два человека, замеченные на обочине сельской дороги, превращались в 10, 300, 600 человек, по мере того как история переходила из уст в уста, пока наконец не появился слух о том, что три тысячи «уравнителей», приверженцев имущественного передела (*partageux*), занимаются мародерством, убийствами и поджогами. Для отражения угрозы было отправлено 30 тысяч солдат. Расследование показало, что один из этих двоих был умалишенным, а второй — отцом, который за ним ухаживал. В тот же самый период крестьянин выдумывал басни, чтобы попугать ребенка; вскоре после этого более тысячи человек взяли за оружие, чтобы сражаться с несуществующими «разбойниками».

Во-вторых, иногда возможно с некоторой точностью определить происхождение, скорость и механизм распространения слухов. «Великий страх» 1789 года возник одновременно, но независимо (скоординированный временем сбора урожая) в семи разных местах, а затем распространился (со средней скоростью четыре километра в час) и охватил большую часть страны. Слух о вторжении в Германию французских безработных рабочих в 1848 году распространялся с похожей скоростью. В обоих случаях делались оценки, насколько замедлялось распространение слуха при пересечении гористой местности и ночью. Часто действия,

к которым побуждала молва, как мы вскоре увидим, сами становились ее возбудителями. Кроме того, слух зачастую распространялся индивидами, искренне или неискренне утверждавшими, что они владеют информацией: официальными лицами, бродягами, странствующими торговцами, солдатами, возвращавшимися с фронта. Слухи также распространяются благодаря звону колоколов, который может быть слышен в соседних деревнях. В более поздние времена важным источником слухов стали газеты и письма. Официальное опровержение слуха лишь подпитывало его.

Лефевр резюмировал часть своего объяснения «великого страха», заметив, что «народ сам себя путал». Вера одних крестьян в приближение разбойников приводила к мобилизации войск, которых другие крестьяне издали принимали за разбойников. Когда в деревне били в колокола, отряды, посылавшиеся соседними деревнями, принимали за врагов. В 1848 году предупредительный выстрел пушки в одной французской деревне был истолкован в соседних как шум битвы. Когда в марте 1848 года слухи о грядущем нашествии французских нищих достигли Германии, дорожные рабочие на французской стороне Рейна в спешке стали переправляться через реку, чтобы вернуться домой к семьям. Другие, наблюдая за этим издали, могли принять их за приближающихся французов.

Во многих случаях содержанием слуха является существование заговора против народа, устроенного правительством или элитой. Естественное событие — плохой урожай, серия пожаров, вспышка холеры — приписывается интенциональной инстанции. (Как отмечалось ранее, люди часто полагали, что элитами движет злонамеренность, а не корысть.) Та же самая ментальная привычка заставляет властей приписывать общий интенциональный источник тому, что на самом деле является слухами, возникшими независимо друг от друга. Из того факта, что похожие слухи одновременно появились в разных частях страны, власти справедливо делали вывод о наличии общей причины. Вместо того чтобы идентифицировать эту причину как единую для всех объективную ситуацию (например, неурожай), власти ошибочно заключали, что это было делом рук интенционального агента. Слухи о заговорах были неотделимы от веры в то, что слухи порождаются заговорами.

Комментируя тенденцию слухов разрастаться, Монтень, возможно, предложил первый анализ микромеханизмов их распространения:

Между ничем и ничтожнейшей из существующих в мире вещей расстояние большее, чем между этой ничтожнейшей и величайшей. Так вот те, кто первыми прослышали о некоем удивительном явлении и начинают повсюду трезвонить о нем, отлично чувствуют, встречая недоверие, где в их утверждениях слабое место, и всячески стараются заделать прореху, приводя ложные свидетельства. <...> Вначале чье-то личное заблуждение становится общим, а затем уже общее — личным.

Вот и растет эта постройка, к которой каждый прикладывает руку, так что самый дальний свидетель события оказывается лучше осведомленным, чем непосредственный, а последний человек, узнавший о нем, — гораздо более убежденным, чем первый. Все это происходит самым естественным образом, ибо каждый, кто во что-то поверил, считает актом великодушия убедить в том же другого человека и ради этого, не смущаясь, добавляет кое-что собственного сочинения, если, по его мнению, это необходимо, чтобы во всеоружии встретить сопротивление другого.

Это до некоторой степени снисходительное объяснение, поскольку Монтень приписывает распространителю слухов всего лишь желание убедить других в том, во что он верит сам. В своем анализе «великого страха» Лефевр указывает на другие, более темные мотивы. Некоторые индивиды могут быть заинтересованы в преувеличении опасности, чтобы не быть обвиненными в трусости. Показать недоверие означало одновременно вызвать обвинения в служении интересам контрреволюции, в желании усыпить бдительность на пике угрозы и оскорбить самолюбие тех, кто бьет тревогу. При анализе слухов в XIX веке утверждалось, что бродяги, как правило, распространяли те из них, которые слушатели хотели слышать: о возвращении Наполеона I — чтобы доставить удовольствие его сторонникам, и о болезни Наполеона III — чтобы удовлетворить его противников. Точно так же торговцы пускали самые сенсационные легенды, чтобы привлечь более широкую аудиторию. Что до молвы о бельгийских франтирерах, они казались неоспоримыми, раз были использованы для кровавых репрессий. Как еще, по мнению немцев, они могли оправдать свои зверства? А если слухи распространяются ранеными солдатами, вернувшимися с фронта, кто решится их опровергнуть?

Авторы, пишущие о слухах, делают акцент на том, как они возникают в «коллективном сознании» в форме предварительной схемы, которая может быть активирована самыми незначительными событиями. Популярное верование, касающееся злонамеренности элит и их склонности устраивать заговоры против народа, входит в ряд общих условий, делающих слухи достоверными, в дополнение к перечисленным выше специфическим факторам. Вера в ненасытную потребность правительства в повышении налогов и в большем количестве солдат также приводила к тому, что внешне нейтральные действия, такие как статистические исследования, вызывали слухи о грядущем росте налогов и объявлении армейского призыва. В 1914 году представления немцев о бельгийском сопротивлении определялись схемой *франтирера*, сложившейся во время Франко-прусской войны 1870–1871 годов. Эти схемы обычно имеют некоторые основания в реальности, даже если специфические верования, которые они внушают, их лишены.

Я несколько раз цитировал поговорку «мы легко верим в то, на что надеемся, и в то, чего боимся». Слухи, которые я привел, иллюстрируют обе возможности. Возвращение Наполеона, снижение налогов по случаю смены режима и присоединение русских солдат к силам союзников — все это свидетельствует о силе надежды. (В случае Наполеона, впрочем, были и те, кто страшился его возвращения.) Но все же один историк слухов во Франции XIX века заявил, что «в целом слухи гораздо более пессимистичны, чем эйфоричны»: страх преобладает над надеждой. Чтобы проверить данное утверждение, можно подсчитать эпизоды принятия желаемого за действительное и принятия нежелаемого за действительное и посмотреть, действительно ли последнее преобладает над первым. Вероятно, такая попытка была бы безнадежной, так как невозможно сделать репрезентативную выборку и трудно понять, какую важность придавать каждому эпизоду.

Более перспективный подход может дать различие убеждений и квазиубеждений, которое я предложил во вступлении ко второй части; главное различие заключается в том, что в качестве предпосылки к действию используются только первые. Если мы изучим примеры формирования слухов, окажется, что *менять свое поведение людей заставляют почти исключительно слухи, основанные на страхе*. В эпизодах образования слухов во Франции XVIII–XIX веков крестьяне собирали урожай до того как он созреет, чтобы помешать его уничтожению «разбойниками», продавали зерно, даже не оставляя ничего на семена, потому что боялись, что все будет полностью конфисковано, прятали ценности, когда ползли слухи, что их обложат налогом, женились, чтобы избежать призыва в армию, и запасались солью, когда появлялись слухи о налоге на соль. Молва о приближении разбойников имела решающее значение в ходе Французской революции. Среди прочего она побуждала крестьян нападать на господские замки, то есть совершать действия, которые, в свою очередь, вызвали к жизни декреты 4 августа 1789 года, уничтожившие феодализм. Этнические бунты сходным образом часто порождаются слухами о грядущем нападении со стороны другой группы⁶. Принятие нежелаемого за действительное, порождаемое страхом, по всей видимости, обладает значительной способностью влиять на поведение. Напротив, принятие желаемого за действительное, порождаемое надеждой, как представляется, чаще используется по причине своей способности к утешению. Данная асимметрия кажется очень устойчивой. Хотя она не доказывает, что слухов, основанных на страхе, больше, она демонстрирует их важность в объяснении поведения.

⁶ Роль эмоций в бунтах подтверждается задокументированной тенденцией, согласно которой они чаще происходят в жаркую погоду.

Главное исключение из этой асимметрии, с которым я столкнулся, имеет место на финансовых рынках, где подпитываемые надеждой слухи часто основываются на спекулятивных пузырях. «Иррациональное изобилие» 1990-х, конечно, использовалось как основа для — крайне саморазрушительных — действий. В то же время финансовые рынки часто демонстрируют подверженность иррациональному пессимизму. Мало что известно, однако, о механике и динамике *интерактивного* формирования убеждений на этих рынках. Представляется, что там имеет место сложное взаимодействие между теми, кто принимает решения на основании слухов, теми, кто действует, ориентируясь на динамику цен, вызванную слухами, и теми, кто руководствуется слухами о возможном изменении цен. Хотя некоторые агенты действуют согласно рациональным убеждениям, опираясь на факты, некоторые из этих фактов являются результатом иррациональных действий, совершенных под воздействием необоснованных слухов.

Асимметрия, как представляется, ограничивается интерактивным образованием убеждений. На индивидуальном уровне принятие желаемого за действительное, конечно же, в меньшей степени способно влиять на поведение, чем принятие нежелаемого за действительное. Более того, повседневные наблюдения подсказывают, что на индивидуальном уровне последний механизм встречается реже первого. Кажется, что интерактивная природа слухов не до конца понятными нам путями порождает паттерны, которые отличаются от тех, что мы наблюдаем, когда индивиды формируют свои убеждения в изоляции друг от друга.

ИНФОРМАЦИОННЫЕ КАСКАДЫ

Слухи также возникают посредством полностью рационального формирования убеждений, через механизм, известный как «информационные каскады». Предположим, что каждый индивид в группе имеет доступ к информации о каком-то частном вопросе. Все формируют свои взгляды последовательно, каждый полагается на свои частные сведения и мнения, выраженные его предшественниками (если таковые есть) по порядку. Так, каждый житель деревни может располагать какими-то частными данными, о присутствии разбойников, и использовать их вместе с услышанным от других для формирования мнения, которое он передаст дальше. Такую же динамику может иметь поименное голосование, если обсуждаемый вопрос решается только на основании убеждений, а не предпочтений. Каждый член собрания будет принимать решение не только на базе имеющейся у него информации, но также с учетом сведений, полученных при голосовании его предшественников. В качестве третьего примера рассмотрим журнального рецензента, готовящего отзыв на ста-

тью, который узнаёт, что рецензент другого журнала выступил против нее (но не узнаёт почему).

В этой ситуации выводы, сделанные другими в ходе выработки позиции, используются как косвенные вводные данные для выработки своих собственных взглядов, при этом неизвестно, какие прямые сведения (частную информацию) использовали те, чтобы прийти к собственным заключениям. Случается, что рациональные индивиды в результате вырабатывают неверные убеждения, в то время как они пришли бы к правильным выводам, будь у них доступ к «сырым данным», которыми располагали их предшественники, а не только к готовым выводам последних. В упомянутом примере второй рецензент мог бы заметить предвзятость или ошибочное суждение, если бы прочел отзыв первого рецензента. Но если ему известен лишь вывод, содержащийся в первом отзыве, и тот факт, что он опубликован в очень уважаемом издании, он должен рационально учитывать отрицательное мнение первого рецензента наряду со своими собственными оценками. Если последние благоприятны, но лишь в небольшой мере, он в конечном счете может порекомендовать отказаться от публикации этой работы. Третий рецензент с благоприятным (в высшей степени) личным отношением также может (рационально) способствовать отказу, если узнает, что до него так поступили двое других рецензентов. Но исход может оказаться субоптимальным (с точки зрения целей научного сообщества), поскольку второй и третий рецензенты выступали за публикацию, а первый был против нее лишь в незначительной степени⁷. Если бы рецензенты прочитали статью в обратном порядке, вывод был бы иным («зависимость от маршрута»).

Библиографические примечания

Эксперименты с нахождением соответствия линий были впервые проведены Соломоном Эшем и описываются в любом учебнике по социальной психологии, например в «Общественном животном» Э. Аронсона (Аронсон Э. Общественное животное. Введение в социальную психологию. М.: Аспект Пресс, 1998). Эксперименты с движущимся светом описаны в работе Р. С. Джейкобса и Д. Т. Кэмпбелла «Сохранение произвольной традиции в нескольких поколениях лабораторной микрокультуры» (Jacobs R. C., Campbell D. T. The perpetuation of an

⁷ С этой точки зрения становятся понятны многие практики, нарушающие конформность. Когда (как это практиковалось в Норвегии) преподаватель университета, ставящий оценки на экзамене, направляет работы для оценки внешнему специалисту, он не показывает ему свои оценки. По той же самой причине, обращаясь за консультацией ко второму врачу, не следует говорить ему диагноз, поставленный первым доктором.

arbitrary tradition through several generations of laboratory microculture // Journal of Abnormal and Social Psychology. 1961. P. 649–658). Касательно распития алкоголя на кампусах см. Д. А. Прентис и Д. Т. Миллер «Плюралистическое неведение и употребление алкоголя на кампусах: некоторые последствия ошибочного восприятия социальной нормы». (Prentice D. A., Miller D. T. Pluralistic ignorance and alcohol use on campus: Some consequences of misperceiving the social norm // Journal of Personality and Social Psychology. 1993. Vol. 64. P. 243–256). По поводу эксперимента со студентами, которым дали читать непонятную статью см. Д. Т. Миллера и С. Макфарланда «Плюралистическое неведение: когда сходство интерпретируется как различие» (Miller D. T., McFarland C. Pluralistic ignorance: When similarity is interpreted as dissimilarity // Journal of Personality and Social Psychology. 1987. Vol. 53. P. 298–305). Сценарий распада основывается на книге Т. Курана «Частные истины, публичные заблуждения» (Kurana T. Private Truths, Public Lies. Cambridge, MA: Harvard University Press, 1995). Наблюдением о влиянии Реформации на веру в то, что король может исцелять, я обязан книге М. Блоха «Короли-чудотворцы» (Блок М. Короли-чудотворцы: Очерк представлений о сверхъестественном характере королевской власти, распространенных преимущественно во Франции и в Англии. М.: Языки русской культуры, 1998). Исследования образования слухов взяты мною из работ Ж. Лефевра «Великий страх 1789 года» (Lefebvre G. La grande peur de 1789. Paris: Armand Colin, 1988); Ф. Пло «Из уст в уста. Рождение и распространение слухов во Франции XIX века» (Ploux F. De bouche à oreille: Naissance et propagation des rumeurs dans la France du XIXe siècle. Paris: Aubier, 2003); Р. Сеневали «„Ложная французская тревога“: революционная паника в Бедене, 1848» (Cenevali R. The «false French alarm»: Revolutionary panic in Beden, 1848 // Central European History. 1985. Vol. 18. P. 119–142); М. Блок «Размышления историка о ложных слухах военного времени» (Bloch M. Réflexions d'un historien sur les fausses nouvelles de guerre // Revue de synthèse historique. 1921. Vol. 33. P. 13–35; «Правда и ложь в Большой войне» (Prochasson C., Rasmussen A. (eds.). Vrai et faux dans la Grande Guerre. Paris: Editions La Découverte, 2004). Подробное перечисление и анализ роли слухов в этнических бунтах можно найти в книге Д. Горовитца «Смертельный этнический бунт» (Horowitz D. The Deadly Ethnic Riot. Berkeley: University of California Press, 2001). Касательно слухов на биржах см. А. М. Роуз «Слухи на биржах» (Rose A. M. Rumor in the stock market // Public Opinion Quarterly. 1951. Vol. 15. P. 61–86). Введение в механизм информационных каскадов можно найти в работе С. Бихчандани, Д. Хиршлайфера и И. Уэлча «Участие на поведении других: конформность, повествование и информационные каскады» (Bikhchandani S., Hirshleifer D., Welch I. Learning from the behavior of others: Conformity, fads, and informational cascades // Journal of Economic Perspectives. 1998. Vol. 12. P. 151–170).

XXIV. Коллективное действие

ЖИВОЙ ФЛАГ

В «Днях на озере Вобегон» Гаррисон Кейлор описывает День флага в своем выдуманном городе. Герман, организатор парада, принес некоторое количество синих, красных и белых кепок и раздал их жителям города, чтобы они могли пройти по улицам как живой флаг, тогда как сам он будет стоять на крыше центрального здания, чтобы это заснять. Сразу после войны люди с удовольствием сотрудничали, но позднее у них появились сомнения:

Одной из причин для обиды было то, что никто из них не увидит Флага, в котором они участвовали; фотография в газете будет черно-белой. Только Герман и мистер Хенсон увидят реальный Флаг, да еще несколько мальчишек, слишком невысоких, чтобы понадобиться внизу. Люди хотели получить возможность подняться на крышу и увидеть зрелище своими глазами.

«Как вы можете подняться наверх, когда вы должны быть внизу? — сказал Герман. — Вы подниметесь наверх, и вам будет не на что смотреть. Разве недостаточно знать, что вы играете свою роль?»

На День флага в 1949 году, едва Герман сказал: «Вот! Стойте так!», один из красных рванул — пробежал четыре пролета по лестнице на крышу, вытянул шею и долго на все смотрел. Хотя после него в толпе осталась дыра, это было чудесное зрелище. Живой флаг заполнил три улицы внизу. Идеальный флаг! Какой яркий красный цвет! Он не мог глаз от него оторвать. «Иди обратно! Нам нужна фотография!» — заорал на него Герман. «Как это выглядит?» — кричали ему люди. «Невероятно! Не могу описать», — отвечал он.

Тогда все захотели взглянуть. «Нет!» — воскликнул Герман, но они проголосовали, и решение было единодушным. Один за другим члены Живого флага поднимались на крышу и любовались им. Это было чудесно. Глаза наполнялись слезами, в голову приходили мысли о величии страны и о месте озера Вобегон во всем этом. Хотелось стоять там весь день и просто впитывать это в себя. Так что час спустя, когда наверху побывали только сорок человек из пятисот, оставшиеся начали нервничать. «Пошевеливайтесь! Что вы там копаетесь? Вы уже все посмотрели! Спускайтесь и дайте другим посмотреть!» Герман посылал людей группами по четыре человека, потом по десять, но через два

часа Живой флаг стал Сидячим, а потом начал распадаться, потому что те, кто посмотрел, подумывали о том, чтобы пойти домой на ужин, что приводило в бешенство тех, кто еще не успел.

«Еще десять минут», — закричал Герман. Но десять минут превратились в двадцать, потом в тридцать, люди потихоньку расходились, и Флаг, открывшейся взору последнего зрителя, был Флагом, простреленным из пушки.

В 1959 году руководство Днем флага перешло к сыновьям Кнута. Герман отдал им коробки с кепками. С тех пор им удалось устроить не сколько хороших флагов, хотя чаще всего явка была невысока. Для хорошего флага нужно по меньшей мере 400 человек. В некоторые годы Кнуты устанавливали правило «не смотреть», в другой раз устраивали лотерею. Один раз они экспериментировали с большим зеркалом, которое два человека держали у края крыши, но, когда люди отступали и смотрели вверх, Флаг, конечно же, исчезал¹.

Вобегонианцы столкнулись с проблемой коллективного действия, хотя и в необычной форме. Каждый испытывает искушение подняться на крышу, а в более позднем эксперименте взглянуть в зеркало. Но если они все это сделают, флаг *распадется* или исчезнет и смотреть будет не на что. Для решения проблемы они прибегают к классическим технологиям координации: ввести правило «не смотреть», чтобы помешать «безбилетникам», смотреть по очереди, провести лотерею. Первое решение не слишком удовлетворительно, потому что нет большого смысла организовать флаг, который никто не может увидеть². Другие решения дают возможность посмотреть, но лишь немногим, потому что иначе смотреть будет не на что. Они, однако, требуют централизованной координации, опирающейся на организации и санкции. Задача в том, чтобы понять, как люди иногда решают проблемы коллективного действия с помощью децентрализованного действия.

РАСПАД И ЛАВИНООБРАЗНЫЙ РОСТ

Если предоставить людей самим себе, то флаг в буквальном смысле распался бы. Распад в метафорическом смысле был также показан в экспериментах с общественными благами. Людям дают некоторую сумму денег и говорят, что, если они вложат ее в общий фонд, сумма увеличит-

¹ New York: Viking, 1985.

² Конечно, его увидит Бог. Во многих средневековых церквях прекрасные капители расположены так высоко, что их детали невозможно рассмотреть без бинокля, который еще не был изобретен, когда они строились. Но это не важно, потому что их может видеть Бог. Вобегонианцы, кажется, не испытывают подобного религиозного рвения.

ся, а потом эту увеличенную сумму в равных долях распределяют между всеми, в том числе теми, кто сам не вносил вклад в общий фонд. (Это «дилемма заключенного» с участием многих игроков, многократно разыгрываемая между одними и теми же подопытными.) Если бы все участники были рациональны и своекорыстны, никто не вложил бы ничего. В действительности, даже в анонимных взаимодействиях участники сначала вкладывают 40–60% полученных денег. В последующих раундах вклады стабильно падают, пока не стабилизируются в районе 10%. Одно из возможных объяснений заключается в том, что все люди готовы вкладывать тогда и только тогда, когда есть определенный уровень вкладов других. Этот уровень может варьироваться от индивида к индивиду. Некоторым может требоваться более высокий уровень вкладов, чтобы сподвигнуть их на участие, тогда как другие вкладывают, если хотя бы немногие сделали свой взнос или даже когда его не сделал никто. Тогда распад может произойти, как показано на рис. XXIV.1.

В самом начале люди полагают, что размер вкладов других будет в среднем составлять OD . Некоторые индивиды — крайние эгоисты и не вносят ничего, но есть и те, кто, делая вклад, руководствуется нормами справедливости. Итоговая сумма их вкладов дает OH . Эта меньшая сумма ($OH = OC$) затем кладется в основу ожиданий в следующей игре. Некоторые из тех, кто был готов отдать деньги, когда ожидаемый вклад составлял OD , теперь выбыли, так что в результате вклад составит OG . В последующих раундах будет выпадать все больше вкладчиков по мере того, как уровень вкладов будет падать ниже того, при котором они готовы делать свои вложения, пока процесс не стабилизируется на $OA = OE$. В этой точке реальные и ожидаемые вклады совпадут.

Можно также наблюдать противоположное явление не «распада», а «лавинообразного возрастания». Примером могут послужить толпы, собиравшиеся каждое воскресенье в Лейпциге перед сносом Берлинской стены в 1989 году. Во время бунтов и революций небольшое количество первопроходцев также может вырасти в движение, которое низвергнет правительство. Но прежде чем попытаться понять природу этих явлений, нужно подробнее описать структуру проблем коллективного действия. Пока мы можем абстрагироваться от последовательного аспекта только что приведенных примеров и охарактеризовать проблему коллективного действия в категориях отношений между индивидуальными выборами и коллективным результатом³.

³ Я буду рассуждать так, как будто индивидуальные выборы бывают только бинарными (либо сотрудничать, либо нет) и как будто исходы постоянно варьируются (общественные блага, такие как чистый воздух, могут обеспечиваться в большей или в меньшей степени). На самом деле, индивиды могут

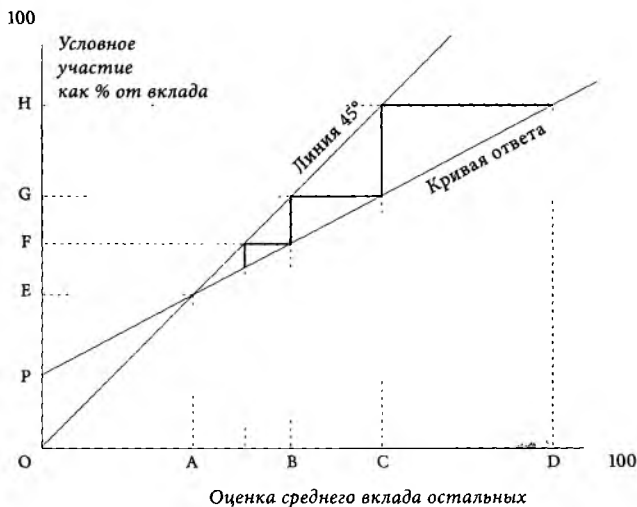


РИС. XXIV.1

Проблема тесно связана с вопросами, которые я затрагивал в предыдущих главах. Проблемы коллективного действия сопровождаются негативными или позитивными экстерналиями, которые порождают «дилеммы заключенного» с участием многих лиц или связанные с ними ситуации. Для характеристики проблемы коллективного действия я ограничусь подмножеством мотиваций, которые могут вдохновлять агентов в таких ситуациях. В частности, предполагается, что агенты мотивированы только затратами на участие и личными выгодами, которые они могут извлечь из исхода коллективного действия. Например, вышедший на забастовку рабочий, как предполагается, должен учитывать только риск потери работы или зарплаты на период забастовки, а также перспективу получения более высокой заработной платы (для себя), если забастовка увенчается успехом. Мы же, наоборот, должны обра-

отличаться друг от друга размерами вклада, который они вносят, а не только тем, сотрудничают ли они. Я игнорирую это осложнение. Кроме того, некоторые общественные блага «комковаты» или дискретны. Если индивиды в каком-то сообществе отстаивают сохранение местной школы, школа либо останется, либо ее закроют. Это осложнение может быть тонким образом обойдено, если истолковать исход как постоянно варьирующуюся вероятность предоставления общественных благ.

тяться к более широкому спектру мотиваций, чтобы объяснить выбор кооперативной стратегии в таких ситуациях.

Обеспечение позитивной экстерналии несет для агента прямые издержки или риски. В эксперименте с общественными благами в качестве издержек выступает жертвование частью полученной суммы. В маршах в Лейпциге это был риск попасть под дубинки полиции. В то же время агент обеспечивает небольшую материальную выгоду тем, с кем он взаимодействует. В эксперименте каждый из остальных участников получал небольшую долю от своего вноса. Подопытный также извлекал материальную выгоду (он получал ту же самую долю), но очевидно, что она была меньше, чем издержки. В Лейпциге каждый новый участник демонстрации становился дополнительным объектом внимания для полицейских сил (если предположить, что их численность была неизменной) и тем самым способствовал тому, что вероятность избиения любого из его товарищей по демонстрации уменьшалась. Таким образом, если мы ограничимся прямыми затратами и прибылями от кооперации, индивидуальная измена будет превосходить индивидуальное сотрудничество. В то же время в обоих случаях универсальное сотрудничество превосходит универсальный отказ от него. Если все внесут полученные вклады в общий фонд, они получают обратно ту же умноженную сумму. Если все жители Лейпцига выйдут на улицы, риск того, что кто-то из них будет избит, снижается до 0, тогда как шансы падения ненавистного режима приближаются к 1.

ТЕХНОЛОГИЯ КОЛЛЕКТИВНОГО ДЕЙСТВИЯ

Мы можем изобразить эту ситуацию на графике. На рис. XXIV.2 видно, как в группе $n + 1$ индивидов варьируется выплата данному индивиду в зависимости от его собственного поведения и действий остальных. Поведение других отмечено на горизонтальной оси, которая отображает количество участников кооперации (среди прочих). Если индивид также относится к участникам кооперации, его полезность, измеренная по вертикальной оси, обозначается вдоль R на линии АВ на графиках. Если он относится к уклоняющимся от сотрудничества, его полезность измеряется по оси L на прямой ОС. Прямые R и L пересекают вертикальные оси в порядке, который определяет обычную (для двух игроков) «дилемму заключенного»: наиболее предпочтительный исход — односторонний отказ от сотрудничества («безбилетничество»), следующий по желательности исход — универсальное сотрудничество, третий по оптимальности — универсальный отказ от сотрудничества, а худший — одностороннее сотрудничество (когда вас используют). В случае игры двух лиц доминирующей стратегией является отсутствие сотруд-

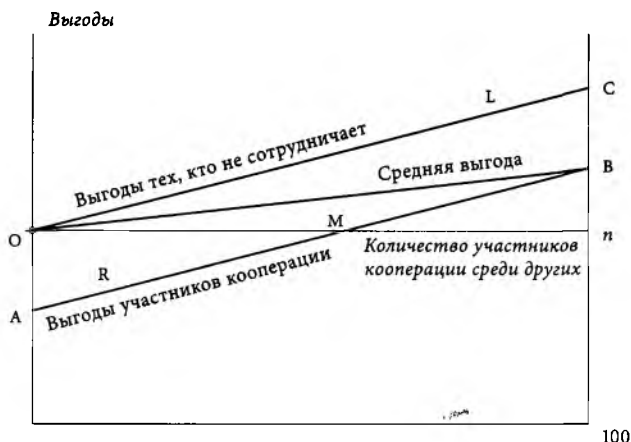


РИС. XXIV.2

ничества, поскольку прямая L повсюду проходит выше, чем R . В отличие от игры с двумя игроками мы, тем не менее, можем выделить некоторое число участников кооперации M , которым выгоднее сотрудничать даже при наличии «безбилетников», которые на них наживаются. Прямая OB показывает *среднюю* прибыль для всех, в том числе тех, кто не сотрудничает, в виде функции от числа участников. Поскольку количество агентов постоянно, OB также отражает *общую* выгоду от кооперации.

Ситуация, показанная на рис. XX.2, отражает особый случай. Он подразумевает, что стоимость кооперации, измеряемая расстоянием между кривыми L и R , постоянна. В других случаях цена кооперации возрастает по мере подключения к ней большего количества людей. Чем больше слушателей присоединяется к телефонным опросам на местном радио, тем больше перегружаются линии и больше времени уходит на то, чтобы дозвониться. Может так случиться, что последние присоединившиеся⁴ действительно снижают среднюю выгоду, потому что их затраты на участие превышают сумму выгод, которые они производят для всех

⁴ Здесь и в других местах слова «первый», «средний» и «последний» могут отсылать к моментам, когда присоединяется следующий сотрудничающий, как при формировании революционного движения. Но они также могут относиться к одновременным актам сотрудничества, как при голосовании. Сказать, что последние избиратели мало что меняют, — значит указать, что выгода, созданная ситуацией, в которой голосует каждый, почти равна выгоде в ситуации, когда голосуют почти все.

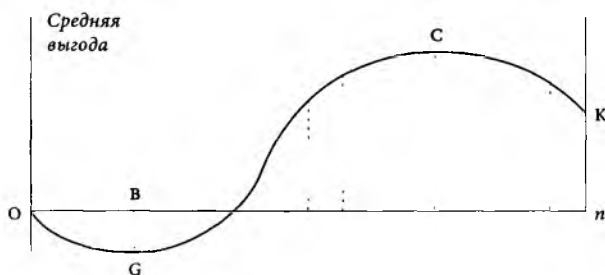


РИС. XXIV.3

остальных (и для себя). Цена также может быть высокой первоначально, но позднее способна снижаться. По мере того как все больше людей выходило на демонстрации в Лейпциге, правительственным силам приходилось распределяться все более «тонким слоем». В этом случае первые сотрудничающие могли так же и по тем же причинам ухудшить ситуацию (с точки зрения сокращения средней выгоды).

На рис. XXIV.2 также видно, что выгоды от сотрудничества, заданные прямой L , являются линейной функцией от количества участников кооперации. Каждый новый участник в одинаковой степени способствует приросту благосостояния всех. Увеличение предельных выгод может быть проиллюстрировано очисткой пляжа от мусора: последняя убранный бутылка производит больший эстетический эффект, чем предпоследняя. Простой пример — звонок в муниципалитет по поводу рытвины в городском районе, где проживает средний класс: первый из тех, кто нашел время позвонить, дает вероятность 0,4 того, что выбоина будет заделана, второй увеличивает вероятность до 0,7, третий — до 0,8, четвертый — до 0,85, пятый — до 0,88 и т. д. Иногда и первые, и последние участники кооперации мало что добавляют, тогда как те, кто оказывается в середине, более эффективны. Несколько революционеров или участников забастовки почти ничего не меняют, а когда почти все присоединились к движению, то, что в него вступило еще несколько ранее не присоединявшихся человек, не имеет значения. Вероятно, для социальных движений такой паттерн является типичным.

Предельные выгоды от сотрудничества в определенном диапазоне могут даже оказаться отрицательными. Одностороннее разоружение только ухудшит положение всех государств, если создаст вакуум власти,

который придется заполнять, развязав войну. Изолированные бунты могут дать властям повод обрушиться не только на реальных бунтовщиков, но и на потенциальных. И наоборот, участников кооперации может оказаться слишком много. Представьте себе, что во время войны все захотят вступить в армию, так что некому будет работать в промышленности и война будет проиграна. Если во время пикника все захотят участвовать в приготовлении обеда, слишком большое количество поваров может испортить блюдо.

Как показывают эти наблюдения, технология коллективного действия меняется от случая к случаю. В дальнейшем я остановлюсь на примере, представленном на рис. XXIV.3, который, как мне кажется, типичен для социальных движений, стремящихся добиться определенных изменений в политике. Первые участники платят высокую цену, подвергаются большому риску и дают небольшую выгоду остальным. В действительности, они могут даже нанести им вред, а не пользу. Их результирующий вклад отрицателен. Последние участники тоже не приносят большой пользы. В некоторых случаях, как я указывал, их затраты могут сокращаться. В других — все, кто ведет борьбу за определенное дело, могут нести значительные потери или подвергаться риску до тех пор, пока противник не капитулирует. Те, кто вступил во французское Сопротивление в 1944 году, нанесли не такой уж большой ущерб немцам, при этом подвергнув свою жизнь большому риску.

ПРЕОДОЛЕВАЯ ПРОБЛЕМУ «БЕЗБИЛЕТНИКА»

Я определил проблему коллективного действия в категориях *целенаправленных рациональных эгоистических мотиваций*. Как правило, этого не достаточно, чтобы подтолкнуть к сотрудничеству, поскольку независимо от того, что делают другие, индивиду выгоднее не участвовать. Позвольте упомянуть два исключения из этого утверждения.

Первое возникает, когда *одни и те же* агенты *множественно* сталкиваются с проблемой коллективного действия. В непрерывном взаимодействии такого рода кооперация может поддерживаться стратегией *переключения*: сотрудничай, пока так поступают все, и отказывайся, когда хотя бы один другой агент уклоняется⁵. Есть некоторые свидетель-

⁵ В отличие от «дилеммы заключенного», разыгрываемой ограниченное число раз, это утверждение относится к ситуации неопределенной длительности. Образуют ли в данном случае стратегии наказания (trigger strategy) равновесие, зависит от структуры количественных выигрышей и от коэффициента дисконтирования во времени агентов. В принципе другие стратегии также могут порождать кооперативное равновесие, но только стратегия наказания обладает необходимым свойством точки координации.

ства того, что именно на этом держатся картели. Решения не участвовать во Второй мировой войне тоже могли быть вызваны опасениями возмездия, но более вероятно, что это было связано с операционными недостатками метода. Однако такие случаи нетипичны. Например, рабочие не пересматривают ежегодно вопрос своего членства в профсоюзе. Бунты и революции являются однократными действиями, к которым не применимы доводы, касающиеся длящихся взаимодействий. Другие кампании, крестовые походы и движения обычно слишком плохо структурированы, чтобы можно было выявлять тех, кто отказывается сотрудничать. Чтобы объяснить, как в таких случаях возникает кооперация, мы должны искать какие-то другие пути.

Второе исключение возникает, если имеется внешний агент, как право, организация, которая в избирательном порядке наказывает тех, кто не сотрудничает, и/или поощряет сотрудничающих. Государство может наказывать тех, кто подвергает других действию негативных экстерналий, — например, тех, кто курит или плюет в общественных местах. Профсоюзы могут предлагать своим членам дешевые отпуска. Революционные движения могут обещать активистам ключевые должности в постреволюционном правительстве. Такие «решения» основаны на централизованном принуждении, а не на нескоординированном действии непосредственно заинтересованных индивидов. Заметим, однако, что само существование организации, способной менять стимулы для социальных агентов, может быть плодом успешного коллективного действия. Точно так же обещания наградить активистов после победы революции могут быть незаслуживающими доверия. Во многих случаях, от Генриха IV в 1594 году до Шарля де Голля в 1945 году, харизматичные лидеры разочаровывали своих сторонников, назначая или оставляя на ключевых должностях чиновников свергнутого режима. Из политических активистов не всегда получаются хорошие администраторы.

При децентрализованных однократных коллективных действиях сотрудничество не возникнет, если все агенты полностью рациональны, имеют целенаправленные эгоистические мотивации, и эти два обстоятельства общеизвестны. (Это высказывание — логическая истина, не эмпирическое утверждение.) Со ссылкой на рис. XXIV.3 мы можем выявить целый ряд мотиваций, которые могут порождать и поддерживать кооперативное поведение. В условном сценарии, который я собираюсь изложить, они срабатывают последовательно, одна за другой, так что каждая (за исключением первой) опирается на присутствие агентов, мотивированных предшествующими. Хотя этот сценарий применим не ко всем случаям, я полагаю, что одна из его черт имеет универсальное действие. Успешное коллективное действие, когда мы его изучаем, не является плодом деятельности индивидов с идентичными мотивациями. Скорее

оно — результат действия *смешанных мотиваций*, как на индивидуальном, так и на интерперсональном уровне⁶. Далее я приведу полдюжины таких мотиваций, которые, по моему мнению, в разных пропорциях присутствуют во многих случаях коллективного действия.

Кооперация нуждается в некотором количестве индивидов, которые мотивированы не только своими личными издержками и профитами. Две категории таких индивидов представляют собой то, что я называю *полными утилитаристами* и *бескорыстными утилитаристами*. Первые будут сотрудничать тогда и только тогда, когда их вклад будет увеличивать среднюю выгоду. На рис. XXIV.3 они требуют, чтобы количество сотрудничающих достигло уровня ОВ, на котором их участие начинает приносить пользу. Вторая категория будет сотрудничать тогда и только тогда, когда их вклад увеличит среднюю выгоду, *независимо от того, какие издержки понесут они сами*⁷. В некоторых приведенных ранее примерах сотрудничество имело негативные последствия только потому, что издержки участников кооперации превышали совокупную выгоду. Бескорыстный утилитарист не отступит, если его действие будет иметь такого рода отрицательный эффект. Он, однако, уклонится от сотрудничества, если оно нанесет прямой вред другим. Многие потенциальные зачинщики социальных возмущений сталкивались с таким затруднением. Их могут, как было отмечено, останавливать опасения спровоцировать репрессивные меры со стороны работодателей или властей, которые ударят по другим членам группы, а не только по ним самим.

Если они не сталкиваются с этой проблемой, бескорыстных утилитаристов может быть достаточно много, чтобы поднять количество участников кооперации до ОВ, точки, в которой присоединяются полные утилитаристы. Если первых слишком мало или описанное выше затруднение препятствует их действию, нужны *безусловные участники кооперации*, которые могли бы сформировать критическую массу на уровне ОВ, что повлекло бы за собой подключение полных утилитаристов. Эти акторы предстают в различных обликах — они могут быть кантианцами, святыми, героями, фанатиками или слегка безумными. Объединяет их то, что они действуют независимо от последствий своих действий и независимо от количества других участников кооперации. Фанатичные нацисты, например, вступали в НСДАП, когда ее перспек-

⁶ Во избежание усложнения дискуссии я буду рассматривать лишь вариации между индивидами.

⁷ Эта асимметрия зафиксирована в нескольких пословицах: «Дружба — это способность забывать то, что отдано, и помнить то, что получено» и «Никогда не забывай об оказанной тебе услуге и немедленно выбрасывай из головы услугу, оказанную тобой».

тивы были весьма туманны. Утилитаристы, наоборот, действуют в зависимости от ожидаемых последствий своих действий. Хотя их ожидания сами по себе могут зависеть от количества других членов кооперации, их действия прямо не иницируются численностью последних (см. главу V). Если кривая OGCK на рис. XXIV.3 не опускается ниже 0, оба типа утилитаристов будут действовать как пионеры-первопроходцы.

Следующая категория включает акторов, которые никогда не станут выступать в роли зачинателей. Это индивиды, чья мотивация обусловлена наблюдением за тем, как сотрудничают другие, или знанием о возможном наблюдении со стороны участников кооперации (см. главу V). Первое подмножество этой группы мотивировано квазиморальными нормами честности: нечестно оставаться в стороне, когда другие подвергают себя риску ради общего дела. Сила этой мотивации может варьироваться, то есть разных индивидов может подталкивать к сотрудничеству разное количество участников кооперации. Те, чье участие способно возникнуть при более низком пороге, могут, следовательно, подтолкнуть тех, чей порог кооперации выше. Второе подмножество мотивировано социальными нормами. Если те, кто отказался сотрудничать, могут быть идентифицированы и подвергнуты социальному ostracism, как это, к примеру, обычно происходит на рабочем месте, участники кооперации могут, пристыдив, склонить их к присоединению. Стыдливость тоже имеет свою градацию. Некоторых можно устыдить одним-единственным замечанием или тем, что их будут избегать, тогда как другие присоединятся, только если будут подвергнуты давлению со стороны многих членов группы, включая (возможно, более всего) тех, кто и сам был принужден участвовать под давлением стыда.

Последняя категория — те, кто вступает в движение из-за «выгоды, которую приносит процесс» в виде радости или какой-либо иной формы индивидуальной привлекательности. Для некоторых людей хождение с песнями по улице или использование всевозможных способов развлечения представляет самостоятельную ценность, независимо от мотивов. Для других вступление в какое-то движение становится способом «расширения сознания» или «формирования характера». Хотя желаемые состояния являются, по сути, побочными продуктами (см. главу IV), что подразумевает, что тот, кто добивается их напрямую, скорее всего, потерпит неудачу, такие люди все равно могут вступить в движение по этой причине. Поскольку эти индивиды внешне неотличимы от других категорий, они также могут поспособствовать вступлению тех, чей порог кооперации еще выше.

Эта типология мотивов не дает объяснения успешного коллективного действия, а лишь создает для него рамку. Существует множество разных мотивационных сочетаний, которые могут создавать требуемый эффект

лавинообразного роста. В некоторых случаях движение может оказаться не в состоянии завербовать достаточное количество сторонников, если первопроходцы слишком немногочисленны, чтобы привлечь утилитаристов, или же и те и другие вместе недостаточно многочисленны, чтобы привлечь тех, чье сотрудничество зависит от окружающих. В других случаях процесс может не начаться вовсе ввиду отсутствия первопроходцев. Исход зависит от распределения мотиваций в группе и от технологии коллективного действия, а также от организации и лидерства, которые я в данном случае сознательно игнорировал, сосредоточившись на ситуациях чистого децентрализованного выбора.

Библиографические примечания

Эксперимент с общественными благами и с эффектом распада описан в «Природе человеческого альтруизма». Э.Фера и У.Фишбаха-ра (Fehr E., Fischbacher U. The nature of human altruism // Nature. 2003. Vol. 425. P.785–791). Об эффекте лавинообразного роста в Восточной Европе см.: Р.Петерсен «Сопротивление и бунт: уроки Восточной Европы» (Petersen R. Resistance and Rebellion: Lessons from Eastern Europe. Cambridge University Press, 2001). Эта работа показывает, как несколько мотиваций могут сосуществовать внутри одного индивида, а не только (как в моем тексте) среди множества индивидов. Я обсуждаю некоторые вариации технологии в «Цементе общества» (The Cement of Society. Cambridge University Press, 1989). Другие вариации рассматриваются в книге Дж. Марвелла и П.Оливера «Критическая масса в коллективном действии» (Marwell G., Oliver P. The Critical Mass in Collective Action. Cambridge University Press, 1993). Изящный обзор темы представлен в «Коллективном действии» Т.Сэндлера (Sandler T. Collective Action. Ann Arbor: University of Michigan Press, 1992). Применение этой теории, учитывающее тот факт, что у индивидов могут быть как разные ресурсы, так и разные мотивации, есть в книге Дж. Боумана «Коллективное капиталистическое действие» (Bowman J. Capitalist Collective Action. Cambridge University Press, 1993). Кооперация со стратегией переключения в картелях обсуждается в работе Ч. Липсона «Дилемма банкиров: частная кооперация при реструктуризации государственных долгов» (Lipson C. Bankers' dilemmas: Private cooperation in rescheduling sovereign debts // Oye K. A. (ed.). Cooperation Under Anarchy. Princeton, NJ: Princeton University Press, 1986). Сюда можно отнести и пример с нью-йоркскими торговцами алмазами, приведенный в главе XXIII. Объяснение коллективного действия с точки зрения вертикальных отношений между лидерами и последователями приводится в книге С.Каливаса «Логика насилия в гражданской войне» (Kalyvas S. The Logic of Violence in Civil War. Cambridge University Press, 2006).

XXV. Коллективное принятие решений

Часто члены группы — от семьи до общества в целом — нуждаются в том, чтобы регулировать совместные вопросы путем принятия общеобязательных решений. Рассмотрим еще раз ограничение потребления воды в период засухи (см. главу V). Иногда эта проблема коллективного действия может быть решена децентрализованным образом, через сочетание моральных, квазиморальных и социальных норм. Однако часто властям города приходится ограничивать потребление, запрещая использование воды для каких-то целей, например для полива лужайки или наполнения бассейнов. Когда коллективное действие не срабатывает, может потребоваться коллективное принятие решений.

В качестве еще одного примера возьмем практику голосования на общенациональных выборах. Выбор между тем, чтобы пойти голосовать или остаться дома, является классической проблемой коллективного действия. Зная, что его или ее голос никак не влияет на исход голосования, каждый гражданин имеет личный интерес, требующий воздержаться от действия. Но если воздержатся все или если явка упадет до очень низкого уровня, демократия может оказаться под угрозой замены диктатурой или олигархией вопреки (почти) всеобщему интересу. Во многих демократиях явка действительно достигает приличных уровней от 50 до 80% в результате децентрализации решений граждан. Некоторые могут спросить себя: «А что, если все воздержится?» Другие могут сказать себе: «Поскольку большинство все-таки дает себе труд пойти и проголосовать, я поступлю честно, если тоже пойду». Кто-то еще может посчитать, что «хотя мой голос и не отражается на жизнеспособности демократии, он становится важен, если его умножить на большое число других граждан, на которых он влияет». В небольшой деревне некоторые могут испугаться, что «если остаться дома, соседи заметят это и вызовут осуждение».

Если эти мотивации, вместе или по отдельности, оказываются слишком слабыми, явка может упасть до катастрофически низкого уровня в ходе процесса, который отчасти сам себя усиливает («если лишь немногие идут голосовать, то почему должен это делать я?»). Чтобы обратить вспять эти процессы, парламент может ввести обязательное голосование и штрафы для тех, кто не ходит на выборы, вынеся данный

вопрос на референдум. Выражая свое мнение о том, должно ли голосование стать обязательным, избиратели сталкиваются с выбором, отличным от стоявшего перед ними при решении вопроса, идти или не идти на выборы при необязательном голосовании. Выбор делается не между «я иду голосовать» или «я остаюсь дома», а между «каждый голосует» и «каждый имеет право оставаться дома»¹. Если многие из тех, кто предпочтет второй вариант в первом случае, предпочтут первый вариант во втором, они *коллективно решают* сделать голосование на выборах обязательным.

Коллективное принятие решений связано с *политическим выбором*. Прежде чем присоединиться к процессу коллективного принятия решений, каждый член группы обладает определенными политическими воззрениями, основанными на фундаментальных предпочтениях в сочетании с комплексом каузальных убеждений о соотношении целей и средств. Основная цель коллективного принятия решений — *обобщить* индивидуальные предпочтения с помощью одного из трех механизмов, которые будут вкратце охарактеризованы далее². Такое обобщение может вызвать *трансформацию* индивидуальных политических предпочтений и создать для индивидов стимул *ввести окружающих в заблуждение* на этот счет. Взаимосвязь между суммированием предпочтений, трансформацией и введением в заблуждение может привести к значительным трудностям.

Во многих случаях, затронутых мною, небольшая группа индивидов принимает решения, которые позднее становятся обязательными для более крупной группы лиц. Иногда они уполномочены для принятия таких решений, будучи представителями или переговорщиками, выступающими от лица более крупного образования. В этой ситуации сдерживающим фактором служит знание о том, что их решения должны быть одобрены избирателями и что они рискуют не быть переизбранными в случае неудовлетворительного результата. В иных обстоятельствах большое общество не имеет других средств воздействия на принимающих решения, влияющие на его жизнь, кроме революции. Но даже здесь мы можем говорить о коллективном принятии решений внутри элиты. За смертью Сталина последовало коллективное лидерство Политбюро.

¹ Иначе говоря, вариант «я останусь дома, но всем остальным придется идти голосовать» не будет присутствовать среди опций на референдуме.

² Хотя выражение «механизмы обобщения», как правило, закреплено за процедурами голосования, я пользуюсь им здесь, чтобы обозначить любой процесс, в котором акторы, имеющие различные исходные предпочтения, взаимодействуют так, чтобы прийти к решению, которое они сочтут обязательным для всех.

Чилийская хунта, правившая с 1973 по 1980 год, имела хорошо структурированный внутренний порядок принятия решений.

Три механизма, которые я собираюсь рассмотреть, — это *спор* (arguing), *торг* (bargaining) и *голосование* (voting). Я полагаю, что этот список исчерпывающий. Хотя некоторые коллективные решения принимаются посредством механизмов, вносящих элемент случайности, как, например, когда собрание подбрасывает монетку, чтобы поставить точку, или когда вожди выбираются среди всего населения по жребию, само решение использовать подобную процедуру должно быть выработано путем споров, торговли или голосования. Я не буду обсуждать очень интересный вопрос о том, не является ли одна из этих процедур сама по себе более фундаментальной в сравнении с остальными, — например, не принимается ли решение о вынесении какого-либо вопроса на голосование в ходе спора. Вместо этого я собираюсь исследовать особенности каждой процедуры, включая важные отклонения от них.

Позвольте привести сначала несколько примеров этих трех процедур. Спор в чистом виде наблюдается (или, по крайней мере, считается за правило) в жюри, от которого требуется единогласие. Даже здесь некоторые члены могут прибегнуть к молчаливому торгу, поскольку им легче держаться до конца, так как они меньше других торопятся завершить участие в жюри и вернуться к обычной жизни³. Поскольку при принятии решения время всегда имеет большое значение и поскольку участники процесса часто дисконтируют будущее с разными коэффициентами — этот пример может в действительности оказаться типичным.

Чистый торг представлен последовательными играми «раздели доллар», в которых стороны одна за другой делают предложения и контрпредложения. Исход определяется механизмом торга и способностью сторон к его ведению, то есть ресурсами, дающими им возможность выступать с правдоподобными угрозами и обещаниями. Процесс проиллюстрирован на рис. XX.2.

Концепция голосования как коллективного принятия решений в чистом виде была представлена у Руссо. Гражданам следовало выработать свои предпочтения в изоляции друг от друга так, чтобы не поддаваться влиянию красноречия или демагогии. Так как голосовать они должны также в изоляции друг от друга, торговля голосами исключается. В настоящих политических системах этот идеал недостижим. Он может быть проиллюстрирован некоторыми решениями с низкими ставками, таки-

³ В ранних английских судах практика морить присяжных голодом до тех пор, пока они не придут к единодушному решению, тоже давала больше возможностей для торга одним присяжным, чем другим.

ми как выборы членов научной академии, чья основная функция — избирать новых членов.

Сочетание спора и голосования без торга может быть проиллюстрировано наймом и решением о предоставлении постоянной должности (*tenure*) на университетских факультетах. Предполагается, что здесь действует лишь оценка заслуг кандидата с последующим голосованием. Хотя этот идеал не всегда соответствует реальности, иногда такое все же происходит. На определенных факультетах существует запрет на оказание взаимных услуг, поддержанный нормой, требующей объяснять причины голосования тем или иным образом.

Сочетание спора и торга без голосования иллюстрируется коллективным торгом относительно заработной платы. Когда профсоюз и менеджмент компании решают, как поделить ее доход, может показаться, что речь идет только о торге. Однако более пристальный взгляд выявляет существенное число споров по конкретным вопросам, таким как финансовое благополучие компании и производительность рабочей силы.

Сочетание торга и голосования имеет место, когда члены профсоюза должны утвердить соглашение о заработной плате, достигнутое их представителями. В таких ситуациях ожидаемый исход голосования выступает в качестве угрозы при торге.

Принятие политических решений на уровне комитета, парламента или всего населения часто включает все три процедуры⁴. И опять это связано с потребностью скорейшего принятия решения. Обычно голосование имеет место, когда вопрос должен быть разрешен срочно, поэтому у участников нет времени совещаться, чтобы прийти к единогласному мнению. С более прозаической точки зрения агенты могут быть не заинтересованы в единодушии. Если решение является более срочным для одних, чем для других, также возникает возможность для торга, поскольку те, кто имеет возможность подождать, могут потребовать уступок в обмен на более быстрое принятие решения. В постоянных комитетах и собраниях торг может возникнуть благодаря практике взаимной поддержки (*logrolling*), что связано с неодинаковой интенсивностью предпочтений по вопросам, которые можно обменивать один на другой. Другие механизмы торга в законодательных органах включают obstructionism (*filibustering*) и «политику пустого стула», которая позволяет группе пользоваться правилами кворума для достижения того, что не может быть достигнуто другими средствами.

В таких ситуациях источники способности к торгу создаются внутри самого собрания. В иных случаях те, кто принимает решения, могут

⁴ Даже всеобщие выборы дают большое пространство торговли. Если голосование публичное, избиратели и кандидат могут торговаться о цене голосов.

пользоваться ресурсами, существующими независимо от собрания — деньгами и людьми. В 1789 году дебаты во французском Учредительном собрании были парализованы королевскими войсками и парижскими толпами, первоначально служившими орудием, направляемым депутатами против войск, а позднее ставшими угрозой, используемой одними депутатами в борьбе с другими. В 1989 году квази- или предконституционные круглые столы в Польше были прекращены в связи с угрозой советского вторжения и перспективой экономического коллапса. Если голос не может быть получен с помощью обещания другого голоса, как при взаимной поддержке, он может быть куплен за деньги, например путем выделения целевых партийных фондов на кампанию по переизбранию.

Как станет понятно из нижеследующих рассуждений, три метода коллективного принятия решения могут рассматриваться как три стадии в идеализированной последовательности, в том смысле, что каждая из них естественным путем возникает из предшествующей. Хотя спор по сути своей направлен на единодушие, то есть он основан на доводах, которые, как предполагается, действуют для всех, эту цель редко удается достичь. Для урегулирования вопроса требуется голосование. Поскольку часто на голосование выносятся множество вопросов, это, естественно, создает почву для торга в форме оказания взаимных услуг.

СПОР

Спор — это усилия по убеждению путем приведения доводов. Еще со времен панегирика Афинам Перикла этот способ принятия решений был тесно связан с демократической политикой:

Одним и тем же лицам можно у нас и заботиться о своих домашних делах, и заниматься делами государственными, да и прочим гражданам, отдавшимся другим делам, не чуждо понимание дел государственных. Только мы одни считаем не свободным от занятий и трудов, но бесполезным того, кто вовсе не участвует в государственной деятельности. Мы сами обсуждаем наши действия или стараемся правильно ценить их, не считая речей чем-то вредным для дела; больше вреда, по нашему мнению, происходит от того, если приступать к исполнению необходимого дела без предварительного обсуждения его в речи.

Связь между институтом публичных дебатов и «мудрыми действиями» может быть не совсем прямой. Часто главный эффект публичности — исключение открытых апелляций к корысти. В публичных дебатах оратор, заявляющий, что «мы должны сделать это, потому что это хорошо для меня», никого не убедит и, более того, будет подвергнут неформальным санкциям и ostracism, которые сделают его менее эффективным

в будущем. Обстоятельства публичности вынуждают даже тех, кто руководствуется только собственным интересом, представлять предлагаемые меры в качестве мотивированных более объективными ценностями. Такое *искажение предпочтений* отличается от процесса *трансмутации* (см. главу IV) точно так же, как обман отличается от самообмана. Только на этот раз выдавать свои мотивы за разумный довод заставляет сам интерес говорящего, а не его потребность в самоуважении. Личный интерес делает искажение трудноуловимым при отстаивании в категориях справедливости политики, которая несколько отклоняется (но не слишком сильно) от той, что идеально с ним совпадает. Искажение может, в действительности, выйти боком, если оно слишком очевидно. Следовательно, *потребность скрыть свои фундаментальные интересы может вызвать перемену в политических предпочтениях* через то, что мы называем «цивилизующей силой лицемерия».

Во многих обществах в качестве избирательного ценза использовалась собственность. Безусловно, в защиту данного принципа можно выдвинуть некие доводы, доказывающие его справедливость. На Конституционном конвенте Мэдисон утверждал, что строгие ограничения, касавшиеся собственности сенаторов, не столько защищали привилегированный класс от народа, сколько народ от самого себя. Но как уже отмечалось, в этот довод есть нечто подозрительное, слишком уж он совпадает с интересами богачей. Тогда, возможно, полезно будет обратиться к грамотности как объективному критерию, который *находится в значительной, но не полной корреляции с собственностью*. На разных стадиях американской истории грамотность также служила легитимирующим заместителем для целей, в которых боялись признаться открыто, как, например, для желания не пускать в политику негров или католиков. Американское иммиграционное законодательство также всегда использовало грамотность в качестве субститута критерия, о котором нельзя было заявить публично. Предложения поставить заслон иммигрантам посредством тестирования их на знание родного языка обычно оправдывали как способ отбора на основании личных достоинств, повсеместно принятой, но предвзятой процедуры. Однако в действительности радетели грамотности руководствовались предрассудками или групповым интересом. Аристократы-нативисты⁵ стремились отсечь, как правило, неграмотных иммигрантов из Центральной и Юго-Восточной Европы. Рабочий класс боялся, что приток неквалифицированной рабочей силы может вызвать снижение уровня оплаты труда.

⁵ Нативизм — теория и политическая доктрина, провозглашающая верховенство интересов коренного населения над интересами иммигрантов. — *Примеч. ред.*

Естественно, было бы ошибкой считать, что спор всегда может быть редуцирован к более или менее тонким способам отстаивания собственных интересов. Если бы это было так, то искажение не потребовалось бы, так как некого стало бы обманывать⁶. Если собеседники мотивированы искренним желанием защищать общественное благо, обсуждения и дебаты могут изменить их убеждения, что приведет к смене политических предпочтений. Вероятность этого особенно высока, если различные члены группы имеют доступ к разной информации, тем самым они могут улучшить качество своих решений, объединив информацию, которой располагает каждый⁷. Если орган представительский, важно отбирать делегатов из самой разной среды. При выборе представителей в общенациональную ассамблею это соображение говорит в пользу пропорционального голосования с низким порогом или без порога вообще⁸. Можно также потребовать, чтобы представители избирательного округа являлись его жителями.

Люди также участвуют в спорах, хотя, вероятно, реже, о фундаментальных предпочтениях и меняют их в результате дискуссий. Зачастую это изменение происходит посредством выявления скрытых аналогий между отдельными событиями или демонстрации внешних сходств. Многие люди, к примеру, выступают против обаятельного использования «органов умерших доноров» для трансплантации. Они полагают, что, если у семьи покойного имеются против этого возражения религиозного свойства, их следует уважать. В ответ на это можно указать на обязательное проведение вскрытия в случае вызывающей подозрения смерти, даже если эта процедура противоречит религиозным чувствам семьи. Если эти навязываемые меры принимаются для определения причины смерти, то есть основания утверждать, что они должны быть приемлемы и в целях спасения жизни. Изменения могут также произойти, когда обнаруживается противоречие общего принципа догадкам, касающимся частных эпизодов. Человек может согласиться с обязательным использованием органов умершего на утилитаристских основаниях, но воспротивится оправданности убийства одного произ-

⁶ Аналогичный довод говорит о том, что, если бы все сигналы животных были бы обманными, обман бы просто потерял смысл.

⁷ Однако не будем забывать, что это улучшение более вероятно, если они объединят свои необработанные данные, а не просто выводы, сделанные на основе «сырых данных» (см. главу XXIII).

⁸ Другие соображения, а именно потребность в эффективном управлении, могут склонять в пользу мажоритарной системы или пропорционального голосования с высоким порогом. При выборах в конституционные собрания, в которых управление играет подчиненную роль, существует тенденция выбирать делегатов на основе пропорционального голосования.

вольно выбранного человека с целью использования его сердца, почек, легких и печени для спасения пяти чужих жизней⁹. В результате первоначальный неопределенный утилитаризм может быть пересмотрен с учетом неконсеквенциалистских ценностей (см. главу IV).

Преимущества спора могут быть, однако, подорваны эффектом *выступления перед аудиторией*. Индивиды, пекущиеся об общественном благе, не меньше других могут быть подвержены влиянию самолюбия, которое помешает им публично признаться в том, что они передумали. В главе III я отмечал, что, как позднее также отмечал Мэдисон, это была главная причина, по которой Конституционный конвент проводился за закрытыми дверями и при условии неразглашения со стороны делегатов. Может показаться, что его довод противоречит традиционному аргументу в пользу дебатов, открытых для публики. Многие законодательные решения имеют сильное краткосрочное воздействие на интересы законодателей. Если процесс принятия решений скрыт от глаз общества, споры об общем интересе могут легко скатиться в торг по поводу неприкрытых интересов. Общественный доступ к наблюдению за процедурами и голосованием ограничивает тайные договоренности и в качестве побочного эффекта способствует поддержанию общественного блага. Как писал Бентам, «чем больше число искушений, которым подвергается отправление политической власти, тем больше необходимость дать тем, кто ее отправляет, самые сильные основания, чтобы им сопротивляться. Но нет более постоянного и более универсального основания, чем высшая зависимость от общественности». Или, как сказал американский судья Луис Брэндейс, «солнечный свет — лучшее дезинфицирующее средство».

Эти замечания вскрывают возникающие в ходе спора проблемы. Если дебаты ведутся публично, от этого страдает качество. Если они проходят за закрытыми дверями, спор может превратиться в торг. Однако это напряжение можно снять, если обсуждаемые вопросы будут оставлять мало места для игры частного интереса. Конституционные собрания в меньшей степени склонны к ангажированным решениям, чем обычная законодательная деятельность, не потому, что у делегатов менее эгоистические мотивации, но потому, что (или в той мере, в какой) их интере-

⁹ Утилитаристы склонны отрицать допустимость подобных выводов. Как правило, они заявляют, что негативные последствия страха и неопределенности, вызванные представлениями о том, что вас могут выбрать в качестве «случайного донора», скорее сведут на нет преимущества этой практики. *Но откуда они знают?* Я думаю, что в своих рассуждениях они движутся назад, от очевидной неприемлемости этой практики к наличию издержек, которые исключили бы ее на утилитаристских основаниях, а не вперед — от демонстрации издержек к отказу от процедуры.

сы меньше воздействуют на текущие вопросы. Кроме того, как я неоднократно отмечал, из-за долгосрочной перспективы конституционных замыслов частный интерес может до некоторой степени имитировать желание способствовать общественному благу.

ГОЛОСОВАНИЕ

Голосование может потребоваться, когда в ходе спора не удалось достичь консенсуса по тому или иному вопросу. Существуют самые разные способы голосования. При народном волеизъявлении вариации касаются права голоса, порядка голосования (тайное или открытое), большинства, необходимого для принятия решения, а также в некоторых системах с референдумом — процента явки. Основными параметрами голосования на общем собрании являются кворум, численные критерии большинства и выбор между поименным голосованием и поднятием рук (и сходными процедурами вроде «выкрикивания» или «сидения и вставания»). Тайное голосование на общих собраниях — редкость, но не такая уж неслыханная. На Учредительном собрании во Франции 1789–1791 годов председатель избирался тайным голосованием. До недавних пор в Италии и Колумбии практиковалось тайное голосование в парламенте. Заметим, что тайное голосование следует отличать от закрытой процедуры, когда в зал заседаний не пускают посторонних. Последняя может сочетаться с открытым голосованием, позволяющим участникам давать надежные гарантии взаимного обмена голосами, которые были бы невозможны при тайном режиме. В свою очередь, если бы заседание было открыто для посетителей, некоторые из них могли бы негативно отнестись к тому, что их представитель голосует вразрез с их позицией по тому или иному вопросу, поскольку они лишены возможности оценить полученные таким способом преимущества по другому вопросу.

В дальнейшем я ограничусь мажоритарным голосованием. Хотя эта практика и не является универсальной, решение утверждать предложения подавляющим большинством (например, в 3/5 или 2/3), само по себе, как кажется, должно быть принято простым большинством. Учредительные собрания, которые зачастую устанавливают квалифицированное большинство для принятия будущих конституционных поправок, практически неизменно используют в своей практике голосование путем простого большинства¹⁰. Идеализированная модель, в которой конституционная ассамблея тайно выносит единогласное ре-

¹⁰ Частичным исключением была южноафриканская конституция 1996 года. Требование о том, чтобы однопалатная система была принята квалифицированным большинством, было изложено в промежуточной конституции

ТАБЛИЦА XX.1

| Состав | Фундаментальные предпочтения | Убеждения | Политические предпочтения |
|-------------|------------------------------|--|---------------------------|
| Реакционеры | Дестабилизировать режим | Двухпалатная система стабилизирует режим | Однопалатная система |
| Умеренные | Стабилизировать режим | Двухпалатная система стабилизирует режим | Двухпалатная система |
| Радикалы | Стабилизировать режим | Двухпалатная система дестабилизирует режим | Однопалатная система |

шение, но, как только завеса секретности снимается, начинает принимать решения мажоритарным голосованием, имеет отдаленное отношение к реальности создания конституций. Я не буду касаться вопроса, сделав лишь одно маленькое замечание, что воздержание или «политика пустого стула» может использоваться меньшинством, чтобы заблокировать решение, которое в случае его участия и голосования «против» было бы принято.

Недостаток консенсуса, который может потребовать решения путем голосования, может возникнуть из-за расхождения фундаментальных предпочтений, расхождения убеждений или из-за того и другого. В качестве примера можно привести дискуссии об однопалатной и двухпалатной системе во французском Учредительном собрании в 1789 году. Говоря широко, Учредительное собрание составляли три примерно равные численно группы. Реакционное правое крыло хотело перевести стрелки часов обратно, в период абсолютной монархии, умеренный центр стремился к конституционной монархии с жестким контролем со стороны парламента, а левое крыло — конституционную монархию со слабым контролем со стороны парламента. По вопросу двухпалатной системы распределение, в очень упрощенном виде, было таким, как показано в табл. XXV.1.

В конце концов, двухпалатная система была отвергнута союзом реакционеров и радикалов. Этот общий феномен — политическое соглашение, основанное на взаимоисключающих различиях в предпочтениях и убеждениях, — имеет широкое распространение. На этой основе

1993 года, которая сама по себе была принята не столько посредством голосования, сколько путем торга.

можно даже достичь единогласия¹¹, хотя очевидно, что это единогласие иного рода, нежели в «ситуации идеальной речи», в которой говорящие мотивированы только общим благом и готовы прислушиваться к доводам.

В предложенном мной условном изложении этих дебатов одно большинство полагало, что двухпалатная система стабилизирует режим, а другое большинство хотело его стабилизировать (см. табл. XXV.1.). Если бы во имя коллективного решения сначала путем (искреннего) мажоритарного голосования обобщили бы убеждения, потом тем же способом — предпочтения и затем, наконец, приняли бы меры, которые, согласно обобщенным убеждениям, лучше всего отражали бы предпочтения, *двухпалатная система* была бы создана. (Во избежание такого исхода реакционеры могли бы высказать ложное убеждение, что двухпалатная система дестабилизирует режим, тем самым создав для него большинство, а следовательно, и большинство для выбора однопалатной системы.) Насколько мне известно, процедура «двойного обобщения» никогда не используется на практике. Учитывая сложности прямого обобщения политических предпочтений, любая система такого рода, вероятно, была бы неуправляемой.

В дебатах по поводу двухпалатного устройства голосующие различаются как по убеждениям, так и по конечным целям. В других случаях они могут совпадать по одному из двух параметров и различаться по-другому. Поскольку напрямую обобщаются и наблюдаются именно политические предпочтения, может оказаться сложно различить оба фактора, участвующих в их формировании. Тем не менее мы можем умозрительно определить воздействие мажоритарного голосования на обобщение убеждений (принятие идентичных целей) и на обобщение фундаментальных предпочтений (принятие идентичных убеждений). Согласно Токвилю, демократия (то есть мажоритарное голосование с широкими избирательными правами) превосходит остальные системы в двух отношениях: «Моральное превосходство большинства отчасти основывается на идее, что в собрании людей больше просвещения и мудрости, чем в уме у одного человека, или что количество законодателей значит боль-

¹¹ Во французском Учредительном собрании это произошло в мае 1791 года, когда радикалы, умеренные и реакционеры объединили силы, чтобы проголосовать за закон, который делал членов ассамблеи невыеборными на время первого созыва законодательного органа. Радикалы ставили перед собой цель ослабить законодательный орган в интересах якобинского клуба, реакционеры — ослабить его в интересах короля. Решение было принято единогласно, поскольку умеренный центр, «упиваясь бескорыстием» (см. главу V), с энтузиазмом голосовал за то, чтобы лишить себя места в будущем законодательном собрании.

ше, чем способ их отбора, [и] на принципе, что интересы многих следует предпочесть интересам немногих». Демократические чиновники могут «совершать серьезные ошибки», но «никогда не будут систематически принимать линию, враждебную большинству»¹².

Рассмотрим для начала обобщение убеждений. Существует обширная дискуссия о том, какое право голоса, широкое или ограниченное, является более удачной процедурой для получения правильных суждений: правда ли, что многие мудрее, чем немногие. По мнению Аристотеля, это был вопрос сопоставления количества (число участников политического процесса) и качества (компетентность участников):

Под качеством я разумею свободу, богатство, образованность, благородство происхождения; под количеством — численное превосходство массы населения. Может случиться, что одна из частей, составляющих государство, будет обладать качественным преимуществом, а другая — количественным; так, например, люди безродные будут превосходить своей численностью людей благородного происхождения, либо неимущие будут превосходить богатых, однако это количественное превосходство не должно быть таким же большим, как качественное превосходство благородных и богатых. Приходится поэтому оба этих превосходства уравнивать. Где количество неимущих превосходит указанное соотношение, там, естественно, рождается демократия.

Говоря современным языком, вопрос может быть сформулирован в терминах «теоремы о жюри присяжных» Кондорсе. Предположим, что члены суда присяжных высказывают свои (независимые) суждения относительно того, действительно ли обвиняемый совершил то, в чем его обвиняет прокурор, и что вероятность правоты каждого из них составляет более 50%. Кондорсе показал, что, если жюри выносит вердикт путем мажоритарного голосования, его шансы принять правильное решение увеличиваются с ростом числа присяжных¹³ и тяготеют к определенности, когда состав коллегии растет до бесконечности. Кроме того, для данной численности присяжных шансы верного решения большинства увеличиваются с возрастанием вероятности принятия правильно-

¹² Тем не менее он не спрашивает, не может ли случайная склонность к *серьезным* ошибкам быть более серьезной, чем *систематическая* предвзятость недемократических режимов. Хотя Токвиль утверждал, что благодаря благоприятному географическому положению Соединенные Штаты могут себе позволить совершать ошибки, в отношении других стран это может быть не так.

¹³ При допущении, в противовес утверждениям следующего абзаца, что увеличение числа голосующих никак не влияет на вероятность правоты каждого из них.

го решения каждым членом¹⁴. Следовательно, как указывал Аристотель, исход можно улучшить либо путем увеличения численности состава суда, либо повышением их квалификации¹⁵.

Пойдя дальше Аристотеля, мы заметим, что квалификация может находиться в прямой зависимости от количества, а не от социально-экономического положения избирателей. Говоря языком социальных наук, компетентность избирателей может быть «эндогенной» системе, а не даваться «экзогенно». Предположим, что нужно выбрать между двумя политическими системами, олигархией и демократией, и что оба решения принимаются путем мажоритарного голосования, но с разным объемом избирательных прав. При демократии избиратели могут рационально решить остаться в неведении, поскольку влияние каждого на исход крайне невелико¹⁶. При олигархии избиратели вложат больше усилий в сбор информации, поскольку влияние каждого из них на результат больше.

Бентам замечал, что этот аргумент также применим к голосованию на собрании: «Чем больше число голосов, тем меньше вес и ценность каждого голоса, тем ниже его цена в глазах избирателя и тем меньше у него стимулов сообразовывать свой голос с истиной и вообще идти его отдавать». Отвечая на довод, что собрание (он имел в виду французское Учредительное собрание 1789 года) должно быть многочисленным, поскольку «вероятность мудрости увеличивается с возрастанием численности участников», он писал, что «подавление, которое та же причина оказывает на силу мотивации проявить свою просвещенность, компенсирует это преимущество». В этом промежутке между качеством и количеством и будет нахо-

¹⁴ Шансы большинства принять правильное решение также возрастают при выполнении требования о квалифицированном большинстве, например 60%. В этом случае, однако, можно получить «подвешенное жюри», в котором ни вина обвиняемого, ни его невиновность не набирают необходимого процента голосов.

¹⁵ Можно также попытаться обеспечить соблюдение условий теоремы Кондорсе, увеличив вероятность, что убеждения голосующих действительно независимы друг от друга. С этой точки зрения предложение Руссо запретить обсуждение до принятия решения может иметь смысл. В то же время, если принятие решения улучшает качество убеждений, нельзя возражать против того, что оно делает их более зависимыми друг от друга. Условия теоремы являются достаточными, но не необходимыми для того, чтобы мажоритарное голосование привело к хорошему результату.

¹⁶ Поскольку решение идти голосовать само по себе иррационально (см. главу XII), могут спросить, не случится ли так, что граждане также иррационально вложат свои усилия в сбор информации об обсуждаемых вопросах. Однако в данном контексте важнее вопрос о том, будут ли граждане больше вкладываться при ограниченном праве голоса, чем при более широком избирательном праве, подобно тому как, если выборы представляются закрытыми, может прийти больше избирателей.

даться оптимальный размер электората, который максимизирует вероятность того, что мажоритарное голосование даст правильные убеждения¹⁷.

Далее рассмотрим сбор предпочтений путем голосования. В данном случае встают две основные, тесно связанные друг с другом проблемы — *искажение предпочтений* и *неопределенность* исхода. Если начать с первого, люди могут иметь стимул голосовать за другие предложения и кандидатов, чем те, которых они на самом деле больше всего хотели бы принять и выбрать. Это явление может быть вызвано выбором путем открытого голосования вместо тайного. В Афинах классического периода решения на собраниях принимались поднятием рук, в связи с чем многие опасались выражать свои истинные убеждения. Вследствие этого, как утверждал Фукидид, «при энтузиазме большинства [по поводу сицилийского похода] те немногие, кто был против, боялись показаться недостаточными патриотами, подняв руки против этого». (В больших афинских судах решения, наоборот, принимались тайным голосованием.) Использование поименного голосования вместо других процедур, например «вставания против сидения», также может внушать голосующим робость. На Учредительных собраниях в Париже (1789–1791) и во Франкфурте (1848) у радикалов была практика требовать поименного голосования по важным вопросам со скрытой, а иногда и открытой угрозой, что они подвергнут голосующих против радикальных предложений риску народного насилия, огласив списки с их именами. Даже если при системе «вставание против сидения» существовало очевидное большинство, поскольку система усложняла выявление того, как голосовали отдельные индивиды, при поименном голосовании исход мог быть пересмотрен¹⁸.

¹⁷ Теоретически оптимум может лежать в одной из двух крайних точек — либо отдельный индивид, либо все взрослое население. При разумных допущениях более вероятно, что здесь будет «внутренний максимум». Если оптимальный размер мал, можно выбирать голосующих наугад среди населения в целом, чтобы они не выражали узкогрупповые интересы. В такой перспективе голосование будет скорее *функцией*, чем *правом*.

¹⁸ Недавний пример: когда в 1964 году обсуждался Акт о гражданских правах, председатель комитета по процедурным вопросам (Говард Смит) ввел, как ему казалось, убийственную поправку: он предложил, чтобы билль запрещал дискриминацию по половому признаку, наряду с расовым и национальным. Лидеры парламента, разделяя веру Смита в то, что такая поправка убьет документ, попытались заставить демократов голосовать против этого предложения, и, по сообщениям ряда источников, многие из членов партии лично готовы были пойти на это только при тайном порядке голосования. Но галереи были заполнены представительницами групп по борьбе за права женщин, которые пришли понаблюдать за подсчетом голосов, и в таких обстоятельствах либеральные члены парламента оказались не готовы согласиться с большинством и провалили поправку Смита. Оказалось, что последняя привела не к поражению билля, а к включению пола в ряд законодательно защищенных признаков.

В ходе тайного голосования также может иметь место искажение. Практически во всех системах волеизъявления могут возникать ситуации, в которых голосующий, выбирая альтернативу, отличную от приоритетной для него, может породить исход, более предпочтительный тому, который выпал бы в случае искреннего голосования¹⁹. (Исключение может возникнуть, если кандидаты или предложения были выбраны при помощи генератора случайных решений с вероятностью выбора альтернативы, пропорциональной доле одобрявших ее. В этом случае проблема «потраченного зря» голоса не возникла бы. Недостатки системы очевидны и объясняют, почему она никогда не использовалась.) Желание, чтобы приоритетный вариант выиграл с не очень большим перевесом, может заставить голосовать против него. Например, в главе XVIII я говорил о том, что социалисты могут голосовать за коммунистов, чтобы сдвинуть платформу своей партии влево. Если известно наверняка, что вариант, стоящий для голосующего на первом месте, не проходит, он может проголосовать за лучшую альтернативу из тех, что имеют шансы победить. Некоторые системы голосования также создают стимул для того, чтобы не проявлять благосклонность кандидату или предложению, предпочитаемому другими, а проявить свои собственные предпочтения (см. последующий пример), или вводят новые альтернативы с единственной целью сделать более вероятным выбор, соответствующий реальным предпочтениям.

В некоторых случаях исход мажоритарного голосования является неопределенным. Предположим, в муниципальном собрании есть три блока примерного одинакового размера, представляющие соответственно бизнес, промышленных рабочих и профессиональных госслужащих. Собрание должно выбрать между строительством крытого бассейна, субсидированием местного симфонического оркестра или строительством поля для гольфа. В соответствии со стереотипными представлениями об этих группах предположим, что (после долгих дебатов) они расставляют варианты, как показано в табл. XXV.2.

Если альтернативы в голосованиях выставляются одна против другой попарно, есть большинство бизнесменов и рабочих, предпочитающих поле для гольфа оркестру, большинство бизнесменов и госслужащих, предпочитающих оркестр бассейну, и большинство госслужащих и рабочих, предпочитающих бассейн полю для гольфа. Таким образом, «социальные предпочтения» являются *нетранзитивными* или *круговыми* (cycling).

¹⁹ Говоря более технически, в не диктаторских и не пробабилистских системах голосования выражение своих истинных предпочтений, на самом деле, не является доминирующей стратегией. Однако можно придумать системы, в которых искренность будет равновесием, то есть наилучшим ответом на искренность других. Детали такой системы сложны и не отражают особенности реальных систем голосования.

ТАБЛИЦА XXV.2

| Группы | Бизнесмены | Рабочие | Госслужащие |
|-----------------|------------|---------|-------------|
| Поле для гольфа | 1 | 2 | 3 |
| Оркестр | 2 | 3 | 1 |
| Бассейн | 3 | 1 | 2 |

В случае индивидуального выбора транзитивность была требованием рациональности (см. главу XI). В настоящем контексте дело не столько в рациональности, сколько в определенности. Если муниципальный совет воспользуется только ранжированием, показанным в табл. XXV.2, непонятно, как вообще он сможет принять решение. Поскольку голосование проводится потому, что совет не может прийти к консенсусу, дальнейшие дебаты едва ли помогут. Если бы можно было измерить *интенсивность*, с которой различные группы предпочитают один вариант другому, или степень, в которой варианты удовлетворяют объективные нужды, можно было бы сказать, что один вариант однозначно превосходит другой. Однако нет общей процедуры, которая позволила бы сравнить степени интенсивности предпочтений или удовлетворения потребностей индивидов²⁰. К примеру, бесполезно спрашивать их о том, как они оценивают тот или иной вариант по шкале от 0 до 10. С одной стороны, мы не знаем, означает ли данная оценка (например, вариант 7) одно и то же для членов этих трех групп. С другой стороны, просьба оценить варианты может побудить их исказить свои предпочтения, например, присудив варианту 10, занимающему верхнюю строчку, и 0 — всем остальным.

Неясно, насколько эта проблема «круговых социальных предпочтений» важна на практике. Она не может возникнуть, если индивидуальные предпочтения «имеют один пик», то есть варианты могут быть ранжированы от «самого высокого» до «самого низкого» таким образом, что предпочтения каждого индивида ровно возрастают вплоть до наиболее предпочтительного исхода, а затем ровно снижаются по мере отдаления

²⁰ Вот почему классический утилитаризм, который велит нам максимизировать общую сумму индивидуальных полезностей, трудно применять на практике. Эта проблема, однако, не должна внушать мысль, что сравнивать степени благосостояния индивидов всегда бессмысленно. Как институты, так и индивиды каждый день с большой уверенностью проделявают такие сравнения. И все же поскольку эти сравнения основываются на догадках, а не на точных измерениях, во многих случаях они не смогут дать однозначного результата.

от него. Во многих случаях это разумное свойство предпочтений. Если индивид предпочитает ставку налога 20%, он выберет 19%, а не 18, и 21, а не 22%. Более того, не было примеров того, чтобы люди на собрании просто подняли руки и заявили, что, раз не отражена «воля народа», никакого решения принято не будет. На самом деле, если статус-кво входит в число вариантов, эта идея представляется противоречивой. Какое-то решение всегда принимается, либо по умолчанию (сохранение статуса-кво), путем использования традиционной процедуры голосования, либо путем манипулирования повесткой дня.

И все же тот факт, что решение принято, не обязательно предполагает, что оно воплощает в себе народную или «общую» волю в каком-то определенном смысле. Для комбинации (искренних) предпочтений, например приведенных в табл. XXV.2, сама идея общей воли бессмысленна. Как часто возникают такие комбинации? Политологи привели множество примеров. Остальные утверждали, что предложенные примеры неправильно описаны и что тщательный анализ опровергает эти утверждения о круговом большинстве. Я опишу два случая, которые представляются действительными примерами круговых предпочтений.

8 октября 1992 года норвежский парламент решил, что будущий аэропорт Осло должен располагаться в Гардермоен (я обозначу этот вариант как опцию G). Другим вариантом был Хобел (альтернатива H) и решение, которое предполагало сочетание Гардермоена и существующего аэропорта Форнебу (альтернатива D). Эти варианты не сравнивались друг с другом все сразу, а должны были рассматриваться последовательно в сравнении со статус-кво. Как только вариант получал большинство голосов, он принимался²¹. Хотя это серийное голосование традиционно для парламента, возможны и другие системы — например, выставление одного варианта против другого в голосовании попарно до тех пор, пока не останется один победитель. Как мы вскоре увидим, при последовательном голосовании порядок, в котором варианты ставятся на голосование, может оказаться решающим.

Выраженные предпочтения сторон, которые, за небольшим исключением, совпали с голосованием депутатов, были следующие:

Трудовая партия (63 депутатов): $G > D > H$
 Коалиция Лево́й социалистической партии, Христианских демократов и Аграрной партии (42 депутата): $D > H > G$
 Консервативная партия (37 депутатов): $H > G > D$
 Партия прогресса (22 депутата): $H > D > G$
 Один независимый депутат: $G > H > D$

²¹ Поскольку все варианты были предпочтительнее статус-кво, можно было быть уверенным, что один будет принят.

Если допустить, что это были *искренние* предпочтения, социальные предпочтения были круговыми: D выигрывает у H 105 к 60, H выигрывает у G 101 к 64 и G выигрывает у D 101 к 64. Перед этим парламент одобрил порядок, в котором должны рассматриваться альтернативы. Трудовая партия предлагала G-D-H, тогда как глава парламента настаивал на D-H-G. Когда предложения вотировались одно против другого, победили лейбористы. Если бы победило предложение председателя, лейбористы, вероятно, голосовали бы за D, поскольку иначе их неспособность набрать большинство для D привела бы к принятию предложения H, стоявшего у них на последнем месте. При одобренном порядке голосования в подобном затруднении оказалась консервативная партия. В конце концов, консерваторы голосовали за G, так как, если бы они голосовали против, выиграло бы предложение D, которое они ставили на последнее место. Хотя теоретически возможно, что лейбористы не были искренни, когда выбирали D как второй по приемлемости вариант, и что они сделали это только для того, чтобы заставить поверить консервативную партию в то, что голосование против G приведет к принятию D, доказательств этого нет. Если бы так было на самом деле, социальные предпочтения не были бы круговыми, поскольку H побивает и D, и G.

Второй пример практически исключает, что круговые предпочтения могут быть чистым артефактом искажения. Он возник в контексте споров о порядке демобилизации из американской армии после Второй мировой войны. Скорая демобилизация была ограниченным благом, которое должно было распределяться по справедливости. Чтобы выработать критерий, армия провела широкомасштабные исследования среди личного состава. В этом исследовании, сопоставлявшем критерии попарно, ранжирование показало некоторую противоречивость. Так, 55% считало, что женатые лица с двумя детьми, ни разу не побывавшие в сражении, должны демобилизоваться раньше, чем одинокие мужчины, прошедшие две кампании; 52% ставили 18 месяцев, проведенные за границей, выше двоих детей; а 60% ставили две кампании выше, чем 18 месяцев, проведенных за границей. Крайне маловероятно, чтобы респонденты искажали свои предпочтения²².

²² Авторы исследования, из которого я брал эти данные, отмечали, что «еда ли можно ожидать высокую степень внутренней последовательности при таком запутанном гипотетическом выборе», указывая на то, что это проблема индивидуально непоследовательного ранжирования. Если бы суммированное большинство составило более 200%, такая догадка была бы оправданна. А поскольку они в сумме дают всего лишь 167%, вполне возможно, что ранжирование было индивидуально последовательным и при этом вызвало коллективную интранзитивность. Исследование было опубликовано в 1949 году, за два года до новаторского исследования Кеннета Эрроу, посвященного общению предпочтений и противоречиям, с которыми оно сопряжено.

ТОРГ

Торг — это процесс достижения соглашения с помощью достоверных угроз и обещаний. Супруг может угрожать подать в суд для получения единоличной опеки над ребенком, пока второй не согласится на совместную опеку. В процессе торга о заработной плате рабочие могут угрожать забастовкой, пунктуальным соблюдением всех правил контракта или отказом от сверхурочных, тогда как работодатели могут грозить отказом обеспечивать сотрудников работой или закрытием предприятия. Менеджмент компании может угрожать работнику увольнением, если он не станет больше работать. Одна страна может угрожать другой вторжением, если та не пойдет на территориальные уступки. На Учредительном собрании делегат от одной из территориальных единиц может угрожать покинуть зал, если не будет принят порядок представительства, выгодный для этой единицы. Американские сенаторы могут угрожать обструкцией, если президент не отзовет предложенную кандидатуру. Конгресс может угрожать отказом от голосования по бюджету, если президент накладывает вето на один из законопроектов.

Если обратиться к обещаниям, член группы, принимающей решения голосованием, может пообещать высказаться за предложение, которое важно для одного из его коллег, при условии что тот проголосует за другое предложение, выгодное первому (взаимная поддержка). Продавец дома может пообещать не пересматривать договор, если покупатель заплатит назначенную цену. Точно так же похититель может отпустить жертву, получив выкуп, а не удерживать ее, требуя новый. Или наоборот, правительство может пообещать выпустить из тюрьмы террориста, как только его соратники отпустят похищенную жертву. Последняя может пообещать не рассказывать полиции о том, как выглядят похитители, если они ее отпустят. Участник «дилеммы заключенного» может пообещать сотрудничать, если другой поступит так же.

Результат торга зависит от достоверности угроз и обещаний. Угроза правдоподобна, если от агента можно ожидать (по любой причине), что он ее осуществит, если другой человек откажется соблюдать условия. «Твоя жизнь или твой кошелек» звучит более убедительно, чем «мой кошелек или моя жизнь». Обещание правдоподобно, если от агента (по любой причине) можно ожидать его выполнения в случае соблюдения условий. «Я создам иллюстрации, если вы написали текст» — правдоподобное обещание, если у давшего его человека есть стимул ему следовать. В игре «Доверительное управление», которую я обсуждал в главах V и XV, у доверенного управляющего нет материального стимула соблюдать обещание вернуть половину прибыли, созданной трансфером инвестора.

Достоверность *угроз* зависит как от объективных, так и от субъективных факторов. Объективные можно разделить на внешние и внутренние опции.

Внешняя опция — то, что получит сторона, если переговоры непредвиденным образом прекратятся. При переговорах о заработной плате внешняя опция для рабочих — зарплата, которую они могут получить в другой компании, или уровень льгот для безработных. При торге об опеке над ребенком внешние опции определяются ожидаемым исходом судебного разбирательства. На Конституционном конвенте в Филадельфии внешняя опция — то, что мог получить штат в случае выхода из союза (и возможного объединения с другой страной). Угроза отказаться от торга правдоподобна только тогда, когда сторона может убедительно показать, что вне данных отношений ей будет лучше, чем если она примет предложение контрагента.

Внутренняя опция — то, чем стороны располагают в процессе самого торга, в зависимости от того, как долго они могут его продолжать. Для рабочих внутренняя опция может определяться размерами забастовочного фонда. Для родителей она заключается в присуждении одному из них временной опеки над ребенком. В Филадельфии внутренняя опция определялась желанием делегатов вернуться к срочным делам в родных штатах. Внутренние опции влияют на достоверность угроз, поскольку они определяют, есть ли у стороны стимул их осуществлять. Угроза забастовки рабочих может показаться недостоверной, если (компания знает, что) у них нет забастовочного фонда, большинство из них женаты и должны выплачивать крупные взносы по ипотеке.

Говоря более обобщенно, угроза навредить другой стороне недостаточно правдоподобна, если при этом агент в значительной мере вредит самому себе. Члены конгресса не могут убедительно угрожать отказом от голосования по бюджету, если президент знает, что эта мера серьезно навредит как их репутации, так и способности решать вопросы. Член французского Учредительного собрания (Ланжуинэ) утверждал в 1789 году, что будущий законодательный орган будет не в состоянии использовать контроль над государственными доходами, чтобы помешать королю воспользоваться правом вето. Поскольку «перестать платить налоги» в этом случае значит «перерезать себе горло, чтобы вылечить рану на ноге», эта угроза не прозвучала бы убедительно.

С субъективной точки зрения угрозы более достоверны, если агент имеет (и другая сторона верит в то, что он имеет) долгий временной горизонт. Более терпеливые участники торга имеют преимущество перед оппонентами, поскольку те могут быть готовы обменять *размеры* уступки на *время*, в которое она делается²³. В проходивших в Париже перего-

²³ Это формулировка может в некоторой степени ввести в заблуждение. В игре с торгом, как на рис. XX.2, соглашение будет достигнуто сразу на основании обратной индукции, в которой играет роль коэффициент дисконтирования во времени сторон. Торг является «виртуальным», а не реальным. И все же более нетерпеливая сторона получит меньше, так как ей важнее получить все тотчас же.

ворах между США и Северным Вьетнамом последний сделал хороший первый ход, когда его делегация арендовала в Париже дом на два года. Тем самым вьетнамцы показали, что никуда не торопятся. Нейтральность по отношению к риску или низкая степень его неприятия тоже может дать преимущество в торге. Боящийся риска родитель, который имеет хорошие шансы получить полную опеку над ребенком по суду, тем не менее, соглашается на совместную опеку (см. главу XI). Следующий субъективный фактор, который может добавить угрозам достоверности, — придание относительно низкой ценности тому, что может предложить другая сторона, в сравнении с тем, от чего нужно отказаться. В торге при разводе один родитель может добиться благоприятной финансовой сделки, потому что второй крайне заинтересован в получении опеки над ребенком.

Наличие дальнего временного горизонта влияет на достоверность также еще одним способом. Если мне не удастся выполнить угрозу в настоящей ситуации, потому что ее цена слишком высока, от этого страдает моя способность выдвигать правдоподобные угрозы в будущем. Возможно, эту идею будет проще понять, если сформулировать ее противоположным образом: я могу выполнить угрозу в текущей интеракции, даже если у меня нет стимула это делать, чтобы создать себе репутацию человека, чьи угрозы следует принимать всерьез. Если другой человек понимает логику этого довода, я могу не выполнять никаких угроз вовсе. Создание репутации происходит только в «виртуальном», а не в прямом смысле. Проще говоря, мне может потребоваться выполнить несколько угроз, «противоречащих моим интересам», чтобы показать миру, что я именно такой человек. В комментариях к игре «Сетевой магазин» в главе XX я рассматривал обе эти возможности.

Достоверность *обещаний* также зависит от объективных и субъективных факторов, хотя и несколько иначе. Она в существенной мере предопределяется способностью агента его сдержать. Рассмотрим для примера провальную попытку взаимной поддержки во французском Учредительном собрании 1789 года. На трех встречах между лидером умеренных Мунье и радикалами Барнавом, Дюпором и Александром Ламетом последние трое сделали следующие предложения. Они дадут Мунье и абсолютное вето для короля, и двухпалатную систему, если взамен он согласится с лишением короля права распускать собрание, с предоставлением верхней палате лишь права временно приостанавливать рассмотрение закона и периодическими созывами конвента для пересмотра конституции. Мунье сразу же отказался якобы потому, что не верил в способность этой троицы выполнить свои обещания, так как на ассамблее не было партий в современном смысле этого слова — как дисциплинированных групп, которые можно заставить голосовать как единое целое.

В качестве еще одного примера рассмотрим обещания иммунитета от уголовного преследования для уходящих лидеров при переходе к демократии. Такого рода обещания делали, принимали и не выполняли в Аргентине в 1983 году, в Уругвае в 1984 и в Польше и Венгрии в 1989 году. (В странах Латинской Америки угроза военного переворота принуждала к выполнению условий.) Ретроспективно понятно, что генералы и партийные функционеры должны были сознавать недостоверность таких обещаний, поскольку вступающие в должность лидеры, с которыми они вели переговоры, не могли гарантировать соблюдение данных ими гарантий судами и законодательными органами. В Польше представители оппозиции, принадлежавшие к левому крылу «Солидарности», заявляли на круглом столе о необходимости следования принципу *«pacta sunt servanda»* — «договоры должны соблюдаться». Когда к власти пришло правое крыло движения, они проигнорировали это обещание.

Субъективные условия достоверности также включают долгий временной горизонт. Если у агента низкий коэффициент дисконтирования во времени и он знает, что ему снова придется взаимодействовать с тем, кому он дал обещание, или с другими людьми, которые могут наблюдать за его поведением, у него есть стимул сдерживать слово. Таким предвосхищением событий может, например, поддерживаться взаимная поддержка в законодательных органах. (В конституционных собраниях, которые созываются лишь один раз, стимул выполнять обещания в рамках взаимной поддержки должен быть слабее и, как следствие, такие гарантии, скорее всего, предоставляются реже.) Агент также может создать для себя систему сдерживающих факторов, которые помешали бы ему нарушать обязательства. Хотя похитители обычно не имеют оснований верить своей жертве, когда та говорит, что не раскроет их личности полиции, обещание может быть подкреплено тем, что жертва сообщит о себе какую-то поддающуюся верификации компрометирующую информацию, которую похитители могут раскрыть в случае ареста.

Поскольку субъективные условия или ментальные состояния, которые влияют на исход торга, не могут наблюдаться непосредственно, участники торга заинтересованы в их искажении посредством вербальных или невербальных форм поведения. Насколько мы знаем, переговорщики из Северного Вьетнама арендовали дом на длительный срок только для того, чтобы создать впечатление, что они не спешат. При взаимной поддержке каждая сторона будет преувеличивать важность того, от чего ее просят отказаться, чтобы выторговать большую уступку у другой стороны. Когда рабочие подчеркивают значение, придаваемое ими дорогостоящим мерам охраны труда, это может являться всего лишь уловкой, чтобы оправдать большую прибавку к заработной плате как цену



РИС. XXV.1

за снятие этих требований. Во многих случаях попытки обмана слишком прозрачны, чтобы сработать. Если разводящийся родитель делает вид, что очень озабочен опекой над ребенком, чтобы добиться более выгодных финансовых условий, вторая сторона может документально подтвердить отсутствие интереса к детям до того, как брак начал разваливаться, засвидетельствовав, к примеру, недавний переход на работу, предполагающую частые командировки. Однако прозорливый родитель может предвидеть эту проблему и заложить фундамент для притязаний на заботу о ребенке задолго до того, как другой супруг поймет, что брак разваливается.

Подобно сторонам в споре, участники торга могут иметь стимул представлять свой интерес в качестве основанного на принципе. Однако у подобного искажения иная причина. В споре стороны хотят отвлечь осуждение за то, что их предложения покоились на голой корысти. В торге же выражение интереса не вызывает осуждения. Предполагается, что компании и рабочие озабочены выгодами и заработной платой, а не общественным благом. Тем не менее участники торга могут получить стратегическое преимущество, сформулировав свои требования как принципиальные. Они могут сделать вид, что, отступая от них, идут на большие жертвы, и исходя из этого требовать от другой стороны

больших уступок в сравнении с теми, которые они получили бы, присутствуя в их деле лишь чистый интерес. Однако, если такую тактику будут использовать обе стороны, торг может быть сорван.

РЕЗЮМЕ

Если резюмировать различные тенденции, отмеченные в этой главе, процесс коллективного принятия решений может быть описан так, как это изображено на рис. XXV.1. Центральная идея состоит в том, что каждый из механизмов коллективного принятия решения — спор, голосование и торг — создает стимулы для искажения некоторых аспектов предпочтений. Другими словами, механизм обобщения сам вносит вклад в формирование вводных данных, поступающих в этот самый механизм. Выраженные политические предпочтения находятся в зависимости как от реальных политических воззрений, так и от механизма, который обобщает эти выраженные воззрения. Влияние искажения на благополучие двояко. Благодаря цивилизующей силе лицемерия эффект может оказаться социально благотворным. В других случаях обобщенное использование этой тактики может создать ситуацию, напоминающую «дилемму заключенного», в которой проигрывают все.

Библиографические примечания

Относительно механизмов, в некоторых случаях опосредующих коллективное действие и коллективное принятие решений, см. «Управление общей собственностью: эволюция институтов коллективного действия» Э. Остром (Ostrom E. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press, 1990). Касательно стран с системой принудительного голосования и применяемых ими санкций против абсентеистов см. http://www.idea.int/vt/compulsory_voting.cfm. Блестящее, хотя местами чересчур эксцентричное, рассмотрение спора и голосования содержится в работе И. Бентама «Тактика законодательных собраний» (Бентам И. *Тактика законодательных собраний*. Челябинск: Социум, 2006). Прочитированные отрывки из Бентама взяты из не менее интересной книги «Права, представительство и реформа» (Rights, Representation, and Reform. Oxford University Press, 2002. P. 35, 122). По поводу представления в ложном свете, вызванного принятием решения, см. главу 5 моей «Алхимии ума» (Alchemies of the Mind. Cambridge University Press, 1999). В связи с различием между обычным законодательным органом и конституционным собранием см. мою статью «Конституционные законодательные органы» (Constituent legislatures // Bauman R., Kahana T. (eds). *Constitution and the Legislature*. Cambridge University Press, 2006). Относительно голосования по двух-

палатной системе в 1789 году см. Ж. Эг্রে «*Революция нотаблей*» (Egret J. *La revolution des notables*. Paris: Armand Colin, 1950). Процедура «двойного сбора» обсуждается в статье К. Листа «*Дискурсивная дилемма и публичное обсуждение*» (List C. *The discursive dilemma and public reason* // *Ethics*. 2006. Vol. 116. P. 362–402). Обсуждение теоремы Кондорсе основывается на работе Д. Кароткина и Дж. Паруша «*Оптимальный размер собрания: дилемма качества-количества*» (Karotkin D., Paroush J. *Optimum committee size: Quality-versus-quantity dilemma* // *Social Choice and Welfare*. 2003. Vol. 20. P. 429–441). Полное обсуждение истории тайного голосования можно найти в работе Х. Бухштейна «*Открытое и тайное голосование*» (Buchstein H. *Öffentliche und geheime Stimmabgabe*. Baden-Baden: Nomos, 2000). Замечание по поводу Закона о гражданских правах более или менее дословно цитируется по работе Д. Брэди и Дж. Фереджона «*Конгресс и политика гражданских прав: исследование эндогенных предпочтений*» (Brady D., Ferejohn J. *Congress and civil rights policy: An examination of endogenous preferences* // Katznelson I., Weingast B. (eds). *Preferences and Situations*. New York: Russell Sage, 2005). По поводу обсуждения стратегического голосования см. главу 2 в книге Д. Остена-Смита и Дж. Бэнкса «*Позитивная политическая теория II*» (Austen-Smith D., Banks J. *Positive Political Theory II*. Ann Arbor: University of Michigan Press, 2005). Книга Дж. Макки «*В защиту демократии*» (Mackie G. *Democracy Defended*. Cambridge University Press, 2003) содержит пространный анализ круговых социальных предпочтений, утверждающий, что почти все примеры хождения по кругу в законодательных органах основываются на неправильном толковании фактов. Пример с аэропортом в Осло взят из работы А. Хилланда «*Парадокс Кондорсе в теории и на практике*» (Hylland A. *The Condorcet paradox in theory and practice* // Elster J. et al. (eds). *Understanding Choice, Explaining Behavior: Essays in Honour of Ole-Jørgen Skog*. Oslo Academic Press, 2006). Пример с демобилизацией американских солдат взят из книги под редакцией С. Стоуффера «*Американский солдат*» (Stouffer S. (ed.). *The American Soldier*. Princeton, NJ: Princeton University Press, 1949). Плодотворная работа на тему торга — Т. Шеллинг «*Стратегия конфликта*» (Шеллинг Т. *Стратегия конфликта*. М.: ИРИСЭН, 2007). Классическая работа по теме торга в действии — Х. Райффа «*Искусство и наука переговоров*» (Райффа Х. *Искусство и наука переговоров*. М., 1982). Неформальным введением в теорию торга является моя книга «*Цемент общества*» (The Cement of Society. Cambridge University Press, 1989).

XXVI. Организации и институты

Примерами организаций являются местный супермаркет, General Electric, Американский конгресс, католическая церковь, немецкий Конституционный суд, Франция и Организация Объединенных Наций. Примерами институтов являются семья, конституции, бумажные деньги, рыночная экономика и правосудие переходного периода (*transitional justice*). Организации и институты можно соотнести с «экземпляром» и «типом», то есть частным случаем идеи и самой идеей. Биржа — это институт (тип), Нью-Йоркская биржа — это организация (экземпляр этого института). Национальное государство — это институт, Франция — организация. Некоторые организации, такие как ООН, не являются экземплярами институтов. И наоборот, некоторые институты (например, рыночная экономика) не имеют организаций-экземпляров. Специальные рынки, однако (например, рынок квот на загрязнение окружающей среды), могут быть организациями. Хотя я буду рассматривать главным образом организации, их легче понять именно на фоне институтов.

Организации — это *коллективные акторы*, определяемые своей способностью к коллективному принятию решений. В некоторых случаях такая способность может выражаться в отсылке к конкретному моменту времени — например, 4 июля 1776 года для США или 24 октября 1945 года для Организации Объединенных Наций. В других случаях — Франция или католическая церковь — об организациях следует говорить как о *возникших* или *развившихся*, а не созданных. Иногда, тем не менее, трудно отследить момент, когда слабоструктурированное объединение индивидов кристаллизуется в коллективного актора с сохраняющейся во времени идентичностью — просто наступает время, когда мы можем сказать, что это произошло.

Институты — это *средства создания вещей*, образующихся или развивающихся постепенно, хотя как только они приняли определенную форму, они могут целенаправленно использоваться и применяться организациями. В качестве примера рассмотрим переходный период в Восточной Европе в 1989–1990 годах. Новые демократии приняли институт рыночной экономики и создали для его осуществления фондовые биржи и другие организации. В свои конституции они включили положение об институте конституционных судов, наделенных полномочиями отменять

неконституционные законы. Имея дело с автократическими предшественниками, они учредили институты правосудия переходного периода.

Эти три института (рыночная экономика, конституционные суды и правосудие переходного типа) существовали не всегда. Хотя Адам Смит, возможно, не так уж и ошибался, когда утверждал в качестве универсального факта «некоторую склонность человеческой натуры... вести обмен, выменивать одну вещь на другую», эту склонность могут сдерживать другие люди. В частности, она может блокироваться недоверием, которое характеризует «войну всех против всех» в гоббсовском «естественном состоянии». Чтобы преодолеть это недоверие, могут потребоваться институты. В примитивных условиях «обмен» может оказаться слишком рискованным, если сильный может просто отнять имущество слабого, ничего не предложив взамен. Тогда лучше держать свое имущество при себе, подальше от чужих глаз. Даже если основная проблема преодолена при помощи полиции, любовь к «безбилетничеству» и недоверие по-прежнему могут создавать серьезные препятствия для эффективного функционирования рынка. Если у меня есть хорошая идея относительно того, как дешевле производить определенный продукт, я могу не спешить ее воплощать, зная, что ею могут воспользоваться другие, сэкономя тем самым на затратах на разработку. Ответом на эту проблему становится институт патентной системы. Если заемщики хотят использовать свой дом и землю в качестве залога, залогодавцы появятся только тогда, когда будут уверены, что собственность не была перезаложена несколько раз. Чтобы решить эту проблему, государство может создать кадастр, в который они смогут заглянуть прежде, чем предоставлять заем. Для того чтобы возникли эти институты, которые в теперешней усовершенствованной форме могут быть прямо заимствованы новыми государствами, потребовались столетия.

Сегодня может показаться очевидным, что, если у страны есть конституция, ей также нужен политический механизм, позволяющий отменять неконституционные законы. Обычно эта задача (судебный надзор) ложится на суд, будь он смешанным, занятым также решением других вопросов, как в США, или специализированным конституционным, как на европейском континенте. Однако исторически конституции существовали задолго до конституционных судов¹. Хотя Конституция США учредила Верховный суд, который был наделен полномочиями по отмене законов штатов, у него нет четко прописанных полномочий объявлять

¹ В Нидерландах до сих пор нет судебного надзора. В Швеции он ограничен конституционной оговоркой, что закон должен быть «очевидным образом» неконституционным, чтобы суд мог его отменить. Предполагается, что такие законы редко проходят.

неконституционными федеральные законы. Когда этот суд в 1803 году присвоил себе эти полномочия в деле Марбери против Мэдисона, это, по сути, был акт самонастройки системы. В 1854 году Верховный суд Норвегии также взял на себя осуществление судебного надзора, который не упоминается в конституции страны. В 1971 году французский Конституционный совет, вопреки открыто выраженным намерениям авторов действующей конституции, позволил себе отменить законодательный акт, нарушавший Декларацию прав человека и гражданина 1789 года. Верховный суд Израиля также позволил себе отменить акт, нарушавший один из «Основных законов» страны, тем самым превратив частичную конституцию в полноценную. Хотя каждое из этих решений в свое время носило противоречивый характер, все они, кроме последнего, стали неотъемлемыми элементами институционального ландшафта.

Сегодня существует представление, что лидеры репрессивных режимов должны ответить перед судом после падения режима. Если новые лидеры страны не проявят инициативы, эту задачу может взять на себя международное сообщество (Руанда и бывшая Югославия). Оно также может оказывать давление на новый режим (Индонезия) или создать смешанные суды с участием местных и международных представителей (Сьерра-Леоне и Камбоджа). Сегодня существует своего рода правосудие переходного периода. Но если мы обратимся к прошлому, то увидим, что внесудебные расправы в целях мести или отстранения от власти всегда были самым распространенным способом борьбы с лидерами репрессивных режимов. Решение союзников в 1814 и 1815 годах не казнить Наполеона, а отправить его в ссылку, было важным шагом в процессе «останавливания руки возмездия». В конце Второй мировой войны многие союзные лидеры хотели, чтобы нацистские преступники были показательно расстреляны как люди, стоящие вне закона, а не предстали перед судом. В конце концов, однако, Нюрнбергский процесс показал пример почти образцового следования потребности выносить приговор на основании фактов и соблюдать нормы отправления правосудия. Некоторые из обвиняемых были в итоге оправданы. В некоторых странах, оккупированных немцами, в 1944–1945 годах присутствовали элементы самосуда, но их очень скоро вытеснила власть закона. Когда спустя сорок пять лет пали коммунистические режимы, все согласилось, что, разбираясь с преступлениями, нельзя подражать их незаконной практике.

ПРОБЛЕМА МОНИТОРИНГА

Организация может иметь *членов* или *сотрудников*. Члены также могут быть сотрудниками, как в рабочих кооперативах. Члены взаимодействуют горизонтально, через процессы спора, торга и голосования, которые

я рассматривал в предыдущей главе. Вертикальные отношения между сотрудниками и их начальниками носят иной характер. Для упрощения предположим, что в организации имеется только один руководитель («принципал») и множество сотрудников («агенты»). Проблема *принципал-агент* возникает, когда, как это часто случается, их интересы расходятся. Рабочие могут быть заинтересованы в умеренном темпе работы, тогда как менеджер, возможно, хочет заставить их прикладывать больше усилий. Если он попытается совместить эти два интереса, заплатив рабочим в зависимости от приложенных усилий, у него могут возникнуть трудности с мониторингом их работы или же этот процесс может оказаться слишком дорогостоящим. Традиционно это входило в задачи бригадира или мастера, но как менеджер может быть уверен, что бригадир не требует от рабочих взятки или не пользуется своей властью для удовлетворения собственных финансовых и сексуальных потребностей? В XIX веке рабочее место часто характеризовалось «тиранией начальника». В таком случае, кто будет надсматривать за надсмотрщиками? Как мы вскоре увидим, такие проблемы могут возникать, даже когда принципалы и агенты совпадают как в рабочем кооперативе.

Сходным образом глава государственного учреждения заинтересован в том, чтобы сотрудники были честными, то есть не брали и не вымогали взятки у населения. Он также заинтересован в эффективности, чтобы размеры общественного сектора были сведены к минимуму. Сотрудники могут иметь противоположные интересы в обоих отношениях. Если они мотивированы только своими эгоистическими экономическими интересами, они будут брать взятки, когда это сходит им с рук. Как следствие их заинтересованности во власти, у агентов также появится стимул раздувать штат своих департаментов и увеличивать количество подчиненных. И снова мониторинг может оказаться очень сложным. Иногда принципал может поймать за руку берущего взятку агента, но на этот метод нельзя полагаться. Он может попытаться сократить возможности для коррупции, например, проводя тендер для государственных контрактов, но это предосторожность не поможет, если агенты составляют контракты таким образом, чтобы они были выгодны определенным поставщикам. Поскольку агенты часто обладают почти монополией на информацию, принципал может оказаться не в состоянии решить, какие заявки на увеличение штата оправданны, а какие нет.

Не одни только подчиненные могут иметь стимулы, расходящиеся с интересами организации. Президентам американских университетов пришлось уйти в отставку, когда выяснилось, что они назначали себе большие зарплаты или делали ремонт у себя дома за счет организации. Один из американских вице-президентов (Спиро Эгню) вынужден был уйти в отставку, когда против него были выдвинуты обвинения в кор-

рупции. Клептократия (правление воров) стала слишком знакомым явлением во всем мире. Хотя в каком-то смысле они являются принципами, эти лидеры тоже могут выступать объектами мониторинга. Однако надзиратели (избиратели, советы попечителей, акционеры, Всемирный банк или МВФ) зачастую терпели очевидную неудачу в регулировании поведения своего топ-менеджмента или глав государства. Как и в других случаях, им может не хватать либо *информации*, необходимой, чтобы сглаживать перекосы, либо *стимулов* это делать.

В рабочих кооперативах может возникнуть конфликт между рабочими как принципами и рабочими как агентами. Выступая на Конгрессе социальных наук в 1863 году, сэр Джеймс Кей-Шаттлворт сказал, имея в виду ланкширскую кооперативную бумагопрядильную фабрику:

[Тогда] встал замечательный вопрос: какие выгоды должны получать дольщики этой фабрики помимо обычных? Первым требованием, выдвигавшимся в таких обществах, было требование отдавать предпочтение семьям дольщиков, чтобы рабочие фабрики отбирались из их числа. <...> На примере своей собственности он наблюдал провал одного из таких начинаний. Существовал соблазн поставить под вопрос принцип кооперации в этом смысле, так как дольщики должны получать преимущественное право найма их семей на работу на фабрике. Прямым следствием этой меры было то, что вместо соблюдения строгой дисциплины и внимательного отношения к машинам, столь необходимого на бумагопрядильных фабриках (он может также упомянуть, что дисциплина в полку бывает менее строгой, чем на бумагопрядильной фабрике), на своих квартальных и полугодовых собраниях рабочие заявляли против надзирателей самые гневные жалобы, а надзиратель, осмелившийся уволить рабочего-дольщика, сталкивался с опасностью быть отставленным на следующем собрании².

Еще одна распространенная проблема кооперативов возникала из нежелания сокращать персонал в периоды низкого спроса. Комментируя упадок «Волверхемптон Плейт-Локсмита», в 1878 г. современник писал:

Если бы дело велось частным мануфактурщиком, он бы, возможно, уволил работников, которым, из-за падения спроса на товар, он не находил достойного применения, и приложил бы все силы к развитию

² В 1875 году исполнительный директор Осборнской кооперативной двигателестроительной компании озвучивал подобные жалобы, так объясняя провал компании: «Правила, касавшиеся графика в Осбурне, включая ежедневный подсчет сверхурочных, чрезмерные ограничения на труд в нашей секции, небольшая доля подмастерьев, частые споры и препирательства, чрезмерное количество сверхурочных, требовавшееся, чтобы мы могли закончить работу по контракту, и, в целом, недостаток дисциплины, возникавший из представления, что каждый сам себе хозяин, — все это увеличивало стоимость продукции».

оставшейся торговли. Но от рабочих, которые образовали акционерное общество, это потребовало бы самопожертвования, к которому они не были готовы. Вместо этого они работали на склад в надежде на оживление спроса. А поскольку тот так и не ожил, тогда как их ресурсы оказались исчерпаны, они с неизбежностью пришли к печальному концу. Долги множились, лучшие работники уходили.

РЕШЕНИЯ ПРОБЛЕМЫ ПРИНЦИПАЛА И АГЕНТА

С халтурой, раздуванием штатов, коррупцией и прочими похожими проблемами можно бороться, воздействуя либо на стимулы, либо на возможности (см. главу IX). Последнее решение трудно проводить в жизнь. Чтобы быть эффективным, агент нуждается в независимости и свободе действия. Рабов редко использовали для труда, требующего прилежания и аккуратности. Может оказаться невозможным выстроить ситуацию таким образом, чтобы агент мог осуществлять цели принципала и при этом был не способен преследовать собственные. К этой задаче можно приблизиться, если, например, те, кто принимает решения, будут оставаться в своей должности столь короткое время, что их будет трудно подкупить. Система судов присяжных и американские коллегии выборщиков (в первоначальном замысле) были оправданы (среди прочего) именно по этим основаниям. Избрание на короткий срок (часто в сочетании с запретом на переизбрание) и частая ротация назначаемых чиновников также должны снижать возможности коррупции. Вышеназванные методы, однако, слишком затратны с точки зрения эффективности. К тому времени, как только чиновники достаточно освоят свою работу, чтобы хорошо ее выполнять, им, возможно, уже придется ее покинуть.

Более перспективным может оказаться воздействие на стимулы. Путем поощрения усердного труда или применения штрафов за халтурную работу стимулы агента и принципала могут быть приведены в соответствие. В профессиях, где индивидуальные усилия могут быть измерены объемом произведенного продукта, проблему можно решить с помощью сдельной оплаты труда. Однако иногда технология позволяет компании измерять только производительность труда группы рабочих. Кроме того, использование индивидуальной сдельной оплаты может оказаться контрпродуктивным. Если вы платите каждому члену футбольной команды (или нападающим) в зависимости от количества забитых голов, выступление команды в целом может оказаться неэффективным. Хороший командный игрок — тот, кто отдаст пас другому, если тот находится в лучшем положении. В этом случае компания может решить проблему, установив систему командных бонусов, чтобы рабочие имели стимул следить за работой друг друга.

И все же, если — как это часто случается — стоимость мониторинга превышает выгоды того, кто ведет наблюдение, у него может не оказаться *материальных* стимулов для его проведения. Футбольный игрок может решить, что не стоит труда (который может быть значительным) бранить товарища по команде за «звездное» поведение. Хотя может сложиться социальная норма, направленная против халтурщиков, эмпирические свидетельства на этот счет неоднозначны. Нормы складываются скорее внутри небольших групп, потому что их членам легче наблюдать друг за другом и каждому из них «безбилетничество» другого наносит гораздо больший вред. Но даже здесь их возникновение необязательно будет спонтанным (см. главу XX). Нормы против звездного поведения или эгоистической игры в футболе могут быть тесно связаны с политикой тренера, наказывающего игроков, которые слишком много думают о себе и слишком мало — о своей команде. Хотя есть анекдотические примеры, когда игроки наказывают тех, кто не сделал пас находящемуся в лучшем положении партнёру по команде, тем, что тоже не дают им пас.

Вопрос о том, кто будет надзирать за надзирателем или наблюдать за наблюдателем, теоретически может быть решен с помощью своего рода *круговой системы*, в которой каждое лицо в цепочке имеет прямой материальный стимул наказывать человека слева, потому что если он не сделает этого, то сам будет наказан человеком справа. На практике подобных схем не существует. Футболисты передают мяч игрокам, которые отдают его тем, кто не передавал пас. Однако соблазн «безбилетничества» может быть смягчен *доверием*. Если компания предлагает зарплаты выше существующего уровня, рабочие (или некоторые из них) могут в ответ прилагать максимум усилий. Загвоздка здесь состоит в том, что у компании может не оказаться стимула вводить такую схему. Хотя общий выпуск продукции (и, следовательно, общий доход) увеличится, возрастут и расходы на заработную плату. Если в результате прибыль снизится, компании выгоднее предлагать конкурентный уровень заработной платы.

ОРГАНИЗАЦИЯ НЕДОВЕРИЯ

Политические конституции часто стремятся к тому, чтобы быть круговыми системами сдержек и противовесов, в которых все институты ведут наблюдение за деятельностью друг друга. Это стремление, однако, не нашло реализации, возможно, ни в одной из существующих конституций. Конституционный механизм обычно покоится на «недвижимом двигателе» или «непроверяемом проверяющем», части, которая наблюдает за другими, сама не будучи объектом наблюдения. В США это

положение занимает Верховный суд. В других странах это парламенты. Например, в Румынии парламент может отменить решение конституционного суда, если тот отменил его законодательный акт (это наследие коммунистической эпохи, когда парламент был формально всемогущим, а де-факто бессильным). Во Франции парламент также был недвижимым двигателем вплоть до 1971 года. И все же, по крайней мере, в англо-американской традиции конституционный замысел покоится на принципе, сформулированном Давидом Юмом: «Это... справедливая политическая максима, что каждый человек должен считаться жуликом». В такой перспективе конституции представляют собой формы организованного недоверия.

Американская конституция является, возможно, наиболее детально разработанным выражением такого отношения. Поразительные черты дебатов на Конституционном конвенте 1787 года и документа, который был принят в его ходе, — постоянная озабоченность возможностью коррупции, взяток и угроз, а также необходимостью уничтожить любые лазейки для их возникновения. Приведем один пример. Специальный комитет по представительству рекомендовал 9 июля, чтобы, «учитывая, что нынешнее положение Штатов может измениться как по уровню благосостояния, так и по количеству жителей, законодательная власть получила бы разрешение время от времени увеличивать число представителей». В ответ Эдмунд Рэндольф выразил «тревогу, что, поскольку количество не будет меняться, пока этого не захочет национальный законодательный орган, не будет недостатка в предложениях для того, чтобы отложить изменения и оставить власть в руках тех, кому она принадлежит». Сходным образом Джордж Мейсон утверждал: «Судя по человеческой природе, мы можем утверждать, что те, в чьих руках власть, не отдадут ее, пока могут ее удерживать». Мягко говоря, мы можем допустить вместе с Юмом, что те, кто стоит у власти, *должны* пытаться ее удержать. Хотя некоторые политики могут работать (и все будут утверждать, что работают) ради общего блага, мы должны экономно использовать доверие. Грубо говоря, мы должны заключить, что те, у кого в руках власть, будут пользоваться ею для ее удержания и по возможности *расширения*. Конституции в такой перспективе — это механизмы, призванные воспрепятствовать этому. Не будучи единственной функцией, которую выполняют конституции, эта зачастую является одной из тех вещей, на которую они нацелены.

Способность этой нормативной рамки *объяснить* принятие реальных конституций зависит от того, в какой степени ее создатели заботятся об общих, а не о своих частных интересах или страстях. Как я отмечал в конце главы XIII, ни в коем случае не следует принимать на веру то, что они якобы свободны от такого рода мотиваций. В то же время нор-

мативное давление, заставляющее действовать так, как будто они руководствуются лишь доводами рассудка, — цивилизующая сила лицемерия может повлиять на конституционное собрание сильнее, чем на обычный законодательный орган. Кроме того, как я утверждал в предыдущей главе, преследование эгоистических интересов оказывает меньшее влияние на решения в процессе принятия конституции. По этой причине конституционные собрания, работающие за закрытыми дверями, часто способны уменьшить искажающее действие страстей, не поплатившись за это торгом, слишком сильно окрашенным корыстью³.

С учетом данного пояснения позвольте мне привести условный пример того, как организованное недоверие можно было бы воплотить в серии взаимосвязанных мер. Многие из этих черт можно найти в современных системах. Остальные включены только для того, чтобы проиллюстрировать, сколько труда придется приложить тому, кто захочет устранить все возможные лазейки.

1. Чтобы предотвратить *вмешательство* правительства в вопросы политической справедливости, конституция должна требовать, чтобы уголовные дела рассматривались судами присяжных или чтобы судьи определялись на дела методом случайного выбора (либо с помощью какой-то еще механической процедуры). При отсутствии положений такого рода правительство может судить своих политических оппонентов и обеспечивать обвинительные приговоры при помощи прикормленных судей.

2. Чтобы помешать правительству *манипулировать избирательной системой*, избирательное законодательство должно быть прописано в конституции, а не определяться обычным законодательством. Изменение границ избирательных округов с учетом изменений в составе населения также должно производиться с помощью механической формулы или воз-

³ Я не вполне разделяю негативный взгляд на страсть. Гегель говорил, что «все великие вещи в мире делались со страстью». Токвиль писал, что во времена кризиса люди либо падают ниже, либо поднимаются выше нормального уровня. Более того, беспристрастность не исключает страсть, как это продемонстрировали французские законодатели в 1789 году, испытывая «опьянение бескорыстием». Один из самых выдающихся из них, граф де Клермон-Тоннер сказал: «Анархия — страшный, но необходимый переход и единственный момент, когда можно установить новый порядок вещей. В тихие времена невозможно принять единых мер». Но даже если энтузиазм ведет к превосходному результату, он может подорвать трезвое мышление, необходимое для его достижения.

лагаться на независимую комиссию. Недавняя американская и французская история изобилует примерами того, как действующее правительство изменяло избирательное законодательство, чтобы остаться у власти.

3. Чтобы помешать правительству *манипулировать потоком информации*, государственные радио и телевидение должны управляться независимой комиссией («британская модель»), а не комиссией, назначаемой правительством («французская модель»). Должна быть также обеспечена защита частных средств массовой информации. Эфирные частоты должны распределяться независимыми комиссиями. Правительству должно быть запрещено распределять такие жизненно необходимые для прессы ресурсы, как бумага и типографская краска, или же такие решения должны возлагаться на «независимый совет по фондированию». Как показал опыт, сочетание свободных от манипуляции выборов и прессы очень эффективно для предотвращения голода и других бедствий. Если те, у кого есть соблазн действовать против интересов общества, знают, что об этом станет известно электорату, который может отправить их в отставку, вероятность совершения подобных действий уменьшается.

4. Чтобы помешать правительству *манипулировать денежной политикой* как инструментом собственного переизбрания, эта политика должна быть прерогативой независимого центрального банка. Альтернативное, но менее привлекательное решение: пусть решения по поводу экономической политики принимает компьютерная программа, созданная тремя экономистами, наугад выбранными в пяти ведущих экономических вузах страны, пересматриваемая каждые пять лет.

5. Чтобы воспрепятствовать тенденциозным и догматичным центральным банкирам *проводить губительную денежную политику*, должна быть предусмотрена возможность их отставки, если за нее выступает подавляющее большинство парламента (2/3 или 3/4). У них должен быть ограниченный срок пребывания в должности (не так, как у бывших председателей итальянского Центрального банка), и они должны назначаться лишь на один срок (не так, как нынешний итальянский глава Центрального банка).

6. Чтобы помешать правительству *манипулировать статистической информацией* как инструментом для переизбрания,

сбор информации должен быть поручен независимому бюро по статистике. В Британии канцлер казначейства принял меры, чтобы сделать Управление национальной статистики независимым от правительства («Файнэншл таймс» от 2 декабря 2005 г.). Заголовок в «Нью-Йорк таймс» (от 22 сентября 2002 г.) гласил: «Некоторые эксперты опасаются политического давления на уголовно-статистические агентства». Также сообщалось о политическом давлении на статистику занятости [«В Бюро трудовой статистики должно быть проведено расследование» («Нью-Йорк пост» от 29 сентября 2000 г.)]. В более тонких случаях выбор модели эконометрического прогнозирования тоже имеет значение. Ясно, что чем больше людям кажется, что в стране ниже преступность и безработица и больше перспективы роста, тем больше шансов на переизбрание власть предержащих. В дополнение к запрету на использование денежной политики для манипулирования реальностью необходимо запретить правительствам пользоваться статистикой для манипулирования представлениями об этой реальности.

7. Чтобы воспрепятствовать *подавлению оппозиции* правительством, основные политические партии должны получать дотации в определенной пропорции от доли их голосов и количества мест в парламенте. Размер дотаций должен быть зафиксирован в конституции и может быть привязан к цене на золото⁴.

8. Чтобы помешать правительству *принимать законы, обслуживающие его интересы*, в конституции должны быть определены некоторые категории законов, которые должны вступать в силу лишь через некоторое время (год или более) после принятия. При помощи этого механизма будет создаваться искусственная завеса неведения, за которой стороны, преследующие свои интересы, могут иметь стимул принимать справедливые решения.

9. Чтобы не дать правительству *обойти эти ограничения* посредством использования большинства в парламенте, для внесения поправок в конституцию нужно требовать квали-

⁴ Сходное стремление сделать конституцию максимально четкой и не подверженной манипуляциям стояло за желанием Мэдисона зафиксировать оклады федеральных судей, «взяв в качестве стандарта зерно или какую-либо другую меру постоянной ценности».

фицированное большинство, отсрочки и т. д. Сами статьи поправок должны быть неприкосновенными, то есть защищены от изменения.

10. Чтобы лишить правительство возможности *игнорировать эти ограничения*, как это было когда-то во Франции, конституция должна предусматривать механизм судебного надзора со стороны независимого конституционного или верховного суда.

11. Чтобы помешать правительству *манипулировать судебным надзором* путем увеличения числа судей в Конституционном или Верховном суде, это число не должно быть отдано на откуп законодателям, как в США и некоторых других странах, а должно быть зафиксировано в конституции.

12. Чтобы помешать догматичным или *идеологизированным* судьям в Конституционном или Верховном суде *игнорировать подавляющее народное большинство*, должна быть предусмотрена возможность отправлять их в отставку, если за это выступает квалифицированное большинство (две трети или три четверти) в парламенте. В отличие от практики, принятой в американском Верховном суде, срок их пребывания в должности должен быть ограничен. Кроме того, они не должны переизбираться.

Как показывает этот пример, конституция может быть крайне замысловатой структурой. При отсутствии одного промежуточного элемента остальные могут стать неэффективными или даже вредными (см. анализ ниже). В то же время есть реальная опасность того, что конституция может стать настолько негибкой, что мешает политической системе реагировать на чрезвычайные ситуации. Конституционные ограничения в денежной политике или избирательном законодательстве могут в конкретном случае иметь катастрофические последствия. На Конституционном конвенте в Филадельфии Джордж Мейсон отмечал, что хотя он смертельно ненавидел бумажные деньги, поскольку он не мог предвидеть все непредвиденные обстоятельства, ему не хотелось связывать руки легислатуры. Он заметил, что последняя война могла бы не произойти, если бы существовал такой запрет. Подобным образом, когда в 1946 году итальянский парламент решил не прописывать в конституции стабильность денежного обращения, одно из возражений заключалось в том, что во время войны государство должно иметь свободу действий. В 1945 году генерал де Голль путем манипуляций с французской избирательной системой оттеснил от власти коммунистов. Много лет спустя комменти-

руя предложение включить закон о выборах в конституцию, он сказал: «Никогда не знаешь, что может произойти. Может так случиться, что однажды снова возникнет необходимость вернуться к пропорциональной системе голосования в интересах нации, как в 1945 году. Мы не должны связывать себе руки». Если оглянуться назад, то Франция должна быть благодарна за это решение.

Проблема может быть резюмирована в двух знаменитых замечаниях, касающихся американской конституции. Комментируя Акт о ку-клукс-клане 1871 года, Джон Поттер Стоктон заметил, что «конституции — это цепи, которыми люди себя сковывают, будучи в здравом уме, чтобы не наложить на себя руки в дни безумия». В 1949 году Роберт Джексон заявил: «Билль о правах — не самоубийственный пакт». Не так-то легко воспрепятствовать тому, чтобы конституция, замышлявшаяся как средство предотвращения самоубийства, не обратилась бы однажды в самоубийственный пакт.

ПРОБЛЕМА ВТОРОГО ЛУЧШЕГО

Оставив этот вопрос в стороне, давайте теперь предположим, что если 12 перечисленных мною условий (или другие условия подобного рода) удовлетворены, политическая система в некотором смысле, который для наших целей необязательно уточнять, будет *оптимальной*. В реальности такой результат едва ли достижим. Трудно все сделать правильно. Однако можно стремиться максимально *приблизиться* к оптимуму при предположительно разумном допущении, что чем больше условий оптимальности удовлетворено, тем ближе мы к нему подходим. Это допущение ошибочно. При самых общих условиях неверно, что ситуация, в которой удовлетворяются многие, но не все требования оптимальности, обязательно или с большей вероятностью превзойдет ситуацию, в которой выполняется меньшее число условий. Если обратиться к 12 пунктам, предложенным выше, априори неясно, что лучше — судебный надзор без контроля над судьями или полное его отсутствие, независимые центральные банки без контроля над банкирами или система, передающая денежную политику в руки правительства.

Исследование Токвилем старого режима во Франции может рассматриваться в этой перспективе. Та система характеризовалась рядом черт, которые было бы трудно найти в хорошо организованном обществе. Королевская администрация была наделена широкими и неопределенными, неограниченными полномочиями. Принцип продажи должностей делал невозможной рациональную бюрократию. Блокирование работы *парламентов*, крайняя политизированность судов, руководствовавшихся преимущественно собственными интересами, делали затруднитель-

ным проведение последовательной политики. И все же, как утверждал Токвиль, с учетом первой из этих черт, присутствие двух других было в действительности благотворным:

Правительство в своем стремлении все обратить в деньги, пустив в продажу большую часть общественных должностей, само лишило себя возможности раздавать и отбирать их по своему усмотрению. Таким образом, одна из его страстей в значительной степени повредила развитию другой: его алчность явилась мощным противовесом властолюбию. Тем самым в своих действиях оно было вынуждено беспрестанно сталкиваться с ограничением, им же созданным, которое оно не могло подвергнуть разрушению. Правительству часто приходилось видеть, как самые решительные его намерения при исполнении утрачивали энергию и силу. Такой странный и порочный характер государственных должностей выступал своего рода политической гарантией против всемогущества центральной власти. Он был своеобразной беспорядочно построенной плотиной, сдерживающей силу правительства и ослабляющей его натиск. <...> Беспорядочное вмешательство судов в дела управления, часто нарушавшее заведенный администрацией порядок, иногда служило защитой свободы людей — это было зло, но оно умеряло зло еще большее.

В своем анализе английского законодательства XVIII века Джеймс Фитц-джеймс Стифен отмечал, что чрезмерный акцент на технические детали закона «смягчает, хотя и иррациональным, прихотливым образом, чрезмерную суровость старых уголовных законов». Комментируя исламское уголовное законодательство, он также писал, что оно «смягчало экстравагантную жестокость своих положений правилами доказательств, которые практически исключали любую возможность проведения этих положений в жизнь». Точно так же Османская империя, царский режим в России, Италия времен Муссолини и Испания эпохи Франко характеризовались как «деспотизм, смягченный бессилием». По сравнению с безжалостно эффективной нацистской Германией это действительно были мягкие режимы.

Токвиль также отмечал, что, когда чиновники старого режима попытались заменить ненавистную всем трудовую повинность на строительстве дорог (*corvée*) налогом, который шел бы на их поддержание, они оставили эти попытки, страшась того, что «как только будет образован такой фонд, ничто не помешает казначейству воспользоваться им для своих нужд, так что вскоре налогоплательщиков обложат новым налогом и при этом сохранят трудовую повинность». Это явление очень распространено. Когда политики предлагают государственным медучреждениям принимать в первую очередь тех, кому нужно помочь вернуться к работе, они обычно обещают направить соци-

альные выгоды от этого таким образом, чтобы они вернулись в сами медучреждения, отчего выиграли бы все пациенты. Администрация больниц обычно относится к таким предложениям скептически, подозревая, что прибыль оседет в правительственных сейфах. С таким же скептицизмом в Калифорнии были встречены предложения распределять воду тем фермам, которые эффективнее ею распорядятся, чтобы затем употребить образовавшийся излишек для улучшения водоснабжения всех остальных. Ситуация с двумя субоптимальными элементами — неэффективная система приоритетов в распределении скудных благ и правительство, не выполняющее свои обещания (или обещания своих предшественников), — может оказаться лучше, чем та, в которой первый элемент устранен.

Эти случаи напоминают те, когда политики устраняют животных, которые создают неудобства людям, а потом выясняется, что еще большее неудобство причиняют организмы, размножение которых сдерживали эти животные. Например, когда Мао Цзедун решил истребить воробьев, так как они массово пожирают зерно, ему позднее пришлось импортировать их заново из Советского Союза, когда вредители, которых они сдерживали, стали размножаться с катастрофическими последствиями для экологии. Общества не меньше, чем экосистемы, могут иметь внешне абсурдные или вредные черты, устранение которых, однако, наносит еще больший вред. Возможно, это одна из причин, по которой Эдмунд Бёрк и его последователи так настаивали на своей критике рационалистского институционального замысла.

Эту линию рассуждений можно продолжить очень далеко. Когда рабочие получили право голоса, противники этого заявляли, что устранение несправедливости по отношению к рабочим дестабилизирует общество в самых разных аспектах и создает еще большую несправедливость по отношению к богатым, чья собственность, как ожидается, может быть конфискована в пользу бедняков. Этот довод кладет на одни весы два набора соображений, обладающих разным весом. Но так как эта книга об объяснении социальных явлений, а не об их оценке с нормативной точки зрения, я не стану разбирать этот вопрос.

Библиографические примечания

Различие между организациями институтами взято из К. Оффе *«Роль институтов в распределении и контроле над социальными силами»* (Offe C. «Institutions» role in the distribution and control of social power // Shapiro I., Skowronek S., Gavlin D. (eds). *The Art of the State*. New York University Press, 2006). «Подъем рынков» является центральной темой книги Д. Норта *«Структура и изменение в экономической ис-*

тории» (North D. Structure and Change in Economic History. New York: Norton, 1981). Возникновение судебного контроля — тема работы под ред. Э. Смита «Конституционное правосудие при старых конституциях» (Smith E. (ed.). Constitutional Justice Under Old Constitutions. The Hague: Kluwer, 1995). Касательно возникновения юстиции переходного типа см. Дж. Дж. Басс «Остановить руку возмездия» (Bass G. J. Stay the Hand of Vengeance. Cambridge, MA: Harvard University Press, 2001). Полезным чтением по теме (определяемой мной несколько идиосинкратично) организаций — «Политическая экономия институтов» под ред. К. Менара (Ménard C. (ed.). The Political Economy of Institutions. Cheltenham: Edward Elgar, 2004). Проблема приведения в соответствие стимулов индивида и организации является темой книги Ж.-Ж. Лаффонта и Ж. Тироля «Теория стимулов в снабжении и регулировании» (Laffont J.-J., Tirole J. A. Theory of Incentives in Procurement and Regulation. Cambridge, MA: MIT Press, 1994). Всеобъемлющее руководство по вопросу коррупции — «Политическая коррупция» под ред. А. Хайденхаймера, М. Джонстона и В. Левиня (Heidenheimer A., Johnston M., LeVine V. (eds.). Political Corruption. New Brunswick, NJ: Transaction Publishers, 1989). Цитаты, касающиеся английских кооперативов XIX века, взяты из книги Б. Джонса «Кооперативное производство» (Jones B. Co-operative Production. Oxford University Press, 1894; New York: Kelley, 1968). Идея «крутовой системы» взаимного мониторинга обсуждалась в работе «Давление равных и партнерство» Э. Канделя и Э. Лазера (Kandel E., Lazear E. Peer pressure and partnership // Journal of Political Economy. 1992. Vol. 100. P. 801–817), а также в статье Р. Калверта «Рациональные акторы, равновесие и социальные институты» (Calvert R. Rational actors, equilibrium, and social institutions // Knight J., Sened I. (eds.). Explaining Social Institutions. Ann Arbor: University of Michigan Press, 1995) (reprinted in Ménard, The Political Economy of Institutions). Относительная важность доверия и стимулов в компаниях обсуждается в работе Е. Фера и А. Фалька «Психологическое основание стимулов» (Fehr E., Falk A. Psychological foundation of incentives // Europe and Economic Review. 2002. Vol. 46. P. 687–724). «Юмовские» истоки американской конституции прослеживаются в книге М. Уайта «Философия, федералист и конституция» (White M. Philosophy, The Federalist, and the Constitution. Oxford University Press, 1987). Цитата из де Голля взята из книги А. Пейрефита «Это де Голль» (Peyrefitte A. C'était de Gaulle. Paris: Fayard, 1994. Vol. 1. P. 452). Идея «второго лучшего» введена Р. Дж. Липси и К. Ланкастером в «Общей теории второго лучшего» (Липси Дж. Р., Ланкастер К. Общая теория второго лучшего // Вехи экономической мысли. Т. 4: Экономика благосостояния и общественный выбор / пер. с англ. под общ. ред. А. П. Заостровцева; Институт «Экономическая школа» ГУ ВШЭ. СПб.: Экономическая школа, 2004). Анализу старого режима у Токвиля посвящена моя статья «Токвиль о 1789 году: предварительные условия, катализаторы и спусковые механизмы» (Tocqueville on 1789: Preconditions, precipitants, and triggers // Welch C. (ed.). The Cambridge Companion to Tocqueville. Cambridge University

Press, 2006). Прочитированные фрагменты из Стифена содержатся в его *«Истории уголовного права в Англии»* (Стифен Дж. Уголовное право Англии в кратком начертании. СПб., 1865). История о Мао Цзэ-дуне взята из Дж. Чана и Дж. Халлидея *«Мао: неизвестная история»* (Chang J., Halliday J. Mao: The Unknown Story. New York: Knopf, 2005. P. 430–431). Я излагаю свои взгляды на институциональные изменения на нормативной основе в главе 4 *«Соломонова решения»* (Solomonic Judgments. Cambridge University Press, 1989).

ЗАКЛЮЧЕНИЕ: ВОЗМОЖНА ЛИ СОЦИАЛЬНАЯ НАУКА?

ЧТО СЧИТАТЬ НАУКОЙ?

Ответ на вопрос, возможна ли социальная наука, или, быть может, она уже существует, зависит от критерия, что считать наукой. Внешний критерий может быть следующим: дисциплина стала наукой, когда: 1) между теми, кто ею занимается, в любой момент есть общее согласие по поводу того, что истинно, что ложно, что предположительно и что неизвестно внутри ее области; 2) есть процесс кумулятивного прогресса, благодаря которому однажды отброшенные теории и объяснения остаются таковыми навсегда; 3) основные теории и объяснения могут быть выражены достаточно четко и ясно, чтобы быть понятыми любым, кто захочет потратить на это время и приложить усилия; 4) «классиков» дисциплины читают преимущественно историки науки. Как утверждал Альфред Уайтхед: «Наука, которая не торопится забыть своих основателей, безнадежна».

Это описание призвано в общих чертах охарактеризовать состояние современных естественных наук. Критерию 1 буквально и полностью не удовлетворяет ни одна из них. Так, ведутся споры по теории струн, прерывистом равновесии или соотношении природы и воспитания в развитии человека. Тем не менее глубина расхождений легко преувеличивается теми, кто находится вне соответствующего научного сообщества, недооценивает огромный массив всеми разделяемых знаний, которые образуют фон научного спора, и зачастую ошибочно принимает пробные вылазки за окончательные утверждения. Критерий 2, как правило, удовлетворяется. В истории естественных наук было немного, если вообще были, обратных поворотов, никакие неоньютонианцы не восставали против Эйнштейна, а неоламаркианцы — против Дарвина. Трансформация того, что в момент возникновения представлялось невразумительной или не очень внятной теорией, в легкоусваиваемый материал из учебников, указывает на общее удовлетворение критерию 3. За редким исключением, в основном в биологии (Дарвин, Д'Арси Вентворт Томпсон, Клод Бернар), критерий 4 также удовлетворяется.

Используя эти критерии, я хочу предложить оценку мягких (soft) качественных (qualitative) и количественных (quantative) социальных наук. После нескольких нелестных замечаний в адрес «мягкой» социальной науки я выступлю в защиту качественной социальной науки. Более неоднозначным образом я выступлю с критикой количественной социаль-

ной науки, по крайней мере в некоторых из ее наиболее заметных форм. Защитник качественной социальной науки оказывается, таким образом, в неудобном положении, так как сражается на два фронта с постоянным риском, что один из оппонентов обвинит его в союзе с другим.

«МЯГКАЯ» СОЦИАЛЬНАЯ НАУКА

«Мягкие» социальные науки больше похожи на некоторые формы литературоведения (или литературы), чем на эмпирические качественные исследования. Постмодернизм, постколониальная теория, теория подчинения (subaltern theory), деконструктивизм, кляйновский или лакановский психоанализ и прочие подобные теории не раз обвинялись многими авторами в обскурантизме, наиболее эффективно, возможно, Аленом Сокалом. Как он отметил в одном интервью, отсутствие общего языка для рациональной дискуссии делает невозможной лобовую критику этих псевдотеоретиков. Вместо этого приходится заставлять их самих пилить свой сук, как он самым действенным образом проделал это, опубликовав в одном из их журналов статью о герменевтике квантовой гравитации, полную бессмысленного, но витиеватого жаргона.

Хотя члены этих замкнутых кружков могут выработать в своей среде стиль обсуждения, частично удовлетворяющий критерию 1, это будет не более чем псевдоинтерсубъективностью, в которой общие языковые рефлексии мимикрируют под рациональные конвенции. Критерии 2–4 удовлетворяются редко, если вообще когда-либо. Ни одна идея не отбрасывается навсегда. Возможно, самое главное — неудовлетворение критерию 3. Деррида может ослеплять своим языком, но его «учению» *невозможно научить*. Нет никаких «самоучителей» по деконструкции (хотя есть множество исследований или «картографирований» и «обговариваний») главным образом потому, что его сторонники предпочитают изъясняться намеками и риторическими вопросами, вместо того чтобы, «высунувшись», сделать определенное утверждение. Возможно, больше всего на учебник похожи пародии Фредерика Крюса «Постмодернистский пух». Что касается критерия 4, культ предков с обязательными ссылками и толкованиями представляется обязательным (ясно, что в этом абзаце я пытаюсь обратить в свою веру уже обращенных).

КАЧЕСТВЕННАЯ СОЦИАЛЬНАЯ НАУКА

К качественной социальной науке я отношу основную массу исторических трудов, а также работ в форме «изучения конкретных случаев» (case studies) в противоположность крупномасштабным исследованиям. Я полагаю, что лучшая форма обучения для того, кто занимается социальной

наукой, — больше и глубже читать об истории, выбирая работы за внутренне качества их аргументации, а не по важности рассматриваемого предмета. Вот некоторые образцы: Джеймс Фитцджеральд Стифен «История английского уголовного законодательства» (Стифен Д. Уголовное право Англии в кратком начертании. СПб., 1865); Е. П. Томпсон «Становление английского рабочего класса» (Thompson E. P. The Making of the English Working Class); Г. Е. М. де Сент-Крюа «Классовая борьба в древнегреческом мире» (St Croix G. E. M. de. The Class Struggles in the Ancient Greek World); Джозеф Левенсон «Конфуцианский Китай и его судьба в современности» (Levenson J. Confucian China and Its Modern Fate); Поль Вен «Хлеб и зрелища» (Veyne P. Le pain et le cirque); Ж. Лефевр «Большой страх» (Lefebvre G. La grande peur); Кейт Томас «Религия и упадок магии» (Thomas K. Religion and the Decline of Magic); Токвиль «Старый порядок и революция» (Токвиль А. де. Старый порядок и революция. М.: Моск. философский фонд, 1997); Макс Вебер «Аграрная история Древнего мира» (Вебер М. Аграрная история Древнего мира. М.: Канон-пресс-Ц, Кучково поле, 2001); Гордон Вуд «Радикализм американской революции» (Вуд Г. Радикализм Американской революции. Американский ежегодник, 1994. М., 1995); Жан Эгре «Французская предреволюция» (Egret J. La pré-révolution française); Дени Крузе «Воины божьи» (Crouzet D. Les guerriers de Dieu); или Мартин Оствальд «От народного суверенитета к суверенитету закона» (Ostwald M. From Popular Sovereignty to the Sovereignty of Law). (Я в некотором роде «высунулся», включив сюда книги, пока еще не признанные классическими.) Эти и ряд подобных фигур объединяет сочетание авторитета в фактических вопросах со способностью обнаружить потенциал для обобщений, а также привести примеры, которые могли бы поставить эти обобщения под сомнение. Знания позволяют им улавливать «говорящие детали» и «значительные отклонения от нормы», создавая тем самым стимулы и одновременно возможности для потенциальных аналитиков по проверке теорий реальностью.

То же самое относится к авторам «исследований конкретных случаев», среди которых величайшим остается «Демократия в Америке» Токвиля. Я бы еще добавил сюда не совсем подпадающую под эту категорию «Капитализм, социализм и демократию» Йозефа Шумпетера (Шумпетер Й. Капитализм, социализм и демократия. М.: Экономика, 1995). Несколько неожиданным, но крайне интересным кандидатом является также «Путешествие по Франции, 1787, 1788 и 1789» Артура Юнга (Юнг А. Путешествие по Франции, 1787, 1788 и 1789. СПб.: ИНАПРЕСС, 1996). Это «характерные портреты» целых обществ или режимов, данные в компаративистской перспективе. «Феодалное общество» Марка Блока (Блок М. Феодалное общество. М.: Изд-во им. Сабашниковых, 2003) также входит в эту категорию. «Зияющие высоты» Александра Зиновьева

не совсем являются портретом советского коммунизма, но это карикатура в хорошем смысле слова — устраняющая все несущественное и выделяющая ключевые черты путем их преувеличения. Исчерпывающе документированной версией является «Черная книга коммунизма» С. Куртуа и др. (Куртуа С., Верт Н., Панне Ж.-Л., Пачковский А. и др. Черная книга коммунизма. М.: Три века истории, 2001). Трилогия Ричарда Эванса о Третьем рейхе обещает представить нацистский режим так, как в более обобщенном виде сделал это Роберт Пэкстон в книге «Что такое фашизм?» (Paxton R. What Is Fascism?). «Муссолини и муссолиниевская Италия» Ричарда Босуорта (Bosworth R. Mussolini and Mussolini's Italy), если ее читать параллельно с книгами Эванса, содержит удивительные догадки, касающиеся различия между режимом, в котором зло, пусть реальное, было преимущественно низкосортным, и режимом, который был злом по самой своей сути.

Исторический анализ и изучение отдельных примеров не только дают материал для обобщений, но сами по себе часто содержат неявные обобщения, к которым автор предлагает читателю прийти самостоятельно. Токвиль, например, был мастером латентных обобщений. В силу высокомерия историка и аристократа он не снисходил до того, чтобы эсплицировать свои теоретические построения, но делал вид, что просто совершает один шаг за другим. Несмотря на то что «Демократия в Америке» постоянно применяет при объяснении схему желанная-возможности, она нигде подробно не оговаривается. Хотя «Старый порядок и революция» широко использует различие между завистью и ненавистью при анализе классовых отношений, читатели должны сами догадаться о совершенно разных причинах и следствиях этих двух эмоций.

В категорию качественной социальной науки я также включаю работы, которые делают эти обобщения *явными*. При их обсуждении я не буду перечислять имена авторов, но вернусь к головоломкам, которые привел во введении. Крайне важно, что на все «разгадки» распространяется условие, которое я упоминал: они удовлетворяют лишь одному минимальному требованию — заключать в себе экспланандум. В некоторых случаях я позволю себе высунуться и сказать, что предложенное мною объяснение верно, но в других случаях — нет.

I. РАЗУМ

- Почему одни игроки полагают, что если пять раз подряд выпадало красное, то более вероятно, что именно оно, а не черное выпадет в следующий раз? *Ответ:* на них действует эвристика наличности.

- Почему другие игроки полагают, что если пять раз подряд выпадало красное, то выше вероятность, что дальше выпадет черное, а не красное? *Ответ:* на них действует эвристика представленности.
- Почему предпочтения иногда меняются с течением времени? *Ответ:* либо из-за гиперболического дисконтирования, либо из-за недолговечности породивших их эмоций.
- Почему многие из тех, кто верит в жизнь после смерти, тем не менее, хотят, чтобы смерть наступила как можно позже? *Ответ:* потому что их убеждение — не то, которое служит побуждением к действию, оно просто приносит им успокоение всякий раз, когда они задумываются о смерти.
- Почему люди неохотно признаются себе и другим в том, что они завистливы? *Ответ:* потому что они заботятся о своем имидже, а в большинстве обществ зависть находится на нижних ступенях нормативной иерархии мотиваций..
- Почему люди неохотно признаются себе и другим в том, что они чего-то не знают? *Ответ:* потому что человек — это животное, распознающее паттерны, а также потому что признание своего невежества в важных вопросах создает психический дискомфорт.
- Почему вера в то, что человеку предопределено попасть в рай или ад, приносила новообращенным кальвинистам XVI века большее успокоение, чем вера в то, что спасения можно достичь добрыми делами? *Ответ:* потому что магическое мышление заставляло их думать, что, присоединившись к церкви, они могут побудить Бога включить их в число избранных.
- Почему иногда верно, что «тот, кто обидел, не может простить»? *Ответ:* потому что горделивые индивиды так неохотно признают совершенный проступок, что придумывают причины, почему обиженный человек заслуживал и продолжает заслуживать такую судьбу.
- Почему в некоторых культурах стыд важнее вины? *Ответ:* потому что общество, в котором вина не концептуализирована, будет в меньшей степени демонстрировать поведенческие практики, связанные с виной.
- Почему победа французской сборной в 1998 году на чемпионате мира по футболу вызвала в стране такую радость и почему тот факт, что французская сборная не смогла выйти из группы на чемпионате-2002, поверг страну в такое уныние? *Ответ:* потому что удивление усиливает как позитивные, так и негативные эмоции.

- Почему женщины часто испытывают стыд, став жертвой изнасилования? *Ответ:* потому что жертвы, как правило, разделяют мнение «жертва сама виновата», которое вытекает из веры в то, что по сути своей этот мир — справедливое место.
- Почему унижительные ритуалы инициации вызывают большую, а не меньшую верность группе? *Ответ:* потому что мысль о том, что человек терпел сильную боль безо всякой цели, вызывает у него когнитивный диссонанс.

II. Действие

- Почему сегодня на бродвейских шоу аплодируют стоя чаще, чем 20 лет назад? *Ответ:* потому что публика хочет чувствовать, что получает за свои деньги некую ценность.
- Почему наказания могут привести к увеличению, а не к уменьшению количества преступлений, которые они призваны предотвратить? *Ответ:* потому что применение суровых наказаний вызывает ненависть, которая может пересилить страх.
- Почему люди так неохотно нарушают правила, которые сами себе установили, даже если их соблюдение не имеет смысла? *Ответ:* потому что они боятся, что единственное исключение может привести к распаду правила.
- Почему существует месть по модели «два oka за одно», а не по модели «око за око»? *Ответ:* из-за неприятия потерь.
- Почему долгосрочный доход по акциям гораздо выше, чем по облигациям (то есть, почему стоимость акций не поднимается, чтобы сравняться по доходности)? *Ответ:* из-за сочетания неприятия риска и «близорукости при принятии решения».
- Почему количество самоубийств снижается, когда опасные лекарства начинают продавать в блистерных упаковках, а не в бутылочках? *Ответ:* потому что многие желания настолько недолговечны, что к тому моменту, как будет вскрыта блистерная упаковка, суицидальный импульс пройдет.
- Почему ни один из 38 очевидцев не вызвал полицию, когда убивали Китти Дженовезе? *Ответ:* потому что каждый из них думал, что, так как никто больше не вмешивается, ситуация, возможно, не столь серьезная.
- Почему были люди, скрывавшие и спасавшие евреев от нацистов? *Ответ:* потому что их об этом попросили, а им было стыдно отказаться.

- Почему президент Ширак в 1997 году объявил досрочные выборы, в результате только потеряв большинство в парламенте? *Ответ:* потому что он не предвидел, что избиратели извлекут из досрочного объявления выборов информацию, которая заставит их голосовать против него.
- Почему при разводе родители часто соглашаются на совместную опеку над ребенком, даже в тех случаях, когда, обратившись в суд, один из них мог бы добиться исключительного права опеки? *Ответ:* из боязни риска.
- Почему вероятность эмиграции малоимущих ниже? *Ответ:* потому что они не могут позволить себе расходы на переезд и использовать себя в качестве залога в обеспечение займа.
- Почему некоторые люди выбирают рождественские вклады, по которым не выплачивается процент и с которых нельзя снять деньги до Рождества? *Ответ:* потому что они знают, что, если они положат деньги на обычный счет, гиперболическое дисконтирование может заставить их закрыть его досрочно.
- Почему люди занимаются такими изначально неприбыльными проектами, как создание авиалайнера «Конкорд»? *Ответ:* из-за гордыни или неприятия потерь.
- Почему в «правосудии переходного типа» (когда агентов авторитарного режима отдают под суд после перехода к демократии) те, кого судят сразу после такого перехода, получают более суровые приговоры, чем те, кого судят позднее? *Ответ:* из-за недолговечности мстительных настроений.
- Почему в шекспировской пьесе Гамлет откладывает месть до последнего акта? *Ответ:* потому что он страдает слабоволием и потому что драматический конфликт должен разрешаться не ранее конца пьесы.

III. Уроки естественных наук

- Почему родители чаще убивают приемных детей и пасынков, чем своих биологических детей? *Ответ:* потому что только последние несут в себе их гены.
- Почему так редок инцест между потомством одних родителей, учитывая возникающие соблазны и все имеющиеся для этого возможности? *Ответ:* потому что естественный отбор поощрял механизм, препятствующий возникновению сексуального влечения между разнополыми особями-одногодками, входящими в одну семью.

- Почему люди вкладывают деньги в проекты под руководством третьих лиц, даже если те свободно могут оставить всю прибыль себе? *Ответ:* потому что групповой отбор поощрял готовность к сотрудничеству.
- Почему люди мстят, хотя месть приносит им материальные убытки и не приносит материальной выгоды? *Ответ:* потому что групповой отбор поощряет стремление наказывать тех, кто отказывается от сотрудничества.
- Почему люди торопятся с выводами, выходя за рамки имеющихся доказательств? *Ответ:* потому что естественный отбор поощряет поиски паттернов.

III. ВЗАИМОДЕЙСТВИЕ

- Почему приверженцы социалистических партий иногда голосуют за коммунистов и тем самым препятствуют победе своей партии? *Ответ:* потому каждый из них оказывается жертвой «синдрома младшего брата», который мешает ему или ей понять, что другие могут сделать то же самое.
- Почему новые независимые государства принимают в качестве официального язык их бывших империалистических угнетателей? *Ответ:* потому что в стране с множеством местных наречий язык колониальной власти оказывается для всех вторым лучшим вариантом.
- Почему палатки с мороженым на пляже часто стоят рядом, тогда как и для покупателей, и для продавцов было бы удобнее, если бы они стояли на отдалении друг от друга? *Ответ:* потому что для каждого продавца индивидуально рациональнее сдвинуться к центру независимо от действий другого продавца.
- Почему люди голосуют на выборах, когда почти точно знают, что их голоса никак не повлияют на результаты? *Ответ:* потому что избиратель характеризуется магическим мышлением или действует согласно категорическому императиву.
- Почему в современных западных обществах экономически успешные люди обычно стройнее остальных? *Ответ:* потому что они верят (справедливо), что другие люди (ошибочно) убеждены в том, что нехватка самообладания в одной области может сказаться и на других областях.

- Почему люди воздерживаются от некоторых сделок, которые могли бы всем принести выгоду? Например, не просят у впереди стоящего продать свое место в очереди к автобусу? *Ответ:* потому что люди с большей неприязнью относятся к открытым проявлением экономического неравенства, чем к скрытым.
- Почему президент Никсон старался казаться Советскому Союзу человеком, склонным к нерациональному поведению? *Ответ:* потому что это поведение придает достоверность угрозам взаимного истребления, которым Советы в противном случае не поверили бы.
- Почему военачальники иногда сжигают за собой мосты (или свои собственные корабли)? *Ответ:* потому что ожидают, что их противник, зная, что они отрезали себе пути к отступлению, воздержится от слишком дорогостоящего сражения.
- Почему люди всегда придают такое большое значение, в сущности, незначительным вопросам этикета? *Ответ:* так как полагают, что нарушитель нормы не беспокоится о том, что о нем подумают.
- Почему пассажиры оставляют чаевые таксистам, а посетители ресторанов — официантам, даже при посещении иностранного города, в который они не предполагают вернуться? *Ответ:* потому что мысль, что другие плохо о них подумают, причиняет им боль.
- Почему фирмы инвестируют в большие запасы товара на складе, даже когда не предвидят остановки производства? *Ответ:* потому что полагают, что это отвлечет рабочих от забастовки и тем самым обеспечит бесперебойную работу.
- Почему в группе студентов каждый думает, что другие поняли сложный текст лучше, чем он? *Ответ:* потому что каждый из них стал жертвой «синдрома старшего брата», который подсказывает, что, если другие не обращаются за помощью, это происходит не из-за робости.
- Почему на многих политических собраниях проводятся поименные голосования? *Ответ:* потому что защитники популярных мер используют эту процедуру, чтобы помешать тем, кто в противном случае проголосовал бы против них.
- Почему взаимная поддержка (logrolling) чаще встречается в обычных законодательных органах, чем на конституционных собраниях? *Ответ:* потому что такая практика поддерживается поведением «услуга за услугу», требующим постоянного взаимодействия в течение длительных периодов времени.

Как качественная социальная наука отвечает критериям 1–4, которые я установил ранее? Будем двигаться в обратном порядке: классики не устарели. Я едва ли могу серьезно относиться к человеку, утверждающему, что сегодня принимать всерьез классические работы не стоит, потому что, если их выводы верны, они полностью инкорпорированы современной мыслью. Это не просто интерес антиквара. Хотя я не утверждаю, что диалог с великими мыслителями прошлого — единственный или лучший способ производства новых смыслов. Так, Томас Шеллинг, кажется, не опирается ни на каких предшественников, по крайней мере, какие-то очевидные связи отсутствуют. Возможно, Кеннет Эрроу заново открыл и обобщил гипотезу Кондорсе, но это не было сделано под влиянием работ последнего. Исследование Даниэля Канемана и Амоса Тверского, насколько мне известно, не опиралось на кого-либо из предшественников. Когда однажды мне представился случай указать Тверскому на то, что одно из его различий (между «эффектом вклада» и «эффектом контраста») предвосхитили Монтень и Юм, он ответил, что рад оказаться в такой хорошей компании. Поскольку именно эти ученые добились, вероятно, самых важных достижений в социальной науке за последние 50 лет, невозможно утверждать, что диалог с прошлым — единственный путь к новым прозрениям. Еще один пример: когда Токвиль взялся за создание «Старого порядка» он, по собственному признанию, не стал читать чужих работ по этой теме, ограничившись архивами и чтением произведений, современных исследуемым событиям. Он скорее предпочел бы, по его словам, переоткрыть найденные другими истины, чем испытывать на себе их чрезмерное влияние. Хотя есть свидетельства того, что в действительности он кратко ознакомился с написанным до него, его шедевр почти ничем не обязан другим авторам.

Диалог с прошлым может, тем не менее, оказаться крайне плодотворным, если только выявить позиции, которые вы хотите оспорить. Трудно себе представить, как такие противники Маркса, как Вебер или Шумпетер, могли бы написать свои произведения, не прочти они его внимательно. Прямое или положительное влияние, конечно, также очень распространено. На некоторые недавние теории эволюции собственности, по всей видимости, напрямую повлиял, а не просто явился предшественником Давид Юм. Работа Поля Вейна о психологии тирании в древности многим обязана гегелевскому анализу отношений раба и господина. Джордж Эйнсли, сделавший много для того, чтобы одно из основных прозрений Фрейда стало аналитически убедительным, смог бы прийти к своим идеям только благодаря более ранней, запутанной фрейдовской версии. Я подозреваю, что нужно получше покопаться в «Тактике законодательных собраний» Бентама. В этих случаях, как и во многих других, которые приходят мне на память, идеи, вдохновленные классика-

ми, в дальнейшем должны защищать себя сами. Удачное использование классиков само по себе не прибавляет авторитета.

Качественная социальная наука в большинстве приведенных мною примеров легко удовлетворяет критерию 3. То, что можно назвать «аналитическим поворотом» в социальной науке, на мой взгляд, основывается не на применении количественной методологии, а на почти маниакальной озабоченности ясностью и четкостью. (Эта озабоченность у некоторых аналитических философов может стать в полном смысле маниакальной.) Все чаще признается, как важно различать концептуальные и причинные связи между объектами исследования. «Контекст» все чаще прочитывается как «туман», а не как «питательная среда». По сравнению с более ранними исследованиями эссенциализма в науке становится все меньше; ученые реже задаются вопросом, чем «в действительности» являются демократия или социализм. Существует общее согласие с тем, что пока определения в какой-то степени ограничены узусом, они не стремятся ухватить стоящие за ними сущности. Определения являются случайными конвенциями, о которых нужно судить лишь с точки зрения того, в какой мере они помогают нам найти хорошие объяснения интересных явлений.

Кроме того, как я полагаю (или надеюсь), существует тенденция к отходу от того, что можно было бы назвать «недедуктивным абстрактным» мышлением. Чтобы считаться состоятельным, абстрактное рассуждение должно подчинить себя дисциплине либо дедуктивной логики, либо постоянной отсылки к фактам, которые подтверждают применимость абстрактных положений и *что-то означают*. В прошлом абстракции часто продолжали жить своей собственной жизнью. Значение понятий менялось по ходу спора, позволяя делать необоснованные заключения. Маркс, например, следующим образом «прослеживал» возникновение частной собственности из отчуждения труда: поскольку при капитализме продукт не «принадлежит» рабочему в том смысле, что работа не является осмысленной, он должен «принадлежать», быть собственностью кого-то другого, капиталиста. Благотворным следствием постоянной профессионализации социальной науки в большинстве западных образовательных учреждений стало то, что делать такого рода заявления становится все труднее. Стало меньше апелляций к аналогии, хотя они и не исчезли окончательно. В то время как идея «человеческого капитала» является ценным продолжением идеи капитала материального, этого не скажешь о «потребительском капитале» (Гэри Бекер), «культурном капитале» (Пьер Бурдьё) и «социальном капитале» (Роберт Патнэм). В лучшем случае это бесполезные и безвредные метафоры; в худшем они задают бесплодные пути исследования и подсказывают ложные каузальные гипотезы.

Труднее оценить, в какой степени качественная социальная наука удовлетворяет критерию 2. Среди историков (включая тех, кто рисует характерные портреты режима) мнения появляются и исчезают. Взять, например, идею, что террор был присущ Французской революции с самого начала, которая оспаривает утверждения о том, он стал результатом предотвратимого возвышения Робеспьера или в равной степени предотвратимого бегства короля, лишившего монархию легитимности и вызвавшего революционные войны. Возможно, согласно последней точке зрения, если бы Мирабо был жив, ничего подобного не случилось бы. В любой момент времени находятся французские историки, придерживающиеся любого из этих взглядов, но мнения большинства циклически повторяются. Взгляды, отброшенные сегодня, могут завтра оказаться преобладающими. Трудно найти какой-либо крупный исторический вопрос, который не порождал бы подобных колебаний. Можно предположить, что так будет всегда. Даже когда имеются в наличии и доступны для тщательного изучения все данные, как в случае истории Древнего мира, новые обобщения, произведенные неисторическими социальными науками, могут поддерживать одни толкования в ущерб другим.

В неисторических науках есть очевидный прогресс, но это не прогресс знания в буквальном смысле слова. Рассмотрим еще раз наше понимание того, почему у рациональных агентов может возникнуть желание отбросить некоторые варианты выбора, — например, сжечь за собой мосты. Понимание того, что это внешнее самоограничительное действие может быть *полностью* рациональным поведением, направленным против врага, а не рациональной защитой от своих собственных иррациональных наклонностей, является необратимым завоеванием в понимании социальных процессов. Другой вопрос, сколько реальных эпизодов сжигания мостов или кораблей она дает нам возможность объяснить. Командующий, подобно Кортесу, может сжечь корабли только за тем, чтобы предотвратить бегство. Представление о возможности кругового большинства тоже непреложно, но возможности ее реального применения, вероятно, ограничены. «Открытие» (как принято называть распространенные преаналитические идеи), что эмоции крайне недолговечны, не может быть повернуто вспять, но в любом конкретном случае могут найтись работающие в иных направлениях механизмы, которые могут компенсировать спонтанное ослабление эмоций. Даже если преобладающее объяснение данного события или эпизода отбрасывается, а потом возрождается заново, составные элементы или механизмы, действовавшие при отвержении и возрождении, остаются. Ассортимент или разнообразие инструментария не уменьшается.

Не очень хорошо обстоят дела у качественных социальных наук с соответствием критерию 1. Как уже отмечалось, среди определенного по-

колениа историков могут быть и на самом деле почти всегда имеются разногласия. Даже если отбросить политические взгляды, которые могут заставить марксистов и либералов по-разному оценивать террор, для научной полемики все равно останется достаточно места. Так, приписывать историческим акторам мотивации и убеждения — довольно искусное занятие. Поведение, которое один историк принимает за простодушное выражение ментальных состояний, другому может показаться стратегическим. Французский аристократ в изгнании во времена террора мог утверждать, что верит в скорую реставрацию монархии, и при этом арендовать квартиру на длительный срок. Современники и историки склонны меньше доверять его искренности, чем искренности того, кто снимает жилье на несколько коротких периодов; но даже этот человек в действительности может так поступать, чтобы создать видимость непоколебимого оптимизма. Его враги в любом случае уверены, что именно это он и делает. Если историку не повезет и он не найдет документ, которому нельзя приписать стратегические цели, например памятку, написанную историческим актором для себя самого, вопрос может решаться только с помощью суждения. И хотя хорошие историки являются таковыми благодаря хорошему умению судить, они могут ошибаться.

Внутри неисторических социальных наук могут также присутствовать существенные разногласия. В господствующей экономической науке часто утверждается, что многие открытия бихевиоральной экономики, дающие, как представляется, свидетельства иррациональности, связаны с тем, что они были сделаны в экспериментальных условиях. В силу самых разных причин то, что люди делают в искусственных лабораторных условиях, может не соответствовать их поведению в повседневной жизни. Экономисты-бихевиористы пытались организовать эксперименты так, чтобы ответить на это возражение: поднимали ставки, изолировали подопытных от социального давления других участников и самого экспериментатора, создавали условия для временного обучения. Кроме того, выдвигалось возражение против того, что участники механически переносили реакции, которые были адаптивными за стенами лаборатории, в обстановку эксперимента, где они становились плохо адаптирующимися, к примеру, используя стратегию «услуга за услугу» в однократных испытаниях, в которых она не имела смысла. Экономисты-бихевиористы отвечали, что в поведении подопытных нет никакого автоматизма, так как можно показать, что они способны прекрасно приспособливаться с учетом стимулов. Эти дебаты продолжаются по сей день, и неспециалисту трудно в них разобраться. Если я склонен встать на сторону экономистов-бихевиористов, вероятно, я делаю это на интеллектном основании: оппонент моего оппонента — мой друг.

КОЛИЧЕСТВЕННАЯ СОЦИАЛЬНАЯ НАУКА

Теперь я хочу обратиться к количественной социальной науке. Существует три ее разновидности: измерение, анализ данных и моделирование. Все три связаны, поскольку анализ данных часто требует измерений (для установления данных) и моделирования (чтобы указать, какие данные мы должны искать). Я сосредоточусь на моделировании, потому что слишком мало знаком с измерениями и анализом данных, но все же сначала скажу о них несколько слов.

Измерение, например, потребления на душу населения, безработицы, распространения коррупции или того, что общество думает о смертной казни, по сути своей очень сложная операция. Чтобы измерения потребления позволяли делать сравнения по времени и по разным регионам, они должны учитывать, что потребители с одинаковым доходом в разное время и в разных местах могут покупать разные товары, что доходы потребителей и имеющиеся в наличии товары также могут различаться. Безработицу трудно измерить, если существует масштабная теневая экономика, большое число студентов, заключенных или людей, прекративших поиски работы. Измерение коррупции обычно основывается на восприятии уровня коррупции, определяемом оценками экспертов и опросами общественного мнения. По очевидным причинам трудно получить независимые свидетельства, чтобы оценить надежность этих источников. Даже там, где они совпадают во мнениях, это может происходить по причине общей предвзятости. Замеры общественного мнения предполагают стабильность последнего. Однако хорошо известно, что незначительные изменения в формулировке вопросов могут привести к большим различиям в ответах. «Поддерживаете ли вы А?» побуждает совсем к другому ответу, чем «Поддерживаете ли вы А, если альтернативой является В?», даже если понятно, что В — единственная альтернатива А. Например, пусть А будет вступление некоей страны в Европейский союз, а В — невступление. Эти проблемы могут очень сильно повлиять на точность анализа данных.

Анализ данных (под которым я понимаю статистический анализ) в определенном смысле является сердцевинной современной социальной науки. Если мы заинтересованы в изучении крупномасштабных социальных явлений, мы знаем, что всегда будем иметь дело с населением, различающимся по ряду параметров, таких как здоровье, доход, семейное положение, предпочтения и место жительства. Любой детерминистский прогноз, вроде того, что столкнувшись с внешним шоком Х (например с увеличением предельной ставки налогообложения), все индивиды ответят поведением В (например, уменьшением своих трудовых ресурсов), обречено на провал. Вместо этого пытаются определить,

какова *вероятность* того, что индивид в данной категории отреагирует каким-то определенным образом, установив корреляцию между принадлежностью к данной категории и реакцией. Как правило, все это сугубо технические процедуры.

Такие опыты могут оказаться крайне опасными, то есть заключать в себе ряд ловушек, в которые могут попасть даже очень опытные ученые. В главах I и II я вкратце описал пять из них: добыча данных, подгонка кривой, произвольность в измерениях зависимых и независимых переменных, проблема различения корреляции и причинности, а также трудности с определением направления каузальности. Распространенная практика использования «запаздывающих» (lagged) значений переменных, устанавливающая корреляцию между одной переменной в момент t и значением другой в момент $t + n$, создает дополнительные возможности для искажений. Кроме того, проблемой может стать гетерогенность единиц анализа. При статистическом анализе причин войны неочевидно, что Первая и Вторая мировые войны не должны рассматриваться наравне с «футбольными войнами» между двумя латиноамериканскими странами. Кто-то, кроме того, может поставить под сомнение практику устранения аномальных случаев как «выпадающего значения» для подгонки данных. Еще одна проблема — предвзятость выборки, которая, например, может иметь место, если опросы общественного мнения проводятся по телефону, который исключает индивидов, не имеющих такового или по какой-то причине не отвечающих на звонки.

Неспециалисту вроде меня невозможно определить, как часто представители социальной науки попадают в эти и другие ловушки анализа данных. Золотым стандартом должно быть предсказание или «постсказание», которое предсказывает одну часть наблюдений, исходя из анализа другой. Выявление паттернов в данных может оказаться неубедительным из-за бесконечных возможностей для подтасовок и жульничества. Эти возможности тоже в принципе могут быть ограниченны, если ученые предварительно возьмут на себя обязательства, касающиеся гипотез и процедур анализа. Насколько мне известно, такой прием применяется редко.

Моделирование является дедуктивной процедурой, которая начинается с допущений и заканчивается предсказаниями. В социальных науках моделирование с точки зрения рационального выбора — самая распространенная, но далеко не единственная форма. В них все чаще выдвигаются эволюционные модели, которые не предполагают рациональных или даже интенциональных реакций на события. Ученые также предложили более механические модели поведения, то есть модели без ясных оснований или мотивировок. Английский физик Льюис Ричардсон, например, предложил модель гонки вооружений, основанную

на двух линейных дифференциальных уравнениях, за которыми не нет ясного психологического объяснения. Я ограничусь моделированием социальных явлений с точки зрения рационального выбора.

Моделирование с точки зрения рационального выбора очень распространено в экономике, в которой оно вышло сегодня далеко за рамки традиционных областей потребительского и производственного поведения. Например, существует экономика суицида, экономика посещения церкви и, шире, экономический анализ любой деятельности, которая так или иначе включает *выбор*. Существует тенденция к устранению концептуального барьера, который отделяет интенциональное действие от рационального: если есть возможность для выбора, делается допущение, что он будет сделан рационально. В политологии теория рационального выбора, без сомнения, является доминирующей, то есть приносит большой престиж и более высокие зарплаты, а также становится обязательной частью образования в элитных учебных заведениях (в США). Ее применение чаще сосредоточивается на политике (и иногда на партиях или даже государствах), чем на избирателях или гражданах. В социологии меньшинство, применяющее теорию рационального выбора к традиционным вопросам дисциплины, было вытеснено экономистами, которые лучше владеют этим инструментарием (изучение доверия является важным исключением). В социальной антропологии, которая (в США) двигалась в направлении «мягкой» социальной науки, теория рационального выбора не утвердилась.

Имеем ли мы дело с наукой? Позвольте сначала рассмотреть критерий 1. Мы можем считать само собой разумеющимся, что моделирование корректно в том смысле, что выводы следуют из предпосылок. Истинны ли выводы? Прежде чем обратиться к этому вопросу, я хочу заметить, что неясно, всегда ли те, кто их делает, *хотят*, чтобы они были истинными, то есть соответствовали реальному миру. Скорее иногда они представляют собой своего рода социальный вымысел — анализ действия и взаимодействия идеально рациональных агентов, которые никогда не существовали и не будут существовать. Анализ постоянно совершенствующихся форм стратегического равновесия, к примеру, едва ли мотивирован желанием объяснить или предсказать поведение реальных индивидов. Скорее он руководствуется эстетическими мотивами. Двое из наиболее опытных теоретиков равновесия, Рейнхард Зелтен (Reinhard Selten) и Ариэль Рубинштейн (Ariel Rubinstein), дали ясно понять: они не считают, что их модели могут что-то сказать о реальном мире. Обращаясь к вопросам реальности, они пользуются некоторой разновидностью бихевиоральной экономики или ограниченной рациональности. Другой пример: теория социального выбора — аксиоматическое изучение механизмов голосования в какой-то момент стало настолько математически изощренным

и до такой степени утратило связь со всяким изучением реальной политики, что один из самых авторитетных журналов в области экономики «Эконометрика» ввел мораторий на публикацию статей по этой теме.

Интересный вопрос для психологии и социологии науки: сколько людей являются *тайными адептами* такой «научно-экономической фантастики», скрывая от себя или от других тот факт, что заняты именно этим. Придумывание хитроумных математических моделей — весьма доходное занятие, но за исключением людей, подобных Зелтену и Рубинштейну, оплачиваемое, только если такую деятельность можно посчитать еще и релевантной; отсюда возникают стимулы для обмана и самообмана. В такой постановке вопрос может показаться неуместным в академическом дискурсе, однако я не понимаю почему. Иногда наступает момент, когда нужно отбросить академический политес — к такому же утверждению приходит Джон Стюарт Милль в своем эссе «О свободе», которое с момента своего выхода в свет считалось библией в вопросе интеллектуальной свободы. Я уже отказался от любых претензий на вежливость по отношению к «мягкой» социальной науке и не вижу причины как-то иначе обращаться с *жестким обскурантизмом*.

У меня нет прямых свидетельств обмана или самообмана, так что мне придется пойти другим путем. Я попытаюсь показать, что то, чем занимаются поклонники рационального выбора, часто настолько далеко от действительности, что их заявления о том, что они имеют дело с реальным миром, невозможно принимать всерьез. Рискую повторить то, о чем я говорил в предыдущих главах, я хочу резюмировать свои доводы в 10 пунктах.

1. Многие модели приписывают агентам мотивации, которых, как можно продемонстрировать, у них нет. Одним из примеров может служить общераспространенное допущение экспоненциального дисконтирования во времени.
2. Некоторые модели приписывают агентам мотивации, относительно которых не предложено никаких доказательств, — например, частная форма функции полезности. Конечно, можно использовать функцию для доказательства теорем вероятности («рациональный потребитель с убывающей предельной полезностью может сделать X»), но только не в подтверждение фактов, связанных с реальным миром. Во многих случаях необходимо также доказывать, а не просто принимать в качестве допущения эгоистичность агентов.
3. Большинство моделей игнорирует достижения в понимании выбора, связанные с теорией перспектив. Поскольку

последняя формулируется в категориях приобретений или потерь по отношению к заданной базовой линии, она не анализирует поведение с точки зрения максимизации объективной функции, определенной на множестве исходов. Формат максимизации может быть действительно необходим для математических выкладок, но к реальному поведению он не имеет никакого отношения.

4. Некоторые модели приписывают агентам способности, которыми те могут обладать, не используя их. Например, вовсе не обязательно, что они будут спонтанно применять обратную индукцию, требующуюся во многих важных случаях.

5. Многие модели приписывают агентам способности, которыми те, как можно показать, не обладают. Этот пункт настолько тривиален, что едва не обескураживает: как экономист может предполагать, что агент способен произвести расчеты, которыми сам экономист заполняет многие страницы глубоко технических приложений? Велико искушение воскликнуть: «Эй! Опустись на землю!»

6. Многие модели приписывают намерения на основании объективного интереса агента в отдельной ситуации, не задаваясь вопросом о том, не могут ли здесь действовать другие мотивации: эмоции или социальные нормы. Попытка объяснить поведение в беспокойных или конфликтных ситуациях, таких как революции, приписыванием агентам рационального долгосрочно-эгоистического мотива (редуцированного к настоящей величине экспоненциальным дисконтированием), выглядит почти смехотворно неадекватной. Другие модели приписывают намерения на основании реальных исходов, пренебрегая тем самым возможностью ошибки в расчетах и просчитанными рисками.

7. Многие модели не учитывают тумана неопределенности, который окружает большинство важных решений, особенно когда значимые последствия наступят в далеком будущем. Попытки преодолеть эту проблему с помощью приписывания исходам субъективных вероятностей зачастую произвольны, как, например, когда ученые апеллируют к принципу недостаточного основания для установления равномерного разброса вероятностей.

8. Некоторые модели рассматривают коллективы (классы или государства), как если бы они были отдельными агентами, оставляя без внимания (в случае классов) проблему «без-

билетника» или (в случае государств) структуру коллективного принятия решений.

9. Некоторые модели предполагают, что отклонения от рациональности либо (1) временные, либо (2) взаимоустранимы при накоплении. Что касается первых, временные отклонения могут влечь за собой длительные результаты в ситуации со структурой «ловушки для дураков». Что касается вторых, бихевиоральная экономика показала, что, поскольку многие отступления носят систематический, а не разовый характер, нет оснований ожидать, что при накоплении они исчезнут. Примером может служить «головоломка премии за приобретение акций».

10. Многие модели довольствуются тем, что объясняют поведение, предполагая, что в игре есть равновесие, при этом не делают следующий шаг, чтобы показать, почему в ситуациях с множеством точек равновесия именно эта конкретная реализуется. Более того, они редко задумываются о том, не изменяется ли мир так быстро, что равновесие не успевает установиться.

Я упоминал некоторые типичные ответы на эти возражения. Один — ссылка на герменевтическую проблему установления мотиваций некруговым способом. Я соглашусь с тем, что это серьезная трудность, но она не является непреодолимой, и даже если бы была таковой, это не оправдывало бы необоснованное приписывание мотиваций. Другая предлагает заменить действительную рациональность (реальных, живых агентов) на «как бы рациональность» (as-if rationality). Я утверждал, что эта стратегия может работать, только если удовлетворены оба условия. С одной стороны, можно указать на механизм, способный имитировать рациональность *вплоть до* приложений с математическими расчетами. Проблема в том, что никто не придумал такого механизма. Апелляции к отбору — не более чем размахивание руками, учитывая огромное расхождение между грубыми и несовершенными механизмами социального отбора и сверхизолированными процессами мышления, приписываемыми агентам. С другой стороны, можно указать на предсказания, которые настолько точны, что мы вынуждены принять теорию, хотя и не понимаем в точности, как она работает. Проблема в том, что социальные науки не дают предсказаний, сколь бы то ни было близких к этому статусу.

На самом деле, если мы внимательнее посмотрим на удовлетворение критерию I, мы заметим, что различные теории с различными предсказаниями и объяснениями сосуществуют во времени. В этом отношении особенно выделяется макроэкономика. Я больше не отслеживаю теку-

щую ситуацию в кейнсианстве, монетаризме и рациональных ожиданиях, но, если кто-то из них обошел других и выдает постоянный поток в высшей степени точных прогнозов, полагаю, я бы это заметил. Является ли потребность в независимом центральном банке доказанной истиной или просто капризом? Уже несколько десятилетий экономисты предлагали радикально противоположные объяснения высокого уровня безработицы в Европе, так и не придя к согласию. Действительно ли люди начинают меньше работать, когда их облаживают более высокими налогами? Мнения экспертов расходятся. Это не вопросы, которые находятся на передовых рубежах исследования, где всегда будут расхождения, но давние проблемы, лежащие в самом центре дисциплины.

Критерий 2 также не удастся удовлетворить, на что указывает приставка «нео», например, в неокейнсианстве или неофункционализме. Критерии 3 и 4 удовлетворяются легко, на самом деле, даже слишком хорошо. Негативной стороной профессионализации американской социальной науки, на которую я указывал ранее, является калечащая узость и самоубийственная озабоченность местом собственного факультета в том или ином рейтинге. Я хочу выдвинуть гипотезу, что на самом деле высокое место в рейтинге — это, по сути, побочный продукт стремления к знаниям, а не стремления к рейтингам. (Как мы видели, Пруст делал подобное замечание о музыкантах.) Более того, так как диалог с прошлым отдан не слишком престижной подотрасли истории экономической мысли или чуть более престижной истории мысли политической, названная узость лишь усугубляется. Но хотя патологии жесткого обскурантизма предпочтительнее недостатков его мягкого варианта, выбирать не приходится.

БУДУЩЕЕ СОЦИАЛЬНОЙ НАУКИ

Может ли ситуация измениться к лучшему? Могут ли устремления социальных наук — дача предсказаний, определенность и точность — быть удовлетворены в какой-то момент в будущем? История науки учит нас с осторожностью высказываться о том, что может и чего не может отдельная дисциплина. «Они все смеялись» над Декартом, когда тот заявил, что животные — машины, но кто будет смеяться последним? Любая попытка решить этот вопрос должна быть осторожной и предположительной.

Разумеется, инкорпорация открытий нейронаук даст более твердые основания психологии и сможет разрешить некоторые существующие споры. Так, сканирование мозга подтверждает скорее квазигиперболическое, а не гиперболическое дисконтирование. Таким же образом можно определить различие между гневом и негодованием, если удастся показать, что эти эмоции активируют разные участки мозга.

Нейрофизиология зависимости, конечно, будет и дальше помогать нам понять это загадочное, саморазрушительное поведение. И все же заметьте, что все эти заявления касаются *побуждения* к действию. Поползновений в направлении феномена «заполнения» также предостаточно. В то же время нейронаука пропозициональных *убеждений* (и взаимодействий мотивированных убеждениями) не существует и в ближайшем будущем не появится. Рассмотрим тривиальный случай пересмотра убеждений. Я думал, что завтра пойдет дождь, но, когда небо прояснилось, я скорректировал свой прогноз и в результате изменил свои планы. Я не думаю, что только нехватка моего воображения (или нехватка научной компетенции) заставляет меня сомневаться, что вскоре мы сможем выявить нейрофизиологию даже этого простого процесса, в то время как многие механизмы формирования убеждений, конечно, еще сложнее. Если я прав, тогда половина модели поведения, основанной на убеждении-желании, еще надолго останется недоступной для нейронауки. То же самое с еще большей уверенностью можно сказать, если выйти за пределы этой модели и задаться вопросом о «нейронах вдохновения», порождающих открытия или произведения искусства.

Есть два вида причин, по которым социальные науки не могут давать предсказаний или объяснений в строгом смысле. Одна заключается в том, что даже для данных убеждений и предпочтений действие может остаться в какой-то степени неопределенным (то есть непредсказуемым). При принятии решений в ситуации неопределенности или ситуации с высокой степенью сложности поведение может быть вызвано тем, что Кейнс назвал «жизнерадостностью», а не некоторыми особенностями ситуации, на которые мы реагируем определенным образом. Люди определенно будут вести себя в этих случаях согласно практическим правилам: проблема в том, что таких правил слишком много. Например, это могут быть конкурирующие между собой точки координации: *делай как раньше*, или *делай, как поступают соседи*. Сказать что люди скорее «удовлетворяют», а не «максимизируют», — значит не сказать почти ничего, пока мы не определим то, что конституирует удовлетворительный уровень.

Другая причина — наше плохое понимание механизмов формирования предпочтений. Индивиды подвержены конкурирующим между собой склонностям, относительную силу которых в любой конкретной ситуации часто трудно оценить. Если я стану угрожать вам, вы испугаетесь или разозлитесь? Если возникнут оба эффекта, какой из них будет преобладать? Если я решу отдохнуть от моей бурной профессиональной жизни, будет ли мой досуг таким же бурным или же, наоборот, крайне расслабленным? Если моя страна из диктатуры превратится в демократию, падение политических авторитетов заставит меня также отвергнуть

религиозные авторитеты или, наоборот, подтолкнет в их объятия? Я настаивал на нашей частой неспособности дать предварительные ответы на эти вопросы, хотя постфактум у нас может появиться возможность выявить преобладающий механизм.

Можно ли снизить степень неопределенности, выявив либо *условия возникновения*, либо *вероятности возникновения*? Я рассматривал первую возможность в главе II и пришел к несколько скептической оценке, так что позвольте перейти ко второй возможности. Можно ли, например, вычислить распределение среди населения уровней удовлетворенности, которые, по крайней мере, помогут нам предсказать общее поведение? В конце концов, в этом и состоит основная задача социальных наук. Точно так же, если мы предположим, что какая-то доля населения под действием угроз скорее испугается, чем разозлится, или что каждый из нас может реагировать и тем, и другим образом, есть ли возможность квантифицировать эти вероятности?

Даже если допустить существование такой возможности, я по-прежнему сомневаюсь, что мы в состоянии достичь больших объяснительных возможностей в силу колоссального значения *контекста*. Ранее я говорил, что прогресс в науке часто основывается на постоянном абстрагировании от контекста. Это в равной мере распространяется и на социальные науки, в том смысле, что мы можем идентифицировать склонности или механизмы только путем отбрасывания многих обстоятельств ситуации или создания обстановки, в которой они отсутствуют. Например, есть сложное интуитивное понятие конформизма. Оно включает склонность полагаться на других как на источник информации, желание быть как другие, стремление не высовываться и нежелание, чтобы другие думали о нас плохо. Если в специфических целях исследования мы хотим определить как конформизм (скажем) вторую черту, мы должны изолировать ее от трех других, создав экспериментальные условия (или найдя реальные), в которой они отсутствуют. Возможно, мы выясним, что одни люди действительно захотят быть как все, тогда как другие захотят выделиться. Возможно, мы даже могли бы определить пропорции или вероятности. Однако в реальных ситуациях присутствуют и другие измерения конформизма, которые могут произвести обратный эффект. Знание того, как люди распределяются вокруг одного из четырех измерений конформизма, тоже может оказаться не слишком полезным для прогнозирования конформистского поведения. Если отбросить шутки в сторону, анализ вполне осуществим. Гораздо сложнее, если вообще возможно, осуществить синтез, сведение воедино.

Легко представить себе ученого, занятого доказательством того, что социальная наука может помочь преодолеть это затруднение, но я не стану больше распространяться на эту тему. Некоторых из нас поражает

и захватывает сложность и изменчивость человеческого поведения. Другие испытывают интуитивную веру в стоящие за всем закономерности, которые, будучи выявленными, дадут возможность социальным наукам встать на равных с естественными — будь то физика, химия, геология или метеорология. Будущее покажет, я же пытался прояснить, на что и почему делаю ставку.