



ЮН

ЭЛЬСТЕР

**КИСЛЫЙ  
ВИНОГРАД**

ИССЛЕДОВАНИЕ  
ПРОВАЛОВ  
РАЦИОНАЛЬ-  
НОСТИ

**Юн Эльстер**  
**Кислый виноград.**  
**Исследование провалов**  
**рациональности**

*Текст предоставлен правообладателем*

*[http://www.litres.ru/pages/biblio\\_book/?art=40652708](http://www.litres.ru/pages/biblio_book/?art=40652708)*

*Кислый виноград. Исследование провалов рациональности [Текст] / пер. с англ. И. Кушнareвой; науч. ред. перевода А. Морозов: Изд-во Института Гайдара; Москва; 2019  
ISBN 978-5-93255-522-4*

### **Аннотация**

Черпая вдохновение в философии, политической и социальной теории, теории решений, экономике, психологии, истории и литературе, классическая книга Юна Эльстера «Кислый виноград» продолжает и дополняет размышления его прославленной ранней работы «Улисс и сирены». Эльстер начинает с анализа обозначений рациональности, чтобы затем взяться за понятия иррационального поведения, желаний и убеждений при помощи крайне изощренных аргументов, подрывающих ортодоксальные теории рационального выбора. Изданный в новом серийном оформлении и со специально написанным по этому случаю предисловием Ричарда Холтона, в котором раскрывается важность этой книги для философских

исследований, «Кислый виноград» был возрожден для нового поколения читателей.

# Содержание

Предисловие к изданию 2016 года	5
Предисловие и благодарности	8
I	11
I.1. Введение	11
I.2. Индивидуальная рациональность: слабая теория	15
I.3. Индивидуальная рациональность: широкая теория	40
Конец ознакомительного фрагмента.	48

**Юн Эльстер**  
**Кислый виноград.**  
**Исследование провалов**  
**рациональности**  
*Под научной редакцией*  
*Артема Морозова*

**Предисловие к изданию 2016 года**

В привычной картине вы действуете рационально, если цель действий – удовлетворить ваши желания. Но порой действия и желания мешают друг другу, размывая эту картину. Именно таким случаям посвящен «Кислый виноград».

Одна из проблем возникает, когда желаемая цель представляет собой нечто, к чему нельзя стремиться без того, чтобы не обречь себя на провал. Стендаль хотел быть рациональным, а дадаисты – спонтанными, но ни того ни другого нельзя добиться, если прилагать к этому усилия. В данном случае проблема концептуальная. В других случаях она каузальная. Мы могли бы быть созданиями, которые засыпают,

пытаясь заснуть, но для большинства из нас, особенно для страдающих бессонницей, попытки еще больше осложняют ситуацию. Остальные ситуации находятся в некоторой сложной промежуточной зоне. Можно ли пытаться расслабиться? Или пытаться быть скромным?

Нашу проблему, хотя и не совсем точно, иллюстрирует случай, вынесенный в заглавие книги. Лиса умирает от голода, но не может дотянуться до винограда, висящего слишком высоко, и потому объявляет его кислым. Виноград не кислый, так что лиса ошибается. Вместо этого она могла бы перестать хотеть виноград или, что еще лучше, сформировать позитивное желание его не есть. Тогда она бы не совершила ошибки, но ее желание было бы удовлетворено. Мы получили бы то, что Эльстер называет – и это уже общепринятый термин – *адаптивными предпочтениями*: желание, адаптировавшееся к тому, что доступно.

Как нам понимать такие желания? Если самое главное – удовлетворение желаний, тогда изменение желания с тем, чтобы оно было сообразно миру, ничуть не хуже изменения мира так, чтобы он был сообразен желанию. Возможно, порой так оно и есть. Стратегии стоиков или буддистов, включающие в себя метажадение иметь желания, сообразные миру, как кажется, включают в себя также совершенно рациональное планирование характера. Но лиса – не буддист. Она просто испытывает воздействие слепого каузального механизма, влечения, как называет его Эльстер, которое «дей-

ствует у нее за спиной» и подспудно сообразует ее желания с миром. В таких случаях, как представляется, рациональность не действует. Более того, настаивает Эльстер, утилитаристская теория, согласно которой правильно удовлетворять желания агентов, терпит здесь двойной провал. Во-первых, она не говорит, что мы должны делать, поскольку, если в конечном счете у агентов бесконечно адаптивные предпочтения, их будут удовлетворять любые наши действия. И, во-вторых, интуиция подсказывает, что неправильно стремиться удовлетворить адаптивные предпочтения; лисе было бы лучше, если бы мы дали ей виноград, даже если она очень хорошо умеет приспосабливаться.

Все это изложение сопровождается большим количеством примеров. В последние годы у философов действия стало принято разбираться в психологии. Эльстер обратился к ней задолго до них, но его книга в равной мере опирается и на историю, экономику и литературу. Несмотря на влияние, которое она оказала, «Кислый виноград» – основополагающая книга, которую еще только предстоит изучить.

*Ричард Холтон*

# Предисловие и благодарности

Действие – это результат выбора в рамках ограничений. Выбор, согласно общепринятому взгляду, воплощает в себе элемент свободы, а ограничения – элемент необходимости. Однако в нестандартных случаях эти тождества не выполняются. Заглавие моей более ранней книги, посвященной рациональному и нерациональному поведению, «Улисс и сирены» – напоминание о том, что порой люди свободны выбирать свои собственные ограничения. В «Кислом винограде» я, наоборот, размышляю над идеей, что предпочтения, лежащие в основе выбора, могут быть сформированы ограничениями. Рассматриваемые вместе, эти два нестандартных явления достаточно важны для указания на то, что общепринятая теория нуждается в фундаментальном пересмотре.

Таким образом, эта книга дополняет мои ранние работы. До некоторой степени она также исправляет то, что мне сейчас видится слишком восторженным применением идеи выбора людьми собственного характера. Глава, посвященная состояниям, являющимся по сути своей побочными продуктами, показывает, что имеются пределы для достижения цели методами планирования характера. От представления о том, что можно быть хозяином своей душе, веет гордыней, равно как и представление о том, что все следствия действия могут быть объяснены им же, пахнет интеллектуальной

ошибкой.

Настоящая книга также представляет собой попытку объяснить некоторые течения в сложных понятиях рациональности, иррациональности и оптимальности. Некоторые из вопросов, поднятых в данной связи, более подробно обсуждаются в моей книге «Объяснение социальных изменений». В особенности это касается анализа функционального объяснения.

Я бы хотел прежде всего поблагодарить Дж. А. Коэна, который сделал обширные замечания к черновикам глав II, III и IV. Без его способности пробуждать меня от врожденной интеллектуальной лени мой уровень аргументации был бы гораздо ниже. Далее я хочу поблагодарить членов семинара по рациональности, проводившегося под эгидой «Дома наук о человеке», за полезные обсуждения и неизменное вдохновение. В частности, я выражаю благодарность Брайену Бэрри, Дональду Дэвидсону, Дагфинну Феллесдалю, Роберту Гудину, Сержу Кольму, Амели Рорти, Амосу Тверски и Бернару Уильямсу. Наконец, хочу отметить то, что сразу бросается в глаза любому читателю: мой огромный интеллектуальный долг перед выдающейся работой Поля Вена «Хлеб и зрелища». Кроме того, я хочу выразить признательность следующим людям, помогавшим мне с отдельными главами. Мои представления о коллективной рациональности из главы I сложились в ходе многочисленных обсуждений с Анундом

Хилландом, Руне Слагстадом и другими участниками проекта «Демократия и социальное планирование», организованного Норвежским исследовательским советом по гуманитарным наукам. Ранний, гораздо более короткий и несколько запутанный, вариант главы II впервые был опубликован в «Информации о социальных науках» (*Social Sciences Information*, 1981). Я благодарен Елене Алмази за ее редакторскую помощь и Вольфу Лепени за полезные замечания. Слегка отличающийся вариант главы III был опубликован в: A. Sen and B. Williams (eds.), *Utilitarianism and Beyond* (Cambridge: Cambridge University Press, 1982). Я получил ценные замечания к черновику этого варианта от редакторов данного сборника, а также от Хермана ван Гунстерена, Мартина Холлиса, Джона Ремера и Артура Стинчкомба. Глава IV была опубликована, тоже в несколько иной форме, в: M. Hollis and S. Lukes (eds.), *Rationality and Relativism* (Oxford: Blackwell, 1982). Я благодарен Мартину Холлису за его редакторские советы.

Университет Осло, Норвежский исследовательский совет по гуманитарным наукам, «Дом наук о человеке» и Колледж Всех Душ в Оксфорде также внесли материальный вклад в написание этой книги.

# I

## Рациональность

### I.1. Введение

Основная тема данной работы – рациональность. Однако ее обсуждение не имеет смысла без предварительного анализа самого понятия рациональности. А оно оказывается чересчур богатым. Существует бесчисленное множество вещей, о которых говорят, что они рациональны или иррациональны: убеждения, предпочтения, выборы или решения, действия, поведенческие модели, люди, даже коллективные образования и институты. Кроме того, коннотации термина «рациональный» варьируются от формальных понятий эффективности и непротиворечивости до содержательных понятий автономии и самоопределения. А где-то на заднем плане этого понятия маячит замечательная пара *Verstand* (рассудок) и *Vernunft* (разум), будь то в кантовском или в гегелевском смысле.

Я начну с рациональности как формальной особенности индивидуальных действий (1.2). В результате мы придем к тому, что, вслед за аналогичной терминологией Ролза<sup>1</sup>, я бу-

---

<sup>1</sup> Ролзом приводится «слабая теория блага для объяснения рационального

ду называть слабой теорией рациональности. Слабая она постольку, поскольку не объясняет убеждений и желаний, образующих причины действий, рациональность которых мы оцениваем, но только оговаривает, что они не являются логически противоречивыми (inconsistent). Собственно, под рациональностью в слабом смысле имеется в виду непротиворечивость: отсутствие противоречий в системе убеждений, непротиворечивость системы желаний и отсутствие противоречий между убеждениями и желаниями, с одной стороны, и действиями, причинами которых они становятся, – с другой.

Широкая теория индивидуальной рациональности идет дальше формальных требований (1.3). Рациональность с ее точки зрения предполагает нечто большее, чем непротиворечивые действия в соответствии с непротиворечивыми убеждениями и желаниями: мы также требуем, чтобы сами убеждения и желания были рациональными в более существенном смысле, что нетрудно показать на примере убеждений. Рациональные убеждения – это, по сути дела, такие убеждения, которые основываются на имеющихся фактах: они тесно связаны с понятием *суждения*. Труднее поддается определению понятие рационального по сути желания. Один из способов решения проблемы – предположить, что *авто-*

---

предпочтения первичных благ», и в то же время он признает, что более широкая теория потребует для объяснения «морального достоинства личности» (Rawls 1971: 396; Ролз 1995: 351).

*номия* является для желания тем, чем суждение является для убеждения, и именно данной посылкой я буду в дальнейшем руководствоваться.

Понятие рациональности может быть также расширено в ином направлении – от индивидуального к коллективному. И снова я начну с более формальных соображений (1.4). На этом уровне рациональность может быть закреплена либо за любым процессом коллективного принятия решений (как в теории коллективного выбора), либо за агрегированным исходом индивидуальных решений. В обоих случаях индивидуальные желания и предпочтения воспринимаются как *данность*, а рациональность определяется преимущественно как отношение между предпочтениями и социальными исходами. В более широкой теории коллективной рациональности (1.5) также придется рассматривать способность общественной системы или коллективного механизма принятия решений приводить индивидуальные предпочтения в соответствие с широким понятием индивидуальной рациональности. Коллективное рациональное установление в этом смысле поощряет автономные желания или же способно отфильтровывать неавтономные.

В настоящей главе я рассматриваю рациональность, в последующих главах – проявления иррациональности. На отношения между ними можно, среди прочего, посмотреть следующим образом. Рациональность указывает агенту, что ему следует делать; если он ведет себя иначе, он иррациона-

лен. Я буду оспаривать этот взгляд. Есть много случаев, в которых рациональность – в слабом или в широком смысле – может лишь исключить некоторые альтернативы, не давая при этом никаких указаний, какой из оставшихся в итоге вариантов следует выбрать. Если мы хотим *объяснить* поведение в подобных случаях, помимо допущения о рациональности следует также учитывать каузальные соображения. Далее я, собственно, утверждаю, что, если мы хотим получить широкое понятие рациональности, такого рода подход становится правилом, а не исключением.

## 1.2. Индивидуальная рациональность: слабая теория

В рамках подхода, намеченного Дональдом Дэвидсоном<sup>2</sup>, рациональным является действие, которое соотносится с убеждениями и желаниями агента (их я собирательно называю *его причинами*). Во-первых, мы должны настаивать на том, чтобы эти причины были причинами для действия; во-вторых – чтобы они реально вызывали действие, причинами которого являются; и в-третьих – чтобы они вызывали действие «должным образом». Подспудно здесь присутствует требование непротиворечивости самих желаний и убеждений. В дальнейшем изложении мы сосредоточимся прежде всего на непротиворечивости, но сначала я хочу немного остановиться на трех условиях, которые вошли в определение рационального действия.

К первому условию можно подойти с двух сторон. Либо можно сказать, что причины являются причинами для действия в том случае, если, учитывая убеждения агента, данное действие является *наилучшим* способом осуществления его желания. Либо (в более слабом смысле) можно сказать, что причины являются причинами для действия, если это *единственный* способ осуществления желания с учетом убежде-

---

<sup>2</sup> См., в частности, работы в сборнике: Davidson 1980.

ний. Данное различие связано с проблемой, поднятой в последнем абзаце раздела I.1, хотя и не совпадает с ней, поскольку вопрос о единственности (существует ли только один рациональный образ действия?) следует отделять от вопроса об оптимальности (является ли рациональный образ действия *наилучшим*?). Вполне может иметься несколько одинаково и максимально хороших альтернатив. Я буду обсуждать данные проблемы ниже. Пока я лишь хочу отметить, насколько слабой оказывается теория рациональности, с которой мы здесь имеем дело. Если агент испытывает навязчивое желание убить другого человека и полагает, что наилучший (или единственный) способ это сделать – воткнуть булавку в куклу, которая представляет последнего, то он будет действовать совершенно рационально, когда ее проткнет. Вы вправе усомниться в рациональности такого желания и такого убеждения.

Второе условие в данном определении требуется для исключения того, что можно назвать «совпадениями первого класса», при которых человек имеет причины для того, чтобы действовать так, как он действует, однако в реальности его действия вызваны не этими причинами, а чем-то другим. Можно случайно сделать то, что у вас есть причины сделать. Кроме того, навязчивое поведение порой может вполне соответствовать обстоятельствам.

Третье условие необходимо для исключения «совпадения второго класса», при котором причины реально вызывают

то действие, причинами которого являются, но «недолжным образом». То, что причины могут вызывать действие «недолжным образом», можно увидеть на примере случаев, в которых они вызывают действие, не являясь при этом его причинами. Дэвидсон, в частности, полагает, что именно так может объясняться слабость воли<sup>3</sup>. Однако данная ситуация сложнее, поскольку действие, которое причины вызывают «недолжным образом», имеет в качестве причин именно их. Чтобы понять, как это возможно, воспользуемся идеей Дэвидсона о нестандартных каузальных цепочках. Возьмем следующий пример из реального мира: «Человек может покушаться на убийство другого человека, выстрелив в него. Предположим, убийца промазал, но выстрел вспугнул стадо диких кабанов, которые затоптали жертву»<sup>4</sup>. В этом случае мы вряд ли скажем, что человек убил жертву умышленно, поскольку имеем дело с неправильной казуальной цепочкой. Что касается случая ментальной каузации, которая нас здесь интересует:

Альпинист может хотеть избавиться от лишнего груза и опасности, связанной с тем, что он удерживает на веревке другого человека, и может знать, что если он ослабит хватку, то действительно избавится от груза и опасности. Альпинист может так разволноваться из-за убежденности и желания, что действительно ослабит

---

<sup>3</sup> Davidson 1980: ch. 2.

<sup>4</sup> Davidson 1980: 78.

хватку, но может так случиться, что он не принимал решения делать это и не сделал это умышленно<sup>5</sup>.

Убеждения и желания могут быть причинами для действия только в том случае, если они непротиворечивы. В них не должно быть логических, концептуальных или прагматических противоречий. Сначала я буду обсуждать критерий непротиворечивости применительно к убеждениям, а затем, уже несколько подробнее, – применительно к желаниям.

Оценить непротиворечивость убеждений нетрудно, по крайней мере на довольно поверхностном уровне, на котором мы делаем допущение о том, что убеждения уже были выявлены. На более глубоком уровне мы должны принять утверждение Дэвидсона о том, что выявление убеждений человека и оценка их непротиворечивости неразделимы. Процесс вменения убеждений должен следовать допущению о том, что они более или менее непротиворечивы<sup>6</sup>. Но как только мы установим точку отсчета или основу для общей непротиворечивости, можно поставить вопрос о локальной противоречивости в убеждениях. Дальнейшее верно только при этом условии.

Мы можем рассматривать убеждения либо как субъективные оценки вероятности, либо как нечто в своем роде, *sui generis*. При первой интерпретации непротиворечивость будет означать просто соответствие законам вероятности: ча-

---

<sup>5</sup> Davidson 1980: 79.

<sup>6</sup> Davidson 1980: ch. 12 and *passim*.

стичные вероятности исключительных и исчерпывающих событий в сумме дают единицу, вероятность сочетания любых двух из них равна нулю и т. д. Аналогично вероятности составных событий должны иметь такое отношение к вероятностям элементарных событий, при котором, например, конъюнкция независимых событий имела бы вероятность, равную продукту составляющих событий.

Если брать убеждения *sui generis*, то очевидным критерием непротиворечивости может показаться следующий: набор убеждений будет непротиворечивым, если имеется возможный мир, в котором все эти убеждения являются истинными, то есть если из них нельзя вывести противоречие. Яакко Хинтикка, однако, показал, что этого недостаточно<sup>7</sup>. Его критерий: убеждения непротиворечивы, если существует возможный мир, в котором они все *истинны и в них верят*. Потребность в последнем условии возникает в случаях убеждений высшего порядка, то есть убеждений об убеждениях. Например, одно время говорили, что у Нильса Бора над дверью висела подкова. Когда его спросили, правда ли он верит в то, что подкова приносит удачу, он ответил: «Нет, но мне сказали, что она приносит удачу даже тем, кто в нее не верит»<sup>8</sup>. Если эту историю немного подкрутить, то она предстанет в следующем виде:

---

<sup>7</sup> Hintikka, 1961. Касательно некоторых способов применения см.: Elster, 1978a, p. 81ff.

<sup>8</sup> Segre 1980, p. 171.

1) Нильс Бор верит, что «подковы не принесут мне уда-  
чи».

2) Нильс Бор верит, что «подковы приносят удачу тем, кто  
не верит в то, что они принесут им удачу».

Между убеждениями, заключенными в кавычки в (1) и (2), противоречия нет, но оно появится, если к ним мы добавим само (1). Таким образом, если допустить (а я думаю, что следует это сделать), что нам хотелось бы распознать некую систему убеждений как противоречивую на интуитивных основаниях, нам нужен комплексный критерий, который даст результат, согласующийся с интуицией.

Для задания критерия непротиворечивости для желаний мы должны сначала внимательнее приглядеться к характеру действия, о котором идет речь. Грубо говоря, действие может рассматриваться либо как *делающее что-то*, либо как *вызывающее что-то*. Когда я беру яблоко из чаши с фруктами, я не запускаю каузальный процесс во внешнем мире: я просто это делаю (своевольно). Наоборот, когда я разби-  
ваю окно, бросая в него пепельницу, я вызываю изменение в мире, запуская каузальный процесс, который вскоре становится независимым от моей воли. (Да, при другом описании эти характеристики можно поменять местами, но меня сейчас интересуют описания, в которых действие выполняется намеренно). Объяснения приведенных действий нельзя

совершенно уподобить друг другу, хотя оба подпадают под общую схему рационального действия. Я хочу яблоко и беру его: этим все сказано. Рискуя показаться занудой, я могу добавить, что я верю в то, что там есть яблоко. А если я хочу получить более сильную форму объяснения, я бы также сказал, что яблоко в данный момент является тем, что я больше всего хочу по сравнению с другими вариантами, которые я считаю доступными. Короче говоря, я *предпочитаю* яблоко. Нет необходимости идти дальше и добавлять (ложно), что я беру яблоко *для того, чтобы* вызвать некоторое ощущение в моих органах вкуса или чтобы максимизировать определенное ощущение. Это было бы истинно только в нестандартных случаях. Однако я должен добавить, что ощущения вкуса могут обладать объяснительной силой с одним промежуточным шагом: они участвуют в образовании и подкреплении предпочтений. На них можно ссылаться при объяснении моего желания, но не при его описании (см. также II.10 ниже).

И все же в случае пепельницы для объяснения действия нам потребуется нечто большее, чем предпочтения, если только это не был *acte gratuit* (произвольный акт), как в «Подземельях Ватикана» Андре Жида. Чтобы понять некое действие, мы должны предположить наличие плана и оговорить будущее положение дел, ради которого тот осуществляется. Цель (разбитое окно) могла быть достигнута многими средствами. Согласно первому объяснению, я просто по-

лагал, что бросить в окно пепельницу – один из способов ее достичь; согласно второму, более амбициозному, я полагал, что это наилучший способ. Можно спросить: «Почему он бросил пепельницу?», намереваясь либо узнать, был ли это акт, выражающий гнев, или же у агента была инструментальная цель разбить окно, либо изучить причины битья окон, либо понять, почему была выбрана пепельница, а не что-то еще. Если сосредоточиться на последнем вопросе, то на первый план выйдут различия между предпочтениями и планами. Выбор пепельницы вместо чашки – это иной вид действия, нежели выбор яблока вместо апельсина. И я буду утверждать, что совершенно разные критерии непротиворечивости применяются к действиям в соответствии с предпочтениями и к действиям в соответствии с планами.

Критерий непротиворечивости для предпочтений как минимум включает в себя *транзитивность*: если я предпочитаю скорее  $a$ , чем  $b$ , и скорее  $b$ , чем  $c$ , то я должен предпочесть  $a$ , нежели  $c$ . Более сложный критерий непротиворечивости требуется, когда определяются предпочтения для вариантов с более сложной внутренней структурой. Я буду рассматривать два таких усложнения, связанных с *вероятностью и временем*.

Предпочтения могут определяться вероятностно, посредством *лотерей* – сочетаний вариантов, причем некоторые из вариантов сами могут быть лотереями. Это важно с практической точки зрения, а также имеет ключевое значение для

построения функции полезности, которая позволяет сравнивать интенсивность предпочтений<sup>9</sup>. Как правило, в рамках подобного подхода принимается *принцип доминирования*: если некто предпочитает  $a$ , нежели  $b$ , и при этом  $p > q$ , то он должен рационально предпочитать вариант с получением  $a$  с вероятностью  $p$  и  $b$  с вероятностью  $(1 - p)$  варианту получения  $a$  с вероятностью  $q$  и  $b$  с вероятностью  $(1 - q)$ . Кроме того, привлекается *принцип редукции*: в случае если составная лотерея (та, что содержит в качестве вариантов другие лотереи) редуцируется к простым лотереям очевидным образом, предпочтения должны оставаться теми же самыми. Оба этих допущения оспаривались<sup>10</sup>.

Предпочтения могут также определяться посредством целых *последовательностей* вариантов, что важным образом подчеркивает фактор времени. В частности, мы можем определить понятие *временных предпочтений* как выражение относительной важности, которую некто в определенный момент придает различным последующим моментам или периодам времени. Такие предпочтения, как правило, предполагают *дисконтирование* будущего, то есть придание меньшего веса будущим потреблению или полезности, чем настоящим. Они становятся жертвой двух видов иррациональности, которые мы можем соответственно назвать «невоздержанностью» (или, более нейтрально, «нетерпением») и *противоре-*

---

<sup>9</sup> Подробно о построении см. Luce and Raiffa 1957: ch. 2.

<sup>10</sup> См., напр.: Dreyfus and Dreyfus 1978; Tversky and Kahneman 1979.

чивостью. Невоздержанность подразумевает такое дисконтирование будущего, которое выходило бы за рамки того, что может быть оправдано статистикой смертности и аналогичными соображениями. Слабая теория рациональности позволяет нам сказать, что невоздержанность иррациональна, только в том случае, если агент в момент совершения невоздержанных действий не считает, что лучше всего было бы подождать, взвесив все за и против. Мы имели бы дело с проявлением слабости воли, упомянутой ранее. С другой стороны, принятие широкой теории рациональности позволило бы нам охарактеризовать невоздержанность как иррациональную даже в отсутствие подобного конфликта<sup>11</sup>. В свою очередь, противоречивые временные предпочтения иррациональны даже согласно слабой теории<sup>12</sup>. Непротиворечивость временных предпочтений определяется требованием о том, чтобы план, построенный в момент  $h$  для разме-

---

<sup>11</sup> Касательно вопроса об иррациональности временных предпочтений см.: Maital and Maital 1978. Авторы выступают за то, чтобы считать временные предпочтения рациональными из-за максимизации полезности, то есть рациональными в слабом смысле слова. См. также доказательства, что дисконтирование будущего логически вытекает из набора разумных (хотя и не до конца убедительных) допущений касательно формы функции полезности, приведенных в: Koopmans 1960 и Koopman, Diamond, and Williamson 1964.

<sup>12</sup> Strotz 1955–1956. См. также: Elster 1979: ch. II.5. Пользуясь случаем, хочу указать на серьезные математические изъяны в моем раннем рассмотрении проблемы противоречивости временных предпочтений. В частности, аргумент, касающийся «размещения непротиворечивого планирования» (Elster 1979: 73ff), в основном ошибочен. Я благодарен Аананду Хилланду за то, что он заметил эти изъяны. Они были исправлены в последующем итальянском издании.

щения потребления в промежутке между моментами  $t_2$  и  $t_3$ , оставался в силе, когда наступает момент  $t_2$ , при условии что не происходило изменений в личности агента или во множестве допустимых решений. Без непротиворечивых временных предпочтений человек никогда не сможет придерживаться прошлых планов. Можно показать, что такие временные предпочтения должны быть экспоненциальными, чтобы будущее дисконтировалось постоянными темпами. Согласно утверждениям Джорджа Эйнсли, неэкспоненциальные предпочтения пронизывают жизнь человека – и агент может стратегическим образом использовать эту особенность для преодоления невоздержанности<sup>13</sup>. Если вкратце, при группировке нескольких будущих выборов увеличится вероятность того, что в каждом случае будет выбираться вариант с позднейшей и более высокой наградой. Однако это решение проблемы импульсивности может оказаться не лучше первоначального затруднения, так как привычка группировать выборы способна привести также к окостенелому и навязчивому поведению. Помимо невоздержанности (*incontinence*) и временной противоречивости (*inconsistency*) время также приносит опасность *непостоянства* (*inconstancy*) или иррационального изменения предпочтений (включая временные<sup>14</sup>). Разумеется, отнюдь не все изменения предпочтений иррациональны; временами иррационально не менять свои предпо-

---

<sup>13</sup> Ainslie 1982.

<sup>14</sup> По поводу этой идеи см.: Meyer 1977 и Samuelson 1976.

чтения в ходе обучения. Но я отложу обсуждение этого вопроса, поскольку в данном случае мы должны были бы обратиться к более широкому понятию рациональности. В разделе I.4 я приведу пример эндогенного изменения предпочтений, которое может быть сочтено иррациональным по чисто формальным критериям, но в целом нам следовало бы обратиться к существенным соображениям автономии.

Теория рационального выбора часто требует, чтобы предпочтения были *полными*, а не только непротиворечивыми – для любой пары вариантов агент должен быть способен выразить предпочтение одного из них или, за отсутствием предпочтения, безразличие. С принятой здесь точки зрения, нет сильных доводов в пользу этого условия. По сути дела, можно утверждать, что иррационально слишком полагаться на один вариант, если очень мало известно об обоих. Как минимум было бы иррационально слишком полагаться на такие предпочтения<sup>15</sup>. В контексте построения моделей ясно, что полное упорядочение имеющихся вариантов – куда бо-

---

<sup>15</sup> Cyert and de Groot 1975: 230ff. Сходный, но существенно отличающийся довод предлагается Токвилем: при демократии люди «боятся самих себя, боятся, как бы при смене их собственных желаний им не пришлось сожалеть о невозможности бросить то, что некогда составляло предмет их вожелений» (Tocqueville 1969: 582; Токвиль, 1992: 423). Отсюда склонность американцев воздерживаться от долгосрочных потребительских благ. Хотя Сизер и де Грут утверждают, что рациональный человек должен предвидеть, что его вкусы могут измениться в связи с приобретением нового опыта, Токвиль считает, что американцы – рационально или нет – действуют исходя из того, что их вкусы иррационально изменятся в будущем.

лее мощное понятие, чем их частичное упорядочение. Но если все-таки руководствоваться реальностью, а не удобствами, по-видимому, приходится выбирать между постулированием упорядочения частичных или неполных предпочтений и постулированием полных предпочтений, подверженных эндогенным изменениям по мере того, как агент больше узнает об альтернативах. Постулирование предпочтений одновременно и полных, и устойчивых кажется слишком далеким от реального мира.

Часто допускается, что помимо непротиворечивости и полноты такие предпочтения обладают свойством *непрерывности*. В самом общем смысле это означает, что если некто предпочитает  $a$ , нежели  $b$ , и  $a$  претерпевает совсем небольшое изменение (сколь угодно малое), предпочтения не должны обращаться<sup>16</sup>. Требование нарушается в случае так называемых неархимедовых предпочтений, одним из важных примеров которых является структура лексикографического предпочтения, включающая в себя иерархию ценностей. Если я умираю от голода и мне предлагается вариант, включающий в себя буханку хлеба и прослушивание записи Баха, и другой вариант, тоже включающий в себя буханку, но с прослушиванием записи Бетховена, тогда любовь к Баху может заставить меня выбрать первый вариант. Однако если в первом варианте не останется ни малейшей крошки хлеба, тогда я переключусь на второй вариант, потому что при голоде

---

<sup>16</sup> Более точную формулировку см. в: Rader 1972: 147ff.

калории несопоставимо важнее музыки. В таком изменении предпочтений нет ничего иррационального, и потому непрерывность не может входить в состав рациональности<sup>17</sup>. Но в контексте построения моделей подобное условие очень важно, поскольку транзитивные, полные и непрерывные предпочтения могут быть представлены при помощи *функции полезности* с реальными величинами.

Здесь возникают два замечания. Во-первых, максимизировать полезность не значит заниматься осуществлением плана, выбором наилучших средств для достижения свободно поставленной цели. В современной теории полезности это, по сути, лишь эвфемизм для предпочтений, и он ничего не говорит о более или менее удовлетворительных ментальных состояниях, которые могли бы рассматриваться в качестве цели поведения. К тому же имеются веские основания полагать, что ординалистская концепция благосостояния заходит слишком далеко. Благодаря интроспекции мы знаем наверняка, что удовольствие, счастье и удовлетворение – важные понятия, однако они нуждаются в концептуальной разработке, осуществить которую не так-то просто. Мой тезис состоит в том, что, даже если кому-то удастся установить кардинальную меру полезности, будет ошиб-

---

<sup>17</sup> Более сильный аргумент см. в: Georgescu-Roegen 1954. Риторика Маркузе (Marcuse 1964; Маркузе 1994) становится понятной именно в этих рамках: если предпочтения можно разметить на реальной линии, мы имеем дело с «одномерным человеком». Аналогично Борч отмечает, что постулат о непрерывных предпочтениях равносителен высказыванию «все имеет свою цену» (Borch 1968: 22).

кой считать, будто бы действие всегда можно будет объяснить в категориях максимизации полезности, подобно тому как инвестиции могут объясняться в категориях максимизации прибыли. Последняя операция (в стандартных моделях<sup>18</sup>) понимается как план, предпринимаемый сознательно и *ex ante*, тогда как сознательная и намеренная попытка максимизировать полезность часто оказывается саморазрушительной. Это банальность, но очень важная: счастье часто ускользает от тех, кто активно к нему стремится. Дальнейшему анализу этой идеи посвящена большая часть главы II. Пока я лишь подчеркну, что, даже если действия могут иногда объясняться как *попытки* максимизировать полезность в смысле *ex ante*, у нас нет оснований полагать, что они всегда будут удаваться; скорее наоборот<sup>19</sup>. С другой стороны, как я отмечал ранее, если максимизирующие полезность последствия поведения и *могут* привлекаться для его объяснения, в таком случае они дают лишь каузальное объяснение предпочтений. Доставляющие удовольствие внутренние состояния играют важную роль в объяснении поведения, но вовсе не в качестве его осознанной цели.

Во-вторых, полезно сравнить *рационального* человека с

---

<sup>18</sup> Обсуждение нестандартных моделей см. в: Elster 1982a: ch. 6.

<sup>19</sup> Следует различать объяснения в категориях предполагаемых и в категориях действительных последствий поведения, как подчеркивается ван Парейсом (van Parijs 1981) и мной (Elster 1982a), хотя, конечно, нет общей презумпции, что предполагаемые последствия не смогут реализоваться, за исключением класса случаев, которые будут предметом главы II настоящей книги.

*экономическим человеком.* Первый обладает – в том слабом смысле, который мы сейчас обсуждаем, – только непротиворечивыми предпочтениями и (забегая вперед) непротиворечивыми планами. Второй – уже более совершенное существо: оно имеет предпочтения, которые не только непротиворечивы, но также полны, непрерывны и *эгоистичны*. Да, экономисты построили множество разнообразных моделей, предполагающих неэгоистические предпочтения<sup>20</sup>, но инстинктивно они всегда пытались вывести любое внешне неэгоистическое поведение из эгоистических предпочтений<sup>21</sup>. Вероятно, это хорошая научная стратегия: когда берешься за объяснение поведения, сначала предполагай, что оно эгоистично; в противном случае предполагай, что оно

---

<sup>20</sup> См. полезное исследование и обсуждение: Kolm 1981a.

<sup>21</sup> См., в частности, важный синтез биологических соображений и соображений, основанных на теории игр (Axelrod and Hamilton, 1981. Они используют модель секвенциональных дилемм заключенного, чтобы показать: (а) что по-настоящему альтруистическая мотивация может возникать из естественного отбора по чисто эгоистическим критериям и (б) что некоторые случаи внешне альтруистической мотивации можно объяснить, не допуская ничего, кроме эгоистической рациональности. Другими словами, если люди ведут себя как альтруисты, то либо потому, что они были запрограммированы заботиться о других, либо потому, что они подсчитали, что выгодно симулировать заботу о других. Первое объяснение, хотя оно в некотором смысле и редуccionистское, дает возможность для рационального сопротивления биологическому редуccionизму, воплощенному во втором. Однако, вероятно, есть случаи, которые сопротивляются и биологическому редуccionизму, если только не постулировать, что ослабляющий приспособляемость альтруизм можно объяснить тем обстоятельством, что «не стоит перегружать зародышевую плазму информацией, необходимой для такого рода приспособления» (Williams 1966: 206).

хотя бы рационально; если же и это не удастся, то, по крайней мере, предположи, что оно намеренно. Но допущение, что все формы альтруизма, солидарности и самопожертвования на деле являются крайне утонченными формами эгоистического интереса, в конечном счете оказывается обосновано банальной уловкой, согласно которой люди заботятся о других людях, потому что не хотят, чтобы чужие страдания причиняли страдание им самим. И даже на эту уловку, как показал Аллэн Гиббард, можно возразить, что рациональные минимизаторы страданий зачастую могли бы воспользоваться более эффективными средствами, чем помощь другим людям<sup>22</sup>.

Теперь я обращусь к планам и критериям их непротиворечивости. Задаваться вопросом о таких критериях значит допускать, что действие может быть намеренным и тем не менее не быть рациональным. Я и впрямь утверждаю, что такие действия существуют, и большая часть главы II посвящена их важному подклассу. Прежде чем я перейду к обсуждению критериев непротиворечивости для планов, позвольте мне указать на различие между этой претензией и претензией иного рода, согласно которой человек может действовать преднамеренно, будучи в целом иррациональным. Нам вновь придется согласиться с Дэвидсоном в том, что глобальная рациональность – это предварительное условие вменения человеку намерений, пусть даже иррациональных. Мы

---

<sup>22</sup> Gibbard 1986.

должны иметь возможность приписывать смысл человеку в целом, чтобы сказать, что некоторые из его планов не имеют смысла.

Рациональный план должен удовлетворять двум критериям. Во-первых, конечное состояние, исходя из которого он определяется, должно быть логически согласованным. Если утверждения Сартра о том, что в основе своей все мы хотим быть одновременно *en-soi* и *pour-soi*, покоиться в себе подобно вещи и в то же время находиться на некотором расстоянии от себя, которое позволяло бы нам этим наслаждаться, верны, мы и правда стремимся к логически или концептуально противоречивой цели. Действие в соответствии с этим желанием обречено на провал, как если бы мы пытались быстро обернуться, чтобы поймать собственную тень. Желание одностороннего признания, то есть желание быть признанным другим, которого вы сами не признаете, – это тоже желание вызвать к жизни состояние, которое не может исполниться, потому что признание по определению взаимно<sup>23</sup>. Так, согласно этому первому критерию, условие непротиворечивости плана состоит в том, что должен существовать возможный мир, в котором он осуществляется.

Однако, как и в случае убеждений, нам требуется второй критерий: должен быть возможный мир, в котором план осуществлен обдуманно, то есть в котором можно найти и

---

<sup>23</sup> Более подробное доказательство того, почему это одностороннее признание – внутренне противоречивая идея, см. в: Elster 1976; 1978a: 70ff.

план, и его выполнение. Возьмем, к примеру, план вести себя спонтанно. Нет ничего противоречивого в конечном состоянии, которое определяет план, поскольку люди часто ведут себя спонтанно. Но попытка быть спонтанным – это саморазрушительный план, поскольку сама попытка мешает осуществлению цели. Есть возможный мир, в котором я веду себя спонтанно, но нет мира, в котором я планирую вести себя спонтанно, и мне это удастся. Планы, нарушающие первый критерий, являются логически или концептуально противоречивыми; планы, нарушающие второй, но не первый критерий, прагматически противоречивы. В главе II я занимаюсь почти исключительно последним видом противоречивых планов.

Я хотел бы завершить этот раздел несколькими замечаниями о двусмысленности в понятии рационального поведения, связанной с различием действия как *единственного* и действия как *наилучшего* способа осуществить желание. Стало быть, меня интересуют единственность и оптимальность рационального поведения. Я выдвину предположение, что агент хочет максимизировать некоторую цель, то есть осуществить план наилучшим способом, и спрошу, как ему можно помешать в достижении цели. Очень похожие доводы можно построить при помощи предпочтений, а не планов, но это я оставляю читателю.

Формы максимизирующего поведения отличаются, во-первых, характером среды и, во-вторых, степенью, в которой

она известна агенту. Среды делятся на пассивные, или параметрические, и стратегические; степень, соответственно, – на определенность, риск и неопределенность. Среди полученных подвидов решение в условиях определенности в параметрической среде – стандартная проблема оптимизации. Но даже в этом простом случае мы сразу замечаем, что здесь нет речи ни о единственности, ни об оптимальности. Скорее всего, тут имеется несколько вариантов, которые в равной степени и максимально хороши с точки зрения выбранной цели<sup>24</sup>. Более того, множество допустимых решений может «плохо себя повести» в том смысле, что оптимального варианта среди них не окажется вовсе. Банальный пример – задача найти самое маленькое вещественное число строго больше нуля; еще более важный пример – отсутствие оптимальных стратегий в экономическом планировании<sup>25</sup>.

В случае параметрических решений в условиях риска максимизируемый показатель становится ожидаемой величиной объективной функции или какой-либо ее модификации, которая учитывает неприятие риска и необратимость.

---

<sup>24</sup> В том, что касается обсуждения этого вопроса, ср.: Ullmann-Margalit and Morgenbesser 1977. Теория общего равновесия признает возможность того, что у производителя и потребителя имеется несколько вариантов, одинаково и максимально хороших с точки зрения максимизации прибыли и удовлетворения предпочтений соответственно. Но вопреки обычным канонам науки данная теория будет сопротивляться попыткам сделать выборы однозначно определенными, поскольку могут разрушиться свойства непрерывности, от которых зависит доказательство существования равновесия.

<sup>25</sup> Heal 1973, ch. 13.

Случай параметрических решений в условиях неопределенности более противоречив, поскольку многие отрицают, что есть такая вещь, как настоящая неопределенность или неведение, то есть случаи, в которых мы не можем приписать никаких численных вероятностей возможному исходу действия. Я не могу излагать здесь свои доводы в пользу того, что такие случаи существуют и что в действительности они очень важны<sup>26</sup>. Итак, предположив, что они существуют, мы знаем: рациональным будет учесть только наилучшие и наихудшие последствия, приписываемые данному образу действий<sup>27</sup>. Поскольку имеется множество способов это сделать – например, избрать действие с наилучшими наихудшими последствиями или действие с наилучшими наилучшими последствиями – отсюда следует, что ни единственность, ни оптимальность не выполняются. Из того обстоятельства, что у нас нет причин выбирать между «максимумом» и «максимумом», не следует вывод, что они оба оптимальны, как в первом случае, рассматривавшемся в последнем абзаце. Уподоблять друг другу эти два случая значит путать безразличие и несопоставимость.

Если среда стратегическая, мы попадаем во владения теории игр. Не вдаваясь в детали, теорию игр можно рассматривать как инструмент, даже *специальный* инструмент для работы сразу с тремя группами взаимозависимостей, которые

---

<sup>26</sup> Дальнейшее обсуждение см. в: Elster 1982a: app. 1.

<sup>27</sup> Arrow and Hurwicz 1972.

пронизывают всю общественную жизнь:

- 1) Награда каждого зависит от наград всех из-за влияний зависти, альтруизма и т. д.
- 2) Награда каждого зависит от действия всех в силу общей социальной причинности.
- 3) Действие каждого зависит от действия всех в силу стратегических рассуждений.

Последний пункт является особым вкладом теории игр. Могу добавить, чтобы не возникло впечатления, будто я считаю теорию игр решением всех проблем: не думаю, что она способна справиться со следующим положением:

- 4) Желания каждого зависят от действий всех.

Индивидуальные предпочтения и планы социальны по своему происхождению, что не означает, что они обязательно являются социальными по своему охвату: цели индивида могут не включать в себя благосостояние других. Глава III посвящена по большей части обсуждению формирования предпочтений.

Взаимозависимость выборов ключевым образом связана с понятием *точки равновесия*, то есть набора стратегий, которые оптимальны друг против друга. В этом случае *решение* игры можно определить как точку равновесия, к которой

будут молчаливо стремиться все агенты. У некоторых игр нет точки равновесия, к примеру: «Каждый игрок записывает число. Тот, кто записал самое большое, получает от каждого из остальных игроков сумму, соответствующую разнице между числами, которые они записали». Так может работать гиперинфляция. У других игр более одной точки равновесия, ни одна из которых не выделяется настолько, чтобы быть выбранной в качестве решения: «Я хочу пойти в кино, вы – в ресторан, но мы оба хотим быть вместе, а не порознь». Зная друг о друге это *и только это*, мы не сможем рационально сойтись в одном или другом месте<sup>28</sup>. Да, выделенное курсивом условие будет редко выполняться, и потому можно действовать, руководствуясь ожиданиями по поводу того, как поступит другой человек, основанными на представлении о его характере. Однако в других случаях один человек может не знать о другом ничего кроме того, что тот рационален, и сознавать, что рациональность взаимна, и тогда не окажется набора ожиданий, которые можно было бы рационально защищать. Здесь оптимальность дает сбой – нет такого образа действий, лучше которого ничего нельзя придумать.

В других интеракциях в рамках теории игр дает сбой единственность – с весьма интригующими последствиями. Почти всегда, когда решение игры состоит в смешанных

---

<sup>28</sup> Более подробное обсуждение этой игры, известной как «Битва полов»: ср.: Luce and Raiffa 1957: 90ff, 115ff.

стратегиях, то есть в выборе для каждого агента между имеющимися у него действиями в соответствии с некоторым (оптимальным) распределением вероятностей, индивид не выиграет и не проиграет, если отклонится от поведения, предусмотренного решением, при условии что другие будут его придерживаться. В частности, индивид не проиграет, выбрав «максиминную» стратегию, то есть альтернативу, приносящую ему максимально возможное благосостояние, при допущении, что другие выберут стратегию, делающую его благосостояние максимально низким. Таким образом, есть искушение выбрать эту стратегию, зная, что, если другие будут придерживаться решения, ничего не проиграешь, а если не будут, то ущерб, по крайней мере, будет ограничен. Однако если вы знаете, что другие так же рациональны, как вы сами, само представление о том, что они могут сделать то же самое, могло бы вас сдерживать, порождая в вас надежду, что и других тоже удастся сдержать подобным образом. Ясно, что ситуация крайне неустойчива. Условие, согласно которому от предательства не должно быть выгоды, гарантирует единичность решения, но индивидуальное поведение привязано к этому куда более стабильно в случае, когда игрок в действительности проигрывает от предательства<sup>29</sup>.

Я говорил о трех случаях, в которых оптимальность не срабатывает: нестабильные множества возможностей, принятие решений в условиях неопределенности и игры без ре-

---

<sup>29</sup> Данная трудность подчеркивается в: Harsanyi 1977.

шения. Они дают *специальный аргумент для разумной достаточности*, оставляя общую аргументацию до следующего раздела. Когда не определен ход действий, «лучше которого нельзя придумать», придется довольствоваться чем-то, что *достаточно хорошо* или удовлетворительно, а не оптимально. В случае планирования речь может идти о замене оптимального плана приемлемым<sup>30</sup>. В случае стратегических интеракций – о максиминном поведении, по сути заменяющем оптимальность осторожностью. Заметьте, однако, что для игр, не имеющих точки равновесия, максиминная стратегия может быть и не определена. Если гиперинфляцию рассматривать как игру, в которой награда достается группе, способной договориться о большей прибавке к зарплате, чем у других, в рамках ее не окажется такого требования, которое гарантировало бы «удовлетворительный» исход для конкретной группы или ограничило бы ущерб, который другие могут ей причинить. В таких извращенных структурах взаимодействия разум отнюдь не помощник – скорее он велит нам изменить ситуацию, когда в ней нет возможности делать рациональный выбор.

---

<sup>30</sup> Hammond and Mirrlees 1973. Строго говоря, понятие приемлемого плана – не пример разумной достаточности, поскольку оно находится в широком классе случаев, детерминированных однозначным образом. Хотя понятие уровня устремлений входит в определение приемлемого плана, он оказывается не зависящим от любого конкретного уровня устремлений. Тем не менее его рациональное основание очень напоминает основание понятия о разумной достаточности.

## **1.3. Индивидуальная рациональность: широкая теория**

Теперь я хочу перейти к изучению более существенных коннотаций рациональности. Сказанное ранее не мешает нам говорить о самоубийстве, убийстве или геноциде как о рациональном поведении. Точно так же не было приведено никаких доводов в пользу того, чтобы исключить из области рационального танец для вызывания дождя у хопи, изучение гороскопа перед тем, как инвестировать в ценные бумаги, или решение вернуться домой, если дорогу перебежала черная кошка. Если говорить строго логически, существует и такая возможность, что весь мир участвует в заговоре, чтобы помешать мне добиться цели, и в слабом смысле, вероятно, для меня было бы вполне рационально действовать на этом основании. Однако ясно, что этот смысл слишком слаб. Нам требуется более широкая теория рациональности, которая выходит за рамки чисто формальных соображений предыдущего раздела и позволяет внимательно присмотреться к существу желаний и убеждений, которые участвуют в действии. Мы хотим иметь возможность сказать, что действовать рационально значит действовать непротиворечиво на основании убеждений и желаний, которые являются не только непротиворечивыми, но и рациональными.

С другой стороны, не стоит до такой степени расширять

понятие рациональности, чтобы оно охватывало все положительные свойства, которые нам хочется придать нашим убеждениям и желаниям. Я утверждаю, что *между слабой теорией рациональности и полной теорией истины и блага имеется место для широкой теории рациональности, а также потребность в ней*. Заявить, будто истинность является необходимым условием рациональных убеждений, значит просить слишком много; но заявить, что достаточно непротиворечивости, значит запросить слишком мало. То же, пускай и более спорно, относится к рациональным желаниям: требование непротиворечивости слишком слабое, требование этического блага – слишком сильное.

Я предлагаю оценивать широкую рациональность убеждений и желаний, глядя на то, как они формируются. Убеждение может быть непротиворечивым и даже истинным, желание – непротиворечивым и отвечающим требованиям морали, – и тем не менее мы можем не решиться с полной уверенностью называть их рациональными, если они были сформированы нерелевантными причинными факторами, слепой каузальностью психики, действующей «за спиной» личности. Акцент здесь должен ставиться на «нерелевантном» и «слепом», а не на каузальности как таковой. Я не утверждаю, что убеждения и желания становятся иррациональными в силу того, что они имеют каузальное происхождение. У всех желаний и убеждений есть (достаточное) каузальное происхождение, но у некоторых из них своего рода неправиль-

ная каузальная история – и потому они иррациональны. Поскольку затруднительно объяснить, что именно может считаться правильной историей, я мало что могу рассказать об этой (важнейшей) проблеме. Здесь и далее я могу поведать обо всех неправильных видах историй. В конце этого раздела я дам краткую типологию того, как убеждения и желания могут быть искажены и извращены. Некоторые из этих искажений будут подробнее рассматриваться в главах III и IV.

Рассмотрим сначала убеждения и их формирование. Ясно, что убеждение может быть истинным и при этом нерациональным, рациональным и при этом неистинным. Рациональность убеждений (будучи взята в своей сути) затрагивает отношения не между убеждениями и миром, но между убеждениями и имеющимися фактами. Более того, как подробнее говорится в главе IV, утверждение о том, что убеждение рационально, должно основываться не на *сравнении* фактов с убеждениями, а на изучении его действительной каузальной истории, ведь человек нерациональным путем может прийти к убеждению, которое также окажется фундаментальным в фактах. Кроме того, недостаточно сказать, что убеждение рационально, потому что вызвано фактами, которые делают рациональной веру в него, поскольку в конкретном случае оно вполне может вызываться фактами недолжным образом, например через механизм компенсирующих ошибок. Все это перекликается во многих аспектах с обсуждением в разделе I.2 отношений между действием и убежде-

ниями с желаниями, которые становятся его причинами.

Положительно охарактеризовать рациональные убеждения можно через отсылку к понятию *суждения*, определяемому как способность синтезировать обширную и весьма разрозненную информацию, которая имеет более или менее четкое отношение к решаемой проблеме, не приписывая ни одному элементу или группе элементов в составе незначительной важности. Ясно, что определение несколько бесполезно, но в реальности этого феномена сомневаться не приходится. Все мы знаем людей, у которых есть это качество, и людей, которые его лишены. При некоторых занятиях без него невозможно обойтись, и те, у кого его нет, вскоре оказываются из них исключены. Крайние случаи – конкурентный рынок, на котором фирмы, управляемые людьми без способности суждения, быстро становятся банкротами, и война, где военачальники и солдаты, лишенные ее, подвергаются большому риску. В определенной степени это также верно и в отношении политиков, для которых способность суждения и здравый смысл, в дополнение к чистому упорству и некоторой наработанной бесчувственности, гораздо важнее ума в его традиционном понимании. Что касается наук, их можно грубо разделить на те, в которых требуется прежде всего логика, и те, которые требуют менее формального применения способности суждения<sup>31</sup>.

---

<sup>31</sup> Что касается общих замечаний о роли суждения в науке, см. заключительную главу в: Newton-Smith 1981.

Однако вскоре окажется, что бойко говорить об «имеющихся фактах» недостаточно, поскольку встает важнейший вопрос о том, какое именно количество фактов следует рационально накопить, прежде чем прийти к убеждению. Этот вопрос допускает разные ответы в зависимости от цели, с которой в дальнейшем будет использоваться данное убеждение, если вообще будет. Для настоящего ученого вопрос о будущем использовании не возникает, поскольку конечная цель его поведения состоит в формулировке истинного убеждения. Поскольку вероятность прийти к истинному убеждению возрастает по мере увеличения числа фактов, может показаться, будто поиски истины сами себя подрывают, так как ученому, занятому ими, придется собирать доказательства вечно, бесконечно откладывая формирование убеждения. Его затруднения напоминают мучения общества, заикленного на потреблении, заставляющем его бесконечно экономить и инвестировать, снова и снова откладывая момент наслаждения плодами, ради которых-то и совершались вклады. В обоих случаях ответ один и тот же: поскольку первоначальная проблема не позволяет никакого решения, рациональная реакция – переформулировать ее так, чтобы найти «удовлетворительные» уровни объема фактов и инвестиций соответственно<sup>32</sup>.

---

<sup>32</sup> Нелегко установить критерий удовлетворительности применительно к фактам в научных теориях. Возможно, самое важное соображение – степень оригинальности и новизны убеждения. Если у идеи есть революционный потенциал, ее рационально поддерживать, даже если она подкреплена лишь умеренным ко-

На первый взгляд может показаться, что легче определить оптимальное количество фактов в практическом контексте бизнес-решений. Здесь пределы для роста числа фактов возникают, прежде чем мы углубимся в парадоксы максимизации истины, поскольку сбор информации создает издержки для фирмы и потому должен предприниматься только в той степени, в которой имеются (или предполагаются) прибыли. Не получать сведений о среде иррационально, а собирать их слишком долго – тоже; должен иметься некоторый оптимальный объем информации, который будет получен фирмой. Но здесь снова встает вопрос, поскольку «выбор самой максимизирующей прибыль информационной структуры требует информации, и неясно, как максимизатор прибыли, желающий получить информацию, ее приобретает – или гарантирует, что он не переплачивает за нее»<sup>33</sup>. Требование оптимального количества фактов сразу же приводит к бесконечному регрессу.

Так звучит *общий аргумент для разумной достаточности*. Он касается не только решений в бизнесе, но и других практических вопросов, в которых возникает конфликт

---

личеством фактов, потому что тогда имеется вероятность того, что научное сообщество в целом сможет исследовать ее полнее, чем это мог бы сделать отдельный ученый за всю свою жизнь. Но дело не только в науке, поскольку, как представляется, способность производить революционные идеи сильно коррелирует с почти навязчивым стремлением не оставлять ни одно возражение без ответа и без изучения.

<sup>33</sup> Winter 1964–1965: 262.

между стремлением вкладывать время и деньги в получение информации и потребностью тратить их, используя полученную информацию<sup>34</sup>. Это означает, что, когда слабая теория рациональности отброшена, связь между рациональностью и оптимальностью полностью разрывается. Рациональное поведение может описываться в качестве оптимизирующего только – или главным образом – по отношению к *данным* убеждениям касательно мира, однако принципы рациональности, управляющие получением убеждений, не могут быть объяснены в категориях оптимизации. Аргумент, по крайней мере, выполняется *ex ante* – в том смысле, что нельзя знать заранее, получение какого объема информации окажется оптимальным. Можно предположить, что он не выполняется *ex post*, во всяком случае в ситуациях, когда ответственные лица, принимая (по своей собственной вине) неразумные решения, оказываются устранены конкуренцией, так что остаются лишь те, кто обладает оптимальной структурой информации. Однако выясняется, что эта привлекательная идея, будучи разработана более тщательно, на деле слабее,

---

<sup>34</sup> «В какой-то момент решение должно приниматься на интуитивных основаниях. Это как пойти в большой лес за грибами. Можно исследовать возможности только в определенной ограниченной зоне, но в какой-то момент приходится останавливать исследования и начинать собирать, потому что дальнейшие исследования возможностей того, можно ли найти больше грибов и грибы лучшего качества при дальнейшем углублении в лес, подорвут саму цель похода. Решение об остановке поисков должно приниматься интуитивно, то есть без исследования того, принесут ли дальнейшие поиски улучшение результатов» (Johansen 1977: 144).

чем кажется, если только не принять весьма недальнозоркие допущения насчет свойств среды. В любой реалистической модели (например, конкуренции между фирмами) среда, с которой сталкивается фирма, меняется настолько стремительно, что за время, которое потребуется для разгрома конкурентов, любая информационная структура перестанет быть оптимальной. По сути, мы имеем дело с приспособлением к движущейся мишени<sup>35</sup>

---

<sup>35</sup> См. подробнее: Elster 1982a: ch. 6.

# Конец ознакомительного фрагмента.

Текст предоставлен ООО «ЛитРес».

Прочитайте эту книгу целиком, [купив полную легальную версию](#) на ЛитРес.

Безопасно оплатить книгу можно банковской картой Visa, MasterCard, Maestro, со счета мобильного телефона, с платежного терминала, в салоне МТС или Связной, через PayPal, WebMoney, Яндекс.Деньги, QIWI Кошелек, бонусными картами или другим удобным Вам способом.