

Ограниченная рациональность и принятие решений

О. П. Кузнецов

Институт проблем управления им. В.А. Трапезникова РАН, г. Москва, Россия

Аннотация. Рассматривается теория ограниченной рациональности (часто называемая поведенческой экономикой), развитая в работах Канемана, Тверски, Талера, Становича. Теория показывает, что человек часто поступает нерационально не только в силу объективных причин (необходимости быстро принимать решения в условиях неопределенности и ограниченной памяти), но и из-за неумения использовать те интеллектуальные инструменты, которыми он обладает. Это связано с особенностями взаимодействия двух типов когнитивных процессов – быстрых (Система 1), действующих при минимальном контроле сознания, и медленных (Система 2), сознательных и требующих когнитивных усилий. Постоянно действующая Система 1 принимает стандартные решения. В нестандартных ситуациях необходимо участие рассуждающей и оценивающей Системы 2. Однако это участие не всегда происходит в полной мере из-за «когнитивной скупости» Системы 2, склонной к минимизации усилий. Эту теорию, естественно, дополняет концепция о важной роли эмоций в процессе принятия решений, предложенная известным нейробиологом А.Дамасио.

Ключевые слова: ограниченная рациональность, принятие решений, поведенческая экономика, быстрые и медленные когнитивные процессы, теория перспектив, фрейминг.

DOI 10.14357/20718594190101

Введение

Теория принятия решений, сформировавшаяся как наука в середине XX века, долгие годы исходила из постулата рационального выбора, предполагающего, что человек, принимая решение, действует рационально. При этом рациональность понималась наиболее «жестким» образом: как стремление к максимизации своей функции полезности. Герберт Саймон в книге [1] впервые обратил внимание на то, что человеческая рациональность ограничена. Развернутую концепцию ограниченной рациональности он изложил в 1978 году в работе [2]. Именно в этом году ему была присуждена Нобелевская премия по экономике «За новаторские исследования процесса принятия решений в рамках экономических организаций».

Основные идеи этой концепции заключаются в следующем.

1. Не имеет смысла заниматься оптимизацией решений, не учитывая возможности реализации процедур, ведущих к оптимальным решениям. А эти возможности объективно ограничены когнитивными способностями человека: ограниченным объемом памяти, невысокой скоростью рассуждений, необходимостью принимать решения в реальном времени и в условиях неопределенности.

2. Теория вычислительной сложности показала, что существуют задачи, в том числе связанные с проблемой выбора, которые обладают неустранимой объективной сложностью, и поэтому методы их точного решения неприменимы в реальной жизни.

3. Еще один круг объективных причин, ограничивающих человеческую рациональность – это резко возросший объем информации и возникшая в связи с этим необходимость совершенствовать процедуры отбора существенной информации («управления внимани-

✉ Кузнецов Олег Петрович. E-mail: olpkuz@yandex.ru

ем») и учета побочных последствий принимаемых решений.

4. «Разумные люди приходят к «разумным» выводам в обстоятельствах, когда нет возможности применить классические модели рационального выбора. Мы мало знаем о том, как они это делают. Еще меньше – о том, имеют ли процедуры, используемые ими вместо неприменимых моделей, какие-либо достоинства» [2].

Если второе и третье положения концепции Саймона говорят об объективных ограничениях, не связанных со свойствами когнитивных механизмов человека, то первое и четвертое указывают как раз на эти свойства как источник ограничений. Тем самым они предсказывают будущую роль когнитивной психологии в изучении прикладных проблем рационального выбора.

Так оно и оказалось. Начиная с середины 70-х гг. XX века именно психологические аспекты принятия решений подверглись наиболее пристальному изучению. Начало этому направлению было положено двумя публикациями Д. Канемана и А. Тверски [3, 4]. Краткий обзор состояния проблемы к середине 80-х гг. имеется в статье [5] и ее русскоязычном эквиваленте [6]. Бурная история этого направления, ярко описанная в [7], началась с резкого отрицания его подходов классиками традиционной (рациональной) теории принятия решений, а увенчалась присуждением двух Нобелевских премий по экономике: в 2002 г. – лидеру этого направления Даниэлю Канеману¹, а в 2017 г. – его последователю Ричарду Талеру.

В западной литературе это направление принято называть когнитивной или поведенческой экономикой. По-видимому, потому, что экономика является главным потребителем прикладной теории принятия решений и именно в экономике последствия принимаемых решений чаще удается оценить количественно. К тому же Ричард Талер является профессиональным экономистом. В названии его книги [7], содержащей большое количество примеров, подтверждающих эффективность поведенческого подхода в экономике, присутствует именно этот термин «behavioral economics». Тем не менее, называть направление психологии «экономикой», хотя бы и поведенческой

(или когнитивной), на наш взгляд, неправильно, прежде всего, потому, что эти результаты описывают механизмы принятия решений человеком практически во всех сферах его жизни и деятельности, а не только в экономике. Поэтому мы предпочитаем пользоваться другим названием этого направления – «теория ограниченной рациональности».

Настоящая статья является кратким изложением основных результатов этой теории. Наиболее систематично они изложены в книгах Д. Канемана [8] и К. Становича [9], который также является ведущим специалистом в этой области.

1. Основные понятия теории.

Два вида когнитивных процессов

В центре концепции ограниченной рациональности лежит ее главное открытие: ограниченная рациональность связана не только с ограниченными когнитивными способностями человека. Часто люди не используют свои когнитивные способности даже в тех случаях, когда их достаточно для получения разумных решений.

«Есть множество людей, которые систематически демонстрируют неспособность к рациональному поведению, при том, что обладают более чем нормальным IQ» [9]. Тем самым приходится разделять интеллект и рациональность. Интеллект – это способность решать поставленные задачи. «Мыслить рационально – значит ставить выполнимые цели, предпринимать необходимые шаги с учетом этих целей и собственных убеждений, а также сохранять убеждения, имеющие под собой отчетливую основу» [9].

Психологической основой теории ограниченной рациональности является представление о двух видах когнитивных процессов. Процессы первого вида (Канеман называет их Системой 1) – это быстрые процессы, которые не требуют сознательного внимания, не зависят от контроля высших уровней и могут происходить параллельно с другими когнитивными процессами. Их называют также автономными. Примеры таких процессов – узнавание лиц и предметов, ориентация в пространстве и др.

Процессы второго вида (Система 2) – это медленные процессы, требующие когнитивных усилий – внимания, концентрации, рассуждений и т.д. Они могут идти лишь по одному или в очень

¹ Эту премию Канеман должен был бы разделить со своим постоянным соавтором Амосом Тверски. Однако к тому времени Тверски скончался, а правила Нобелевского комитета не предполагают посмертных награждений.

малом количестве одновременно. Примеры таких процессов – решение задач, заполнение анкет, выбор покупки в магазине и т.д. Одно из важных свойств Системы 2 – умение не смешивать представление о реальном мире с представлением о воображаемых ситуациях, т.е. способность создавать альтернативные модели мира, которые отделены от реальной картины мира и поддаются умственным манипуляциям.

Система 1 работает постоянно и автоматически; ее нельзя отключить, хотя можно подавить. Она быстро оценивает текущую ситуацию, прогнозирует ее ближайшее развитие и передает оценки и прогнозы Системе 2. Пока ситуация близка к обычной (скажем, вы совершаете обычный маршрут с работы домой), Система 2 использует лишь малую часть своих возможностей и принимает предложения Системы 1 совсем или почти без изменений.

Таким образом, в привычных ситуациях Система 1 способна быстро принять решения и осуществить соответствующие действия в реальном времени. Эта способность объясняется в основном наличием обширного репертуара стереотипных ситуаций и соответствующих им образцов поведения (планов, решений и действий), которые человек накапливает на протяжении своей жизни. Распознавая знакомую ситуацию, ассоциативные механизмы Системы 1 вызывают связанные с ней образцы поведения, уже осуществлявшиеся ранее и закрепленные благодаря своей адекватности. Высокая скорость решений и действий, связанных с вызванными образцами, объясняется тем, что они запускаются, так сказать, целиком, без сознательного контроля за составляющими их элементами. Например, быстрое выполнение целенаправленных двигательных действий (моторных актов) обусловлено тем, что «в зоне F5 содержится некий словарь моторных актов, отдельные «слова» которого представлены популяциями нейронов. Некоторые из них кодируют основную цель выполняемого действия (удержание, хватание), другие – способ выполнения конкретного моторного акта (точный захват и т.д.), и, наконец, есть группа нейронов, определяющих временную развертку целостных моторных актов на элементарные движения» [10].

Система 2 мобилизуется, когда ситуация нетипична (например, вы ищете нужный дом в незнакомом городе) или обнаруживается событие, нарушающее модель окружающего мира в

представлении Системы 1 (например, на вашем пути с работы домой возникло какое-то препятствие – перекрыта улица и т.п.). В случае трудностей Система 2 перехватывает управление, и обычно последнее слово остается за ней.

Описанное разделение труда между Системой 1 и Системой 2 в привычных ситуациях дает наилучшую производительность при минимуме усилий. Поскольку большую часть своей жизни обычный человек проводит как раз в привычных ситуациях, то это разделение, сложившееся в ходе эволюции, в среднем весьма эффективно. Однако в нестандартных ситуациях характер взаимодействия должен меняться, и в зависимости от того, как это происходит (или не происходит), могут возникать ошибочные решения.

Причины ошибок различны. Их подробный анализ и составляет предмет исследований в области ограниченной рациональности. Краткое описание начнем со свойств Системы 1, являющихся источником этих ошибок.

2. Свойства Системы 1

Основная функция Системы 1 – отслеживание текущей картины мира, говорящей, что все идет нормально. Основными когнитивными механизмами отслеживания являются узнавание и категоризация. Быстрый процесс узнавания (до сих пор не разгаданный и имеющий мало общего с компьютерным распознаванием образов) заключается в отнесении наблюдаемого объекта (процесса, ситуации) к определенной категории. Категоризация объекта означает автоматический перенос на объект свойств и образцов поведения, свойственных объектам данной категории. Скажем, яблоки – круглые, зеленые или красные, их можно грызть, резать, печь, покупать, выращивать и т.д. С другой стороны, в каждой категории существуют прототипы – «хорошие» (центральные, репрезентативные, типичные) примеры [11-13]. Прототипы играют роль норм: нормальность объекта или ситуации – это близость его к прототипу. Обнаружение аномалий, т.е. сильных отклонений от прототипов происходит очень быстро – за 0,2 сек. При этом не обязательно обнаруживаются отклонения по конкретным признакам; часто бывает достаточно ощущения «что-то не так». Это ощущение хорошо описано в книге [14]: «Предположим, пока вас не было, я заметил что-то в вашей входной двери.... Когда вы

вернетесь домой и попытаетесь открыть дверь, вы быстро обнаружите, что что-то не так. Вы не сразу поймете, что именно не так, но сам факт изменения заметите очень быстро. Когда ваша рука потянется к дверной ручке, вы обнаружите, что она не на своем месте. Или, если вес двери изменился, вы толкнете ее с неправильным усилием и будете удивлены. Вы заметите любое из тысячи изменений за короткий промежуток времени».

Важной мерой адекватности для Системы 1 является *когерентность* (согласованность) воспринимаемых данных. Для текущих ситуаций – это согласованность с образцом, прототипом; для рассказов, обсуждений, аргументации – отсутствие видимых противоречий или ссылка на авторитетный (по мнению рассуждающего) источник. При этом количество и качество данных не имеет значения. Система 1 не отслеживает альтернативы. Более того – они для нее не существуют. Это выражается в принципе WYSIATI – What You See Is All There Is): *Что ты видишь, то и есть*. Всего остального как бы нет.

Другой мерой адекватности для Системы 1 является ощущение *когнитивной легкости* – ощущение, что можно обойтись без Системы 2. И наоборот – когнитивное напряжение мобилизует Систему 2. Когнитивная легкость может вызываться когерентностью, но у нее могут быть и другие причины, не контролируемые сознанием, – яркая подача материала, умело расставленные акценты, отвлекающие от сомнительных мест в аргументации, да и просто хорошее настроение.

Стремление к когнитивной легкости заставляет Систему 1 подавлять неоднозначность, упрощая картину мира и делая поспешные выводы на основе ложной индукции, т.е. неправомерного обобщения. Примером является известный еще с 20-х гг. XX века «*эффект ореола*» (гало-эффект): влияние первого впечатления о человеке на оценку других его качеств. Другой эффект стремления к когнитивной легкости – *подстановка*: если на вопрос не находится быстрого удовлетворительного ответа, Система 1 находит более легкий родственный вопрос и отвечает на него.

Когнитивная легкость связана со слабым участием (или отсутствием) сознательного контроля. Многие процессы, обеспечивающие работу Системы 1, являются бессознательными. Поэтому человек далеко не всегда может ра-

зумно объяснить, почему он принял то или иное решение. Одним из наиболее изученных таких процессов является *прайминг*² (эффект предшествования) – неосознаваемое влияние на поведение (действие, принятие решения и т.д.) предшествовавшего события или контекста. Хорошо известен эксперимент, когда одной группе студентов предлагали завершать фразы, содержащие слова, связанные со старостью, а другой – аналогичные фразы с нейтральными словами. Выходя из аудитории, студенты первой группы шли гораздо медленней второй, хотя до начала эксперимента они шли одинаково быстро. Эффект прайминга издавна используется в рекламе и торговле, когда потребителя различными якобы нейтральными приемами, умело подводят к нужному выбору.

3. Рациональность и интеллект

Вернемся теперь к различиям между рациональностью и интеллектом, которые наиболее четко сформулированы в концепции К. Становича [9]. Суть ее представлена на Рис. 1 (Рис. 3.4 в [9]), где процессы Системы 1 названы автономным мышлением. Процессы Системы 2 делятся на два уровня – алгоритмическое и рефлексивное мышления. Алгоритмическое мышление – это набор интеллектуальных инструментов, механизмов обработки информации. Рефлексивное мышление управляет этим набором. Оно рассматривает цели и убеждения, а также принимает решения о выборе действий с учетом целей и убеждений. Именно оно инициирует (процесс В) вмешательство (процесс А) в работу Системы 1, т.е. передачу управления Системе 2, а также процессы разделения С и D, запускающие моделирование – создание альтернативных моделей мира для обсуждения различных сценариев поведения. Разделение и моделирование – важнейшие когнитивные способности, позволяющие оценивать и критически рассматривать собственные убеждения и желания.

Тесты на интеллект (IQ) позволяют оценить только когнитивные способности алгоритмического уровня. Алгоритмический уровень решает только те задачи, которые ему поставят рефлексивный уровень. Степень включенности

² Подробный обзор различных видов прайминга содержится, например, в статье [15].

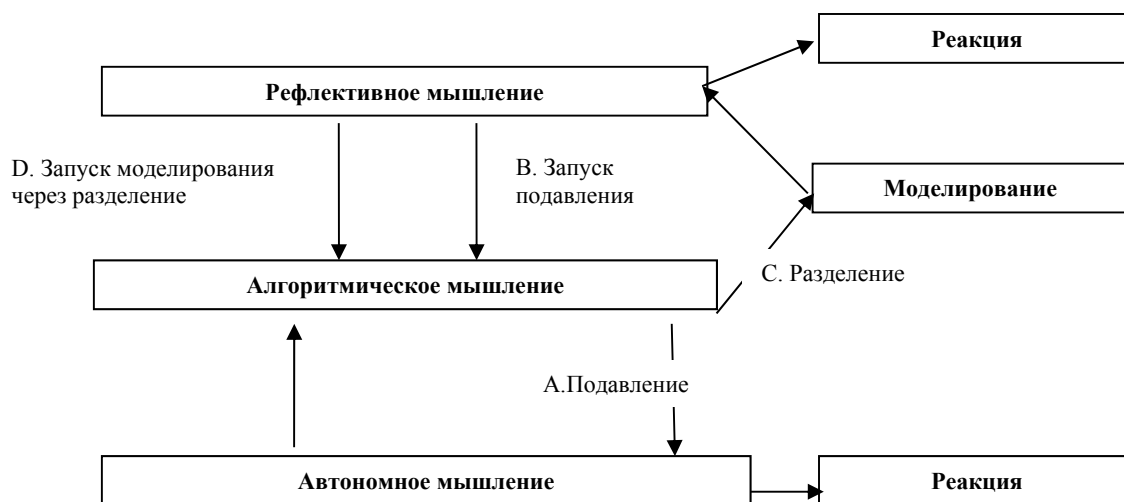


Рис. 1.

рефлексивного уровня и определяет рациональность мышления. При этом следует различать инструментальную и эпистемическую рациональность. Инструментальная рациональность – это способность оптимизировать процесс достижения уже поставленной (возможно, ложной) цели. Эпистемическая рациональность – это соответствие убеждений структуре реального мира и связанная с этим способность выбирать адекватные цели.

Экспериментально подтверждено, что корреляция между интеллектом, измеряемым тестами IQ, и склонностью мыслить непредвзято (рассматривать в равной мере доводы за и против) минимальна. Более того, хорошо известно, что умные люди часто ведут себя удивительно глупо (эффект «умного дурака»). Схема Становича объясняет эти явления. IQ-интеллект – это набор инструментов. «Оценивая его, мы оцениваем двигатель, а не водителя» [9]. Рациональность – это способность пользоваться этим набором всякий раз, когда этого требует ситуация.

4. Система 2.

Ошибки когнитивной скупости

В первом приближении эти ошибки можно разделить на три группы.

1. Система 2 может недооценить нестандартность ситуации и положиться на быстрые решения Системы 1. Но Система 1 не анализирует эффективность принимаемых решений. Она принимает их потому, что они включены в стереотипный образец поведения, а включены в него они потому, что оказались эффективными

раньше в аналогичных ситуациях. У нее есть свои способы реакции и на нестандартные ситуации (например, выбор наиболее близкого по ее мнению образца поведения, уклонение от активного поведения и т.д.), которые часто оказываются неудачными.

2. У Системы 2 есть свои систематические ошибки. По большей части они связаны с неиспользованием имеющихся у нее возможностей, причина которого – склонность к минимизации когнитивных усилий. Эту склонность Канеман [8] называет когнитивной ленью, а Станович [9] – когнитивной скупостью.

3. Система 2 конкретного человека может не иметь необходимых средств (знаний и умений, приобретаемых как результат образования и опыта) для решения возникающих задач и к тому же пользоваться неадекватными средствами (гадалки и знахари для решения медицинских проблем, конспирологические объяснения социальных и политических процессов и т.д.).

Первые две группы ошибок (излишняя доверчивость к Системе 1 и неиспользование своих возможностей) объясняются когнитивной скупостью. Третья группа связана с содержанием набора умственных стратегий и программ, способствующих (или не способствующих) рациональному поведению.

Рассмотрим сначала ошибки когнитивной скупости. Вот первый выразительный пример [9] – задача: женатый А смотрит на Б, Б смотрит на холостого С. Имеет ли место ситуация: тот, кто состоит в браке, смотрит на того, кто в браке не состоит? Варианты ответов: 1) да; 2) нет; 3) невозможно определить. Большинство испытуемых дает третий ответ, потому что не-

известен статус Б. Однако правильный ответ – «да», который получается перебором возможных случаев статуса Б: либо женатый А смотрит на холостого Б, либо женатый Б смотрит на холостого С. Способность автоматически включать перебор случаев у большинства людей отсутствует (он требует концентрации внимания и удержания в сознании всех результатов перебора). Как правило, включение происходит, только когда явно сказано, что это надо сделать.

Эффект привязки. Привязка – это принятие заданной точки отсчета («якоря»). Знакомый большинству людей пример этого эффекта – ситуация купли-продажи, когда покупатель попадает под влияние запрошенной цены. «Цифра, с которой начинается торговля, имеет сильнейшее влияние на исход торга. Тот, кто держит в руках якорь, держит в руках всю сделку» [9]. Канеман [8] считает, что здесь работают два разных механизма. Если ведущую роль играет Система 1 (из-за «лени» Системы 2), то она, стремясь к когерентности, принимает цену-якорь как верную (работает прайминг); остальные альтернативы (корректировки исходной цены) не рассматриваются (принцип WYSIATI). Если же ведущей является Система 2, то корректировки возможны, но только в окрестности первоначальной цены; из-за когнитивной скупости смена точки отсчета как альтернатива не рассматривается.

Фрейминг. Привязку можно считать частным случаем фрейминга. «Базовый принцип фрейминга – пассивное принятие заданной формулировки» [8]. Предложенная субъекту рамка (frame) воспринимается как основа модели мира, и дальнейшие рассуждения ведутся внутри этой рамки, поскольку поиск альтернатив требует больших когнитивных усилий и дополнительной информации. Как и привязка, фрейминг является результатом взаимодействия поспешных выводов Системы 1 и ленивой Системы 2.

Фрейминг в дискуссии заключается в стремлении навязать партнеру свою картину мира через свой дискурс [16] и, в первую очередь, через словарь, содержащий скрытые оценки («временное правительство» или «хунта», «революция» или «мятеж» и т.д.). В политических дискуссиях и информационных войнах определенный дискурс навязывается всему обществу.

Эффект доступности – оценка частоты встречаемости события на основе легкости, с которой в голову приходят примеры. Не слишком частые события могут благодаря СМИ привлечь слишком большое внимание и потому казаться более частыми, чем на самом деле. Например, гораздо большее число людей боится летать на самолете, чем садиться в автомобиль, потому что авиакатастрофы обсуждаются гораздо шире и «громче», при том, что количество жертв ДТП во много раз больше.

Эффект репрезентативности – сходство представленного описания со стереотипом провоцирует поспешный вывод и игнорирование другой информации.

Влияние аффекта – люди часто принимают решения и формируют мнения и оценки, руководствуясь лишь эмоциями.

Эгоцентрические предубеждения – уклонение от объективности. Люди склонны рассматривать любую ситуацию, исходя из собственного положения. Например, они подмечают противоречия в поведении своих оппонентов и не замечают противоречий в поведении своих сторонников; позволяют себе то, что не позволяют другим. Это выразительно иллюстрируется одним из опросов в Канаде: 70% водителей за рулем курят или говорят по телефону, однако примерно тот же процент водителей сильно раздражается, когда видят других, курящих или говорящих по телефону за рулем.

5. Теория перспектив Канемана-Тверски

Одним из главных результатов Канемана-Тверски является разработанная ими «теория перспектив» (prospect theory)³ – исследование принципов, на основе которых люди оценивают альтернативы и принимают решения.

1. Наличие точки отсчета, относительно которой производится оценка. Для финансовых исходов такой точкой обычно является текущее положение, иногда – ожидаемый исход. Исходы выше точки отсчета считаются выигрышами, исходы ниже точки отсчета – проигрышами. Психологическое ощущение бедности – это существование ниже собственной точки отсчета. В этой ситуации небольшие доходы воспри-

³ Первая версия этой теории была опубликована в статье [4] в 1979 г.

нимаются как сокращение потерь, а не как выигрыш; любые покупки – это потери возможности купить что-то другое.

2. Принцип относительной чувствительности к выигрышам и потерям: чувствительность приращения богатства определяется не его абсолютной величиной, а его долей в существующем богатстве: приращение одного миллиона к одному миллиону (т.е. в два раза) воспринимается гораздо сильнее, чем приращение того же миллиона к 10 миллионам. Этот принцип был предложен еще Д.Бернулли, однако у него он симметричен: при одном и том же богатстве одинаковые потери и выигрыши воспринимаются одинаково. Канеман и Тверски на основании своих исследований предлагают противоположный принцип – неприятие потерь.

3. Неприятие потерь: потери кажутся крупнее, чем такие же выигрыши. Экспериментально полученная оценка коэффициента неприятия потерь колеблется от 1,5 до 2,5: потеря \$100 равносильна выигрышу \$200.

Причиной глубоко укоренившегося в сознании человека принципа неприятия потерь является его биологическая природа – унаследованные в ходе эволюции механизмы выживания, которые в первую очередь требуют избегать угроз. Исследования психологов показывают, что «информация о плохом обрабатывается тщательнее. Мотивация личности сильнее, когда она избегает плохих самоопределений, чем когда работает над положительными. Дурные стереотипы и неприятные впечатления быстрее формируются, чем хорошие, и более устойчивы к попыткам их устранения» [8].

6. Система 2. Ошибки содержания

Ошибки содержания [9] – это дефекты или недостаток когнитивных инструментов (умственных программ).

Дефицит критического мышления, связанный с незнанием правил научного мышления: неумение оценить корректность рассуждений, уклонение от рассмотрения альтернатив и поиска опровержений предлагаемым суждениям.

На вопрос, верен ли силлогизм: «Всем живым существам нужна вода. Розам нужна вода, следовательно, розы – живые существа». 70% студентов отвечают «да» [9], хотя такой же по структуре силлогизм «Всем насекомым нужен кислород. Мышам нужен кислород, следова-

тельно, мыши – насекомые» эти же студенты считают неверным. Такая оценка первого силлогизма вызвана непониманием разницы между логической корректностью рассуждения и фактической истинностью заключения.

Известный тест Уэйсона [17] хорошо иллюстрирует уклонение от опровержений. Он заключается в том, что испытуемым показывают 4 карты, на лицевой стороне которых написаны *A*, *D*, 4, 7. Сообщается гипотеза, которую надо проверить: «если на одной стороне – гласная, то на другой – четное число». Требуется повернуть только те карты, которые, по мнению испытуемого, необходимы и достаточны для верификации гипотезы. Правильное решение – повернуть *A* и 7. Однако большинство поворачивают либо только *A*, либо *A* и 4. Поворот *A* – это поиск подтверждения гипотезы; это делают все. Поворот 7 – это поиск опровержения гипотезы. Поворот 4 ничего не дает для верификации гипотезы: как известно в логике, при истинном *q* («число четно») импликация «если *p* то *q*» истинна для любого *p*, т.е. любая буква на обороте не опровергнет и не подтвердит гипотезу.

Неспособность к вероятностному мышлению. Вероятностное мышление человеку несвойственно. Люди склонны искать закономерности даже там, где их нет. При взгляде на последовательности рождений мальчиков и девочек в роддоме МММДДД, ДДДДДД, МДММДМ большинство испытуемых только третью последовательность склонны считать случайной, хотя тем, кто знаком с теорией вероятностей, ясно, что все три последовательности равновероятны.

А вот еще более выразительный пример, приведенный в [8]. Исследование частоты рака почки, проведенное более чем в 3000 округах США, показало, что самый низкий уровень заболеваемости обнаружен в сельских малонаселенных округах, расположенных в традиционно республиканских штатах на Среднем Западе, Юге и Западе. Самый высокий уровень заболеваемости обнаружен тоже в сельских малонаселенных округах, расположенных в традиционно республиканских штатах на Среднем Западе, Юге и Западе.

Ясно, что каузального объяснения здесь нет: не могут противоречащие друг другу факты иметь одну и ту же причину. Ключевое слово в примере – малонаселенность. Дело в том, что, чем меньше выборка, тем выше в ней вероят-

ность появления экстремальных значений. Например, при одновременном бросании четырех монет возможны $2^4 = 16$ комбинаций «орел-решка»; вероятность появления экстремумов, т.е. комбинаций OOOO или RPPP равна $2/16 = 1/8$. При бросании 10 монет эта вероятность равна $2/2^{10} = 1/512$. Этот факт говорит о том, что результаты экспериментов, полученные на малых выборках, могут сильно отличаться от результатов на больших выборках. Пренебрежение величиной выборки в исследованиях, ложная уверенность в том, что «закон больших чисел с тем же успехом применим и к малым», Канеман и Тверски иронически называли «Законом малых чисел».

Одна из типичных ошибок при оценке вероятностных ситуаций – игнорирование контрольной группы. Вот пример из [9].

*200 человек получили лечение,
им стало лучше.*

*75 человек получили лечение,
им не стало лучше.*

*50 человек не получили лечение,
им стало лучше*

*15 человек не получили лечение,
им не стало лучше*

Дало ли лечение эффект?

Обычный ответ – «да, дало». Он учитывает только результаты получивших лечение и не замечает, что у не получивших лечения (в контрольной группе) результат практически тот же самый.

Особенно сильные отклонения от реальности происходят при оценках, связанных с учетом априорной вероятности, т.е. требующих использовать правило Байеса. Пример:

Вирус XYZ вызывает серьезное заболевание, которому подвержен 1 человек из 1000. Существует анализ, который указывает носителя этого вируса, но в 5% случаев ошибочно указывает, что человек – носитель вируса. Анализ конкретного человека показал, что он – носитель. Какова вероятность того, что он действительно носитель?

Обычный ответ – 95%. Правильный ответ (по правилу Байеса) – 2%: если провести анализ 999 из 1000, у которых вируса нет, то некорректный анализ получают 50 человек. Таким образом, из 51 пациента вирус есть только у одного ~ 2%.

Дефективные (ложные и вредоносные) умственные программы. Одной из распростра-

ненных причин нерационального поведения являются различные дефективные схемы поведения (Станович [9] называет их зараженными программами). К ним относятся мошеннические схемы (финансовые пирамиды, схемы быстрого обогащения, схемы ухода от налогов и др.), лженауки и псевдонаучные мифы (астрология, спиритизм, снежный человек, НЛО и уфология), *мемы*.

Термин «мем» был предложен Докинсом в книге [18] и очень быстро получил широкое распространение. Мем – это набор слов и словосочетаний, за которыми стоит комплекс взаимосвязанных идей и убеждений (мемплекс), являющийся элементом культуры (или субкультуры). Примеры – «демократия», «права человека», «растленный Запад». Возникло целое направление в гуманитарных науках, посвященное исследованию мемов – меметика. Это направление изучает законы распространения мемов в различных обществах и причины, по которым одни мемы чрезвычайно живучи, а другие не приживаются. Меметика подвергает сомнению традиционные представления о том, что человек всегда сам сознательно решает, каких убеждений ему придерживаться. Распространенность определенных мемов во многом объясняется их внутренними свойствами, использующими те или иные особенности конкретной культуры. Мемы отнюдь не всегда полезны. Существуют паразитические мемы, которые работают не на помощь своему носителю, а на продление своего жизненного цикла. Поэтому важно избегать программ, не поддающихся оценке и опровержению, например, программ, опирающихся на слепую веру. Особенно опасны мемы-разрушители: агрессивные программы, делающие культурную среду враждебной для конкурирующих мемов; мемы различных сект и агрессивных версий ислама, мемы мученичества и вознаграждения в загробной жизни.

7. О роли эмоций при принятии решений

Общепринято считать, что решения надо принимать без участия эмоций, «с холодной головой», поскольку они нарушают нормальный ход рассуждений, а зачастую заменяют их: вместо ответа на сложный вопрос (что я об этом думаю?) выдается ответ на более легкий вопрос (что я в

связи с этим чувствую?). Традиция разделения разума и чувств идет от Платона и Аристотеля и окончательно закреплена Декартом. Теория ограниченной рациональности в изложении Канемана и Становича также склонна смотреть на эмоции как на помеху рассуждениям (см. выше «влияние аффекта»). Однако исследования последних десятилетий говорят о том, что отсутствие эмоций также отрицательно влияет на процесс принятия решений.

Ярким примером служит «случай Эллиота», пациента невролога А. Дамасио, описанный в [19]. Эллиоту удалили небольшую опухоль в лобной доле коры головного мозга. До операции он занимал руководящую должность в крупной корпорации. После операции он потерял способность принимать даже самые простые решения, бесконечно размышляя над тем, писать синей ручкой или черной, что выбрать на обед в ресторане и т.д. При том, что его IQ остался высоким – т.е. его способность рассуждать не была нарушена, его нерешительность стала патологической, что, в конечном счете, разрушило его жизнь: его уволили с работы, от него ушла жена, все его начинания оказывались неудачными⁴. Дамасио обратил внимание на то, что патологическая нерешительность Эллиота сопровождается патологическим хладнокровием. О драматических событиях своей жизни он рассказывал бесстрастно, как посторонний наблюдатель. При демонстрации фотографий различных ужасных случаев, которые должны вызывать сильные негативные эмоции, Эллиот оставался совершенно спокойным. Стандартная кожно-гальваническая реакция, автоматически возникающая при эмоциональных возбуждениях, у него отсутствовала.

Исследование этого и других аналогичных случаев привело Дамасио к концепции, утверждающей важную роль эмоций в принятии решений и подробно изложенной в книге [19]. Эту концепцию здесь мы изложим следующим образом.

1. Чисто рациональный подход к принятию решений, как правило, связан с рассмотрением слишком большого числа факторов, которые надо принимать во внимание, и различных вариантов последствий каждого из возможных решений. Учитывая ограниченный объем рабо-

чей памяти человека, этот подход либо не даст никакого решения, либо затянется на слишком большое время.

2. Эмоция, возникающая при рассмотрении некоторой альтернативы в процессе принятия решений, является ее грубой оценкой – положительной, если эмоция положительная, и отрицательной в противном случае. Тем самым она является сигналом тревоги или одобрения, т.е. дополнительной информацией, используемой при рассмотрении вариантов. Поскольку эмоции связаны с определенными состояниями тела (помимо кожно-гальванической реакции, может участиться сердцебиение, повыситься кровяное давление и т.д.) Дамасио назвал такие эмоции-подсказки *соматическими маркерами*.

3. Эмоции делятся на первичные и вторичные. Первичные эмоции – эволюционно наиболее старые. Они связаны с биологическими инстинктами выживания (связанными в первую очередь с угрозами жизни и вызывающими страх, испуг, боль) и закреплены генетически, т.е. являются врожденными. Неврологически они реализуются лимбической системой мозга. Вторичные эмоции формируются в результате личного социального опыта. Повреждения в лимбической системе нарушают выработку первичных эмоций; повреждения в префронтальной коре нарушают выработку вторичных эмоций. У пациентов типа Эллиота отсутствуют именно вторичные, «социальные» эмоции.

4. Хотя разные типы эмоций формируются в разных частях мозга, врожденные механизмы участвуют в образовании схем, фиксирующих личный опыт. «Записи опыта, и ответы на них, если они должны быть адаптивными, должны оцениваться и формироваться фундаментальным набором предпочтений организма, которые считают выживание первостепенным. Поскольку оценка и формирование жизненно важны для существования организма, гены также предписывают, что врожденные схемы должны серьезно влиять на все множество схем, которые могут быть модифицированы опытом». [19]. Поэтому «аппарат рациональности, традиционно предполагаемый в неокортексе, по-видимому, не работает без аппарата биологической регуляции, который традиционно считается подкорковым. Природа, похоже, построила аппарат рациональности не просто над аппаратом биологической регуляции, но также из него и вместе с ним. Механизмы поведения,

⁴ Подробно случай Эллиота описан также в популярных книгах [20,21].

выходящего за пределы побуждений и инстинктов, используют и подвалы, и верхние этажи: неокортекс вовлекает старый мозг в свою деятельность, и рациональность возникает из их совместной активности» [19].

5. Соматические маркеры могут действовать на сознательном уровне, либо явно вмешиваясь в процесс рассуждений путем оценки конкретного варианта, либо в виде сигнала неопределенной тревоги (упомянутого выше ощущения «что-то не так»). При этом они мобилизуют внимание и рабочую память. Но они могут действовать скрыто, бессознательно: «вместо проявления воспринимаемых изменений телесных состояний они могут тормозить регуляторные нервные схемы, находящиеся в коре, которые управляют выбором поведения. Этот скрытый механизм – источник того, что мы называем *интуицией*» [19]. Фактически соматические маркеры являются проявлением скрытых схем, возникших в результате накопленного опыта. Подсознательность многих таких схем – причина того, почему человек не всегда может объяснить, как он принял то или иное решение.

6. Подсказки соматических маркеров могут быть источником ошибок. Они могут придавать излишний вес деталям, которые в действительности несущественны, или, при сильной эмоции, «выбить из колеи», лишив на время способности хладнокровно рассуждать. Поэтому последнее слово остается за рациональным мышлением.

7. Роль эмоций и, в частности, соматических маркеров в процессе принятия решений говорит об «ошибке Декарта»: в когнитивных процессах участвует не только мозг, но и тело. «Тело неизбежно реагирует как на внешние воздействия, так и на наши действия и размышления. Поэтому разум немыслим без некоторой воплощенности (embodiment)» [19]. Используя известный образ советской фантастики, можно сказать, что «голова профессора Доуэля» невозможна. Дамасио отмечает, что этот же принцип «воплощенности разума» высказывался Дж. Лакоффом [12, 22], Э. Рош и др.

На наш взгляд, теория Дамасио естественно дополняет концепции Канемана-Тверски-Становича, если считать, что соматические маркеры – это инструмент Системы 1, служащий сигналом для Системы 2 и, как уже было указано, мобилизующий внимание и рабочую память.

Заключение

Исследования ограниченной рациональности за последние 35-40 лет радикально изменили представления о когнитивных механизмах принятия решений. Традиционные постулаты теории принятия решений, исходящие из того, что человек, оценивая ситуацию и принимая решения, ведет себя предельно рационально, т.е. стремится к максимизации своей функции полезности, подверглись серьезному пересмотру. Теория ограниченной рациональности показала, что человек часто поступает нерационально не только в силу объективных причин (необходимость быстро принимать решения в условиях недостаточной информации, ограниченного времени и т.д.), но и из-за неумения использовать те интеллектуальные способности, которыми он обладает. Это объясняется особенностями взаимодействия двух типов когнитивных процессов – быстрых (Система 1), действующих при минимальном контроле сознания, и медленных (Система 2), сознательных и требующих когнитивных усилий. Постоянно действующая Система 1 обычно принимает стандартные решения, основанные на прошлом опыте. В нестандартных ситуациях включается рассуждающая и оценивающая Система 2. Однако она не всегда использует свои средства в достаточном объеме из-за своей «когнитивной скупости». Система 2 склонна к минимизации когнитивных усилий.

Эти же особенности взаимодействия двух систем объясняют существенные различия между рациональностью и интеллектом. Интеллект, измеряемый тестами IQ, – это набор когнитивных инструментов; рациональность – это умение их применять. Именно рациональность играет определяющую роль в успешности принятия решений.

Новая теория уже находит реальные приложения. В 2008 г. вышла книга [23], в которой авторы излагают идею «подталкивания» – набора приемов, помогающих людям избегать типичных ошибок при достижении своих целей. Эта идея основана на двух основных принципах: 1) выгодный вариант по возможности должен быть легким; 2) выгодность варианта должна быть обоснована накопленной статистикой, расчетами или научно обоснованным экспериментом. Характерный пример такого приема – выбор опции по умолчанию, который

особенно эффективен при бинарном выборе. Если человеку предстоит выбор между альтернативами (опциями) А и В, причем для выбора В надо проделать некоторые процедуры (заполнить большую анкету, принести некоторое количество справок и т.д.), а в случае отказа от этих процедур автоматически (по умолчанию) реализуется альтернатива А, то часто человек в силу когнитивной (или обыкновенной) лени выбирает А, даже если при некотором размышлении В окажется более выгодной. Типичным примером такого выбора является выбор «оставить все как есть» [24], потому что обычно по умолчанию реализуется именно он. Задача подталкивания заключается в том, чтобы сделать опцией по умолчанию выгодную для пользователя опцию. В последней главе книги [7] описывается опыт реализации идеи подталкивания командой специалистов, работавших консультантами при правительстве Великобритании.

Помимо полезных прикладных рекомендаций по принятию решений, вытекающих из теории ограниченной рациональности, особое внимание следует обратить на опасности, о которых говорит эта теория. Они связаны с тем, что ограничения рациональности дают серьезные возможности для манипулирования сознанием путем использования принципа WYSIATI, фрейминга, привязки и других эффектов, связанных с когнитивной скупостью. Да и методы подталкивания в руках недобросовестных людей могут оказаться эффективным инструментом манипуляций.

«Когнитивный скряга, буквально говоря, думает не своей головой. Мозг его обрабатывает те стимулы, которые преподнесены ярче, наиболее легкоусвояемые факты или самые сильные сигналы. Когнитивный скряга легко поддается манипуляциям тех, кто управляет ярлыками, контролирует яркость изображения и держит в руках якоря» [9]. Это активно используется различными социальными манипуляторами – от создателей рекламы до политической демагогии и информационных войн.

Литература

1. Simon H.A. Models of Man / N.Y.: Wiley, 1957.
2. Simon H.A. Rationality as Process and as Product of Thought. Richard T.Ely Lecture // American Economic Review, May 1978, v.68, no.2, p.1–16.
3. Tversky A., Kahneman D., Judgement Under Uncertainty: Heuristics and Biases // Science, 1974, No.185, pp. 1124–1131.
4. Kahneman D., Tversky A. Prospect Theory: An Analysis of Decision under Risk // Econometrica, 1979, v. XLVII, pp. 263–291.
5. Larichev O.I. Psychological Validation of Decision Methods // Journal of Applied Systems Analysis, 1984, v. 11, pp. 37–46.
6. Ларичев О.И. Психологические обоснования методов принятия решений // Труды конференции «Психология и автоматизация организационного управления». Часть I / Под ред. Е.Д. Жаркова, В.Е. Лепского. — Одесса: Научный совет по комплексной проблеме «Кибернетика» при Президиуме АН СССР. 1985, с. 39–59.
7. Thaler R.H. Misbehaving: The making of behavioral economics / New York: Norton. 2015. (Русский перевод: Талер Р. Новая поведенческая экономика. Почему люди нарушают правила традиционной экономики и как на этом заработать / М.: Эксмо, 2017).
8. Kahneman D. Thinking, fast and slow. / Farrar, Straus and Giroux, 2011. (Русский перевод: Канеман Д. Думай медленно... решай быстро / М.: АСТ, 2013).
9. Stanovich, Keith E. What Intelligence Tests Miss: The Psychology of Rational Thought / Yale University Press, 2009. (Русский перевод: Станович К. Рациональное мышление / М.: Карьера Пресс, 2012).
10. Rizzolatti Giacomo, Sinigaglia Corrado. Mirrors in the brain — how our minds share actions and emotions. / Oxford University Press, 2008. (Русский перевод: Дж.Риццоллатти, К.Синигалья. Зеркала в мозге. / М.: Языки славянских культур, 2012).
11. Rosch E. Cognitive representations of semantic categories // Journal of Experimental Psychology, 1975. 104, pp.192–233.
12. Lakoff J. Women, Fire, and Dangerous Things: What Categories Reveal About the Mind / University of Chicago Press, 1987. (Русский перевод: Лакофф Д. Женщины, огонь и опасные вещи: Что категории языка говорят нам о мышлении / М. Гнозис, 2011).
13. Кузнецов О.П. Когнитивная семантика и искусственный интеллект // Искусственный интеллект и принятие решений. 2012, №4, с.32–42.
14. Hawkins Jeff. On Intelligence. — New York. Henry Holt and Company, 2005. (Русский перевод: Хокинс Джефф. Об интеллекте. М.: Вильямс, 2007).
15. Фаликман М.В., Койфман А.Я. Виды прайминга в исследованиях восприятия и перцептивного внимания // Вестник МГУ. Серия 14. Психология. 2005, №3, с.86–97, №4, с. 81–90.
16. Кузнецов О.П. О концептуальной семантике // Искусственный интеллект и принятие решений, 2014, №3, с.32–39.
17. Wason P.C. Reasoning about a rule // Quarterly Journal of Experimental Psychology, 1968, v.20, p.273–281.
18. Dawkins R. The selfish gene / New York: Oxford University Press. 1976.
19. Damasio, Antonio R. Descartes' error: emotion, reason, and the human brain / Putnam Publishing, 1994.
20. Lehrer J. How we decide / New York: Houghton Mifflin Harcourt, 2009. (Русский перевод: Лерер Дж. Как мы принимаем решения. М.: Астрель, 2011).

21. Kean S. The Tale of the Duelling Neurosurgeons / New York. Little, Brown and Co., 2014. (Русский перевод: Кин С. Дуэль нейрохирургов / М.: Эксмо, 2015).
22. Lakoff G., Johnson M. Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought / Basic Books, 1999.
23. Thaler R.H., Sunstein C.R. Nudge: Improving Decisions about Health, Wealth, and Happiness. / Yale University Press, 2008.
24. Kahneman D., Knetsch J.I., Thaler R.H. The endowment effect, loss aversion, and status quo bias. Journal of Economic Perspectives, 1991, v.5, 193-206.

Bounded rationality and decision making

O. P. Kuznetsov

V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Science, Moscow, Russia

Abstract. This paper considers the theory of bounded rationality, also known as behavioral economics developed by D.Kahneman, A.Tversky, R.Thaler and K.Stanovich. This theory demonstrates that human beings often act nonrationally not only due to objective reasons, such as the necessity of decision making under uncertainty and constraints on memory, but also due to inability to make use of their own intellectual tools. This is explained by how fast and slow types of cognitive processes interact. The fast system (System 1) is implemented with minimal control of consciousness and the slow system (System 2) involves consciousness and cognitive efforts. System 1 is working constantly and is making routine decisions while System 2 is involved in non-standard situations. However, System 2 minimizes efforts and does not always fully use its reasoning and estimating abilities due to ‘cognitive miserliness’. The role of emotions in decision making as a concept introduced by the prominent neurobiologist A. Damasio contributes largely to this theory.

Keywords: Bounded rationality, decision making, behavioral economics, fast and slow cognitive processes, prospect theory, framing.

DOI 10.14357/20718594190101

References

1. Simon H.A. Models of Man. / N.Y.: Wiley, 1957.
2. Simon H.A. Rationality as Process and as Product of Thought. Richard T.Ely Lecture // American Economic Review, May 1978, v.68, no.2, p.1–16.
3. Tversky A., Kahneman D., Judgement Under Uncertainty: Heuristics and Biases. // Science, 1974, No.185, pp. 1124-1131.
4. Kahneman D., Tversky A. Prospect Theory: An Analysis of Decision under Risk. // Econometrica, 1979, v. XLVII, pp. 263—291.
5. Larichev O.I. Psychological Validation of Decision Methods. // Journal of Applied Systems Analysis, 1984, v. 11, pp. 37–46.
6. Larichev O.I. Psichologicheskie obosnovaniya metodov prinyatiya reshenii (Psychological Validation of Decision Methods) // Trudy konferencii «Psichologia i avtomatizatsia organizatsionnogo upravleniya». Chast' I / E.D.Zarkov, V.E.Lepskii (ed.) — Odessa: Nauchnyi Sovet po kompleksnoi probleme «Kibernetika» pri Prezidiume AN SSSR. 1985, s. 39–59.
7. Thaler R.H. Misbehaving: The making of behavioral economics. / New York: Norton. 2015.
8. Kahneman D. Thinking, fast and slow. / Farrar, Straus and Giroux, 2011.
9. Stanovich, Keith E. What Intelligence Tests Miss: The Psychology of Rational Thought. / Yale University Press, 2009.
10. Rizzolatti Giacomo, Sinigaglia Corrado. Mirrors in the brain — how our minds share actions and emotions. / Oxford University Press, 2008.
11. Rosch E. Cognitive representations of semantic categories. // Journal of Experimental Psychology, 1975. 104, pp.192-233.
12. Lakoff J. Women, Fire, and Dangerous Things: What Categories Reveal About the Mind. / University of Chicago Press, 1987.
13. Kuznetsov O.P. Cognitive Semantics and Artificial Intelligence// Scientific and Technical Information Processing, 2013, Vol. 40, No. 5, pp. 269–276.
14. Hawkins Jeff. On Intelligence. — New York. Henry Holt and Company, 2005.
15. Falikman M.V., Koifman A.Ya. Vidy praiminga v issledovaniyah vospriyatiya i vnimaniya (Species of priming in the investigations of perception and attention). // Vestnik MGU. Seria 14. Psihologia. 2005, №3, 86-97, №4, 81-90.
16. Kuznetsov O.P. Conceptual Semantics. // Scientific and Technical Information Processing, 2015, Vol. 42, No. 5, pp. 307–312.
17. Wason P.C. Reasoning about a rule. // Quarterly Journal of Experimental Psychology, 1968, v.20, p.273-281.
18. Dawkins R. The selfish gene. / New York: Oxford University Press. 1976.

19. Damasio, Antonio R. Descartes' error: emotion, reason, and the human brain. / Putnam Publishing, 1994.
20. Lehrer J. How we decide. / New York: Houghton Mifflin Harcourt, 2009.
21. Kean S. The Tale of the Duelling Neurosurgeons. / New York. Little, Brown and Co., 2014.
22. Lakoff G., Johnson M. Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought. / Basic Books, 1999.
23. Thaler R.H., Sunstein C.R. Nudge: Improving Decisions about Health, Wealth, and Happiness. / Yale University Press, 2008.
24. Kahneman D., Knetsch J.I., Thaler R.H. The endowment effect, loss aversion, and status quo bias. Journal of Economic Perspectives, 1991, v.5, 193-206.