



# НЕЙРОЭКОНОМИКА: НЕЙРОБИОЛОГИЯ ПРИНЯТИЯ РЕШЕНИЙ<sup>1</sup>

**КЛЮЧАРЕВ В. А.**, *Университет Базеля, Швейцария, Санкт-Петербургский государственный университет, Санкт-Петербург*

**ШМИДС А.**, *Эразмус Университет Роттердама, Нидерланды*

**ШЕСТАКОВА А. Н.**, *Московский городской психолого-педагогический университет, Москва  
Санкт-Петербургский государственный университет, Санкт-Петербург*

Происходящий в настоящее время синтез экономических наук, психологии и нейробиологии в рамках новой дисциплины – нейроэкономики, а также развитие новейших нейроимиджинговых методов обещают появление новой междисциплинарной модели, объясняющей механизмы принятия решений. Результаты нейроэкономических исследований демонстрируют, что принятие решений обусловлено работой относительно независимых нейрональных систем мозга, отобранных посредством естественного отбора. Взаимодействие «когнитивных» и «эмоциональных» нейрональных механизмов принятия решений определяет степень рациональности поведения человека.

**Ключевые слова:** нейроэкономика, нейробиология, принятие решений, социальное влияние, временное дисконтирование, моральная дилемма, функциональная магнитно-резонансная томография.

Большинство из нас считает привычку задумываться над последствиями наших поступков ужасно скучной... порой до смерти.

*Бертран Рассел*

## Введение

В 1961 году больной эпилепсией W. J. перенес знаменитую операцию по рассечению мозолистого тела (комиссуры, соединяющей большие полушария мозга), приведшую к разделению мозга на два независимых полушария (Gazzaniga, 2005). В поведении пациента появились необычные особенности: иногда его руки соревновались друг с другом при выполнении заданий. Например, когда испытуемому нужно было, следуя инструкции, собрать предложенную конструкцию, используя набор цветных кубиков, его правая рука настойчиво старалась изменить комбинацию, которую до этого собрала левая. В медицинской практике известен аналогичный феномен, называемый «синдром чужой руки», – редкое заболевание, встречающееся, например, у пациентов с нейродегенеративными нарушениями мозга и выражающееся в непроизвольной конфликтной активности рук, словно движениями пациента управляет внешняя сила. Так, одному такому пациенту приходилось всякий раз тормозить машину, когда его левая рука непроизвольно и хаотически начинала хвататься за руль во время движения автомобиля; другая пациентка восклицала, испугавшись непроизвольных действий своей руки: «Остановите это чудовище! Эта рука убьет меня!» (Leiguarda et al., 1993). Приведенные выше любопытные клинические наблюдения наиболее ярко иллюстрируют господствующие в нейробиологии взгляды на поведение человека и механиз-

<sup>1</sup>Работа выполнена при финансовой поддержке федеральной целевой программы «Научные и научно-педагогические кадры инновационной России», ГК 02.740.11.5233, ГК 14.740.11.0232, ГК 02.740.11.5148.

мы принятия решений, в основе которых лежит представление о независимых нейрональных системах мозга, отобранных посредством естественного отбора и существующих параллельно в нервной системе человека.

Достижения нейробиологии последних лет позволили приблизиться к пониманию биологических основ принятия решений. Сравнительно недавно появилась новая междисциплинарная область нейробиологии – нейроэкономика (Glimcher and Rustichini, 2004). Она объединила нейробиологию, экономические науки, психологию и ряд других дисциплин, знание которых легло в основу новых представлений о механизмах принятия решений и позволило успешно моделировать поведение человека и животных. Возникшая на стыке наук, нейроэкономика в широком смысле является нейробиологией принятия решений (decision neuroscience) (Rilling et al., 2008).

На первом этапе своего развития нейроэкономические исследования сосредоточились на изучении нейробиологических механизмов простейших (перцептивных) решений, а также на природе иррациональности, например, на причинах более сильной эмоциональной реакции, сопровождающей потери по сравнению с приобретениями (loss aversion) (Tom et al., 2007). Интенсивно изучалось также *временное дисконтирование* (temporal discounting) – непропорциональное предпочтение сиюминутного вознаграждения относительно отложенного во времени (McClure et al., 2004). Позднее в орбиту нейроэкономических изысканий попали исследования мозговых механизмов определения субъективной полезности (subjective utility) в процессе выбора из возможных альтернатив, исследования того, как мозг учитывает риски при принятии финансовых решений, и исследования поведения потребителя (Knutson et al., 2005; Kuhnen, Knutson, 2005), включая восприятие цены и рекламы (Klucharev et al., 2008; Plassmann et al., 2008). К сожалению, в настоящем обзоре не представляется возможным охватить все направления исследований нейроэкономики. Поэтому в первую очередь мы сосредоточимся на исследованиях дуалистичности механизмов принятия решений. Наша цель – показать, что нерациональное поведение, т. е. поведение, не укладывающееся в традиционные нормативные экономические теории, является результатом эволюционного отбора, закрепленного в структуре и функциях нейрональных сетей нашего мозга.

### Нейроэкономика как новая дисциплина

Исходя из современных представлений нейроэкономики, принятие решений и как следствие – выбор оптимального поведения происходят на уровне специализированных нейронных сетей. В целом нейроэкономический подход является относительно механистичным: нейронные сети обладают способностью регулировать оценку всех «за» и «против» при выборе того или иного поведения. Согласно данному подходу, нейрон, или нейронная сеть, принимающая решение, аккумулирует информацию о возможных поведенческих альтернативах и делает выбор в пользу наиболее оптимальной. В первых нейроэкономических исследованиях, проводимых на животных, были использованы простые модели принятия решений. Например, обезьян обучали следить взглядом за передвигающимся в том или ином направлении объектом (Sugrue et al., 2005). Экспериментаторы пытались обнаружить нейроны, отвечающие за принятие элементарного решения – перевести взгляд вправо или влево. На экране перед обезьяной появлялось облако из движущихся точек, когерентностью движения которых манипулировал экспериментатор. Иногда движение было абсолютно случайным, а иногда точки начинали двигаться в одном направлении (вправо или



влево). Результаты исследований показали, что обезьяны, стимулируемые за правильное выполнение задачи фруктовым соком, способны быстро выявлять доминирующее направление движения точек и следовать за ними взглядом. Увеличивая или уменьшая количество точек, двигающихся в одном направлении, исследователи манипулировали степенью сложности задачи и продемонстрировали, что при выполнении задания активность нейронов внутритеменной борозды (*lateral intraparietal area, LIP*) постепенно возрастала, нейроны аккумулировали информацию до определенного *порога принятия решения* – переход через данный порог запускал движение глаз. Нейробиологи предложили относительно простую модель принятия решения: обработка и хранение информации об альтернативных направлениях движения происходят на уровне специализированных нейронов-детекторов, кодирующих движение зрительных изображений влево или вправо независимо друг от друга. Например, чем больше точек двигается влево, тем более активным становятся нейроны, детектирующие движение влево. А теперь представим себе, что эти детекторы оказывают друг на друга тормозящее (подавляющее) воздействие: в результате наиболее активные нейроны, к примеру, специализирующиеся на движении влево, будут тормозить (подавлять) нейроны, кодирующие движение в противоположную сторону (вправо), активность которых меньше в данный момент времени. Таким образом, мы имеем дело с элементарным алгоритмом сравнения альтернатив, где выживает и выигрывает сильнейший за счет подавления конкурента. В целом «левосторонние» нейроны запускают команду перевода взгляда влево, сдерживая (подавляя, затормаживая) при этом использование альтернативных вариантов. В данной модели нейроны LIP являются субстратом принятия решений, осуществляя сравнение зрительной информации и выбор в пользу одной из альтернатив. В последние годы было предложено несколько моделей нейронных сетей, обеспечивающих принятие решения (Bogacz, 2007), в основе которых лежит представление о том, что решение генерируется в тот момент, когда разница между альтернативами достигает определенной пороговой величины. В экспериментах с обезьянами (Glimcher, 2003) было показано, что принятие решения можно уверенно предсказать исходя из активности нейронов LIP: после того как активность нейрона, кодирующего принятие решений, достигает порогового уровня, данное решение неотвратимо. Более того, оказывая воздействие на определенные нейроны, например, электрическим разрядом, можно повлиять на принятое обезьяной решение. Таким образом впервые удалось продемонстрировать существование нейронов, кодирующих принятие элементарных решений.

В экспериментах на животных наградой, как правило, является пищевое подкрепление. Интересно, что величина подкрепления пропорциональна активности нейронов LIP. Более того, активность нейронов LIP отражает одновременно и величину ожидаемого подкрепления, и вероятность его получения. Таким образом, нейроэкономические исследования вплотную приблизились к концепции *ожидаемой полезности* альтернативы (*expected utility*) – одному из ключевых представлений в экономической теории. Согласно данной концепции, среди всех возможных альтернатив или сценариев поведения оптимальный выбор должен быть сделан в пользу альтернативы с наиболее высокой ожидаемой полезностью. Действительно, как было показано в нейроэкономических экспериментах, в активности нейронов LIP наблюдаются те же самые закономерности, которые могут быть предсказаны классической экономической теорией. Конечно, упрощенное представление о том, что только нейроны LIP активно вовлечены в процесс принятия решений, было бы неверным. Целый ряд других областей мозга также вовлечен в процесс оценки ожидаемой полезности.

сти, и к ним в первую очередь относятся так называемые дофаминергические области, богатые нейромедиатором дофамином. У человека базальные ганглии (striatum) и нижние области лобной коры (orbitofrontal cortex) играют ключевую роль в оценке полезности выбора в момент принятия решения и таким образом управляют нашими решениями (Rangel et al., 2008).

Нейроэкономика предлагает свою достаточно механистическую модель оценки мозгом субъективной полезности, в соответствии с которой нейронные сети обладают способностью сравнения имеющихся альтернатив. В такой модели на вход нейронной сети поступает сенсорный, мотивационный, когнитивный или любой другой сигнал, а на выходе мы получаем результат сравнения в пользу наиболее оптимального решения. Такая модель получила несколько нелогичное название *диффузной* (термин «диффузный» происходит от одноименной статистической концепции). В настоящее время эта простая модель считается базовой нейроэкономической нейронной моделью теории принятия решения (рис. 1).

Если рассматривать проблему принятия решений с точки зрения нейроэкономики и предположить, что наши решения могут быть предсказаны исходя из активности определенных нейронных сетей, то раскрытие соответствующих нейронных механизмов откроет новые горизонты в понимании природы рационального поведения человека.

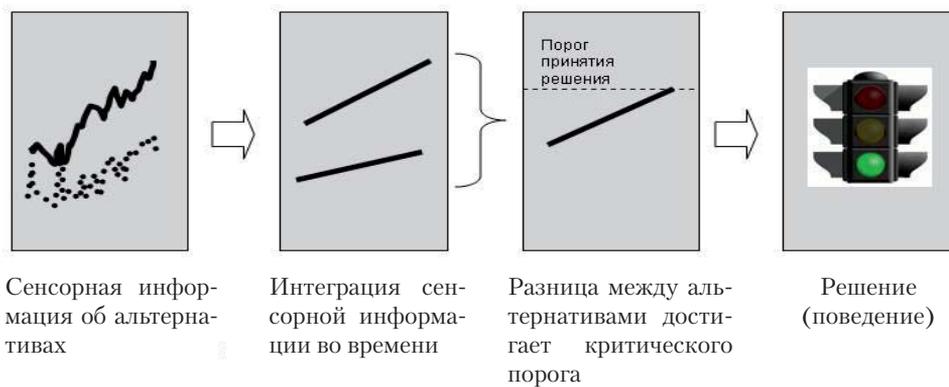


Рис. 1. Схематическое представление последовательности процессов, лежащих в основе принятия элементарных (перцептивных) решений (адаптировано с изменениями; см.: Vogacz, 2007). На первой стадии в мозг поступают сенсорные сигналы с кодированной информацией об имеющихся альтернативах; линиями схематично обозначена динамика активности сенсорных нейронов. Сначала сенсорная информация концентрируется вокруг одной из двух альтернатив. Так как сенсорный вход нервной системы характеризуется высоким уровнем шума, то порой, в определенные моменты времени, предпочтительней может оказаться вторая альтернатива, однако шум, очевидно, затрудняет операцию выбора. На второй стадии происходит интегрирование информации в течение определенного периода времени. От стадии к стадии шум уменьшается. На третьей стадии происходит извлечение информации и сравнение ее с выбранным критерием. Весь процесс похож на работу светофора на перекрестке, регулирующего движение: система находится в ожидании получения достаточного количества информации для принятия окончательного решения или продолжения дальнейшего интегрирования до тех пор, пока не накопится необходимое количество аргументов в пользу одной из альтернатив



## Общие принципы нейробиологической теории принятия решений

Как было показано выше, первые исследования механизмов принятия решений касались самых простых ситуаций и проводились с участием животных (обезьяны должны были следить за движущимися на экране дисплея объектами), в то время как в реальной жизни человек вынужден принимать решения под влиянием гораздо большего количества факторов и в более сложном контексте.

Для того чтобы разобраться, как происходит принятие решений, целесообразно разделить этот процесс на несколько этапов (Rangel et al., 2008). На первом этапе происходит формулирование задачи как таковой, формируется представление о цели и контексте решения (1). Здесь интегрируется информация о внутреннем состоянии организма и факторах окружающей среды, к примеру, голода или уровня угрозы, в контексте будущего действия. На следующем этапе определяется ценность (value или valuation) выбора той или иной поведенческой альтернативы (2). На третьем этапе сравниваются альтернативные варианты решения и происходит выбор наилучшего (action selection) (3). После осуществления выбранного действия происходит обработка его результатов и оценка эффективности (4). На последнем этапе происходит обучение, т. е. обновление хранящейся в памяти информации, с тем чтобы все последующие действия выполнялись с наибольшей эффективностью (5).

Нейроэкономика в первую очередь изучает процессы, происходящие на этапе определения субъективной ценности альтернатив (второй этап в описанной выше иерархии процессов) как наиболее значимой стадии, ведь именно на ней происходит *оценка* и *выбор* действия, результат которого приведет к достижению индивидом наибольшей выгоды. Классическая нейробиология убедительно продемонстрировала, что существуют параллельные системы оценки ценности поведения и, более того, что само поведение зависит от того, какая система оценки ценности используется в текущий момент времени (здесь и далее мы используем «субъективную ценность» и «полезность» как синонимы). Несмотря на то, что трудно оценить точное количество подобных параллельных систем оценки полезности, стоит выделить и обсудить наиболее важные.

В первую очередь необходимо упомянуть классический *павловский условный рефлекс*, а также концепцию *целенаправленного поведения* (goal-directed behavior). Упрощая, можно сказать, что классическая павловская схема «стимул – реакция» (по крайней мере, в той ее западной интерпретации, которая широко используется в современной нейроэкономической науке) в большей степени характеризуется автоматичностью процессов, тогда как целенаправленное поведение лежит в основе более осознанного целеустремленного поведения.

В павловской рефлекторной системе наибольшей ценностью обладает ограниченное количество автоматических поведенческих актов, сформировавшихся в процессе эволюции и адаптированных к окружающей среде, типичным примером которых является пищевое поведение – слюноотделение и выделение желудочного сока при посещении ресторана. Области мозга, обслуживающие так называемую павловскую систему, включают миндалину (amygdala) и нижние области стриатума (ventral striatum), т. е. те структуры, которые отвечают за выполнение автоматических поведенческих рефлекторных актов по типу «стимул – реакция». В данной системе адаптивная ценность поведения закодирована и закреплена генетически.

Целенаправленное поведение, наоборот, характеризуется большей пластичностью. По определению данное поведение направлено на достижение конкретной цели: выбор,

скажем, музея для посещения – пример подобного целенаправленного поведения; причем чем чаще выполняется то или иное *целенаправленное* действие, тем вероятнее оно станет привычкой, т.е. действие приобретет независимость от ценности его результата. В итоге вырабатывается поведение с особой оценкой полезности – *автоматизм*, не требующий достижения цели. Данный механизм может быть рассмотрен подробнее на следующем примере: для нашего организма такое привычное (автоматическое) действие, как выпить кофе в рабочий перерыв, обладает высокой ценностью независимо от того, хотим мы в действительности взбодриться в данный момент или нет. Однако при изменении условий в цепочке «действие – награда» такая система может дать сбой, результатом чего становятся ошибки в предсказании ценности наших решений. Представьте, что в кафе, где вы привыкли перекусывать, качество обслуживания резко ухудшилось. Рациональная часть вашего Я будет подталкивать вас к поиску нового места, однако вполне вероятно, что в силу привычки вы продолжите посещать любимое заведение.

Считается, что в основе работы разных систем принятия решений лежит работа независимых сетей головного мозга. Поведение, связанное с автоматизмами, обусловлено активностью в латеральной области стриатума (*dorsolateral striatum*) – области, играющей ключевую роль в планировании и изменении поведения. Целенаправленное поведение в большой степени контролируется лобными областями коры и медиальной областью стриатума (*dorsomedial striatum*).

Таким образом, можно провести компьютерную аналогию и образно сказать, что кандидатами на управление нашими действиями выступают три «процессора» (Rangel et al., 2008): павловский классический рефлекс, автоматизмы и целенаправленное поведение, которые, по всей видимости, почти всегда задействуются параллельно. Зачастую альтернативы, исходящие от этих разных систем определения ценности, совпадают. Например, когда человек испытывает чувство голода, то, в соответствии с пищевым рефлексом, привычками и целенаправленным поведением, для него особую ценность приобретает процесс поглощения пищи. Но нередко мы сталкиваемся с обратной ситуацией, когда наблюдается рассогласование в работе упомянутых выше систем. Рассмотрим поведение человека, имеющего пристрастие к алкоголю, который, несмотря на то, что через час ему необходимо сесть за руль, выпивает кружку пива в баре. С одной стороны, привычка подталкивает его к употреблению еще одной порции спиртного, с другой стороны, целеустремленное поведение определяет необходимость сокращения количества выпитого, так как человек не может не осознавать пагубность этой привычки и опасность последствий опьянения. В результате возникают конфликт и чувство вины. Следовательно, для принятия наиболее оптимального решения необходимо, чтобы активность системы целенаправленного поведения была сильнее активности системы привычек или автоматизмов, чтобы одна система подавляла другую, например, по описанному выше диффузному механизму принятия решений. Нейроэкономика, в свою очередь, стремится объяснить нерациональное поведение человека конфликтом между различными системами определения субъективной полезности.

### **Рациональные и эмоциональные системы мозга**

Нерациональность человеческого поведения не раз подтверждалась экспериментально. Нобелевский лауреат Дениэл Канеман (Kahneman, 2003) выдвинул предположение о существовании двух эволюционно и структурно различающихся систем, обуславливающих принятие решений: а) быстрой, автоматической, или бессознательной (Система 1), и



б) медленной, целенаправленной, произвольной (Система 2). Установление прямого соответствия между тремя биологическими системами определения ценности результата поведения, описанными в предыдущей главе, и дуалистической моделью принятия решения Д. Канемана является непростой задачей. С некоторым приближением классический условный рефлекс и привыкание могут быть объединены в единую автоматическую Систему 1, тогда как целенаправленное поведение (реализуемое в том числе и через инструментальные рефлексy) в таком случае будет являться частью произвольной Системы 2.

В настоящее время многочисленные нейроэкономические исследования посвящены изучению взаимодействия рациональной и эмоциональной систем в рамках дуализма Канемана. Среди наиболее часто используемых экспериментальных моделей можно назвать игру «Ультиматум» (Ultimatum Game), наиболее ярко демонстрирующую возникновение подобного рода конфликтов: два игрока получают инструкцию поделить между собой определенную сумму денег, например, 100 рублей. Один из игроков (proposer) первым предлагает способ дележа, причем он абсолютно свободен в своем решении. Предположим, игрок захочет оставить 80 рублей себе, а 20 рублей отдать своему партнеру по игре. Партнер (responder), которому сделали предложение, оказывается перед выбором: согласиться или не согласиться; однако если респондент не соглашается, то, по правилам игры, денег не получает никто и игра заканчивается.

Как было показано в поведенческих исследованиях, в среднем предлагаемая сумма колеблется вокруг половины (50%) от общего количества денег; чаще всего явно невыгодные предложения (менее 20%) неизбежно отвергаются даже в тех случаях, когда предложенная сумма превышает месячный доход игрока (Knoch et al., 2006). С точки зрения классической экономики такое поведение нерационально, ведь небольшая сумма денег все же лучше, чем их отсутствие. Таким образом, можно предположить, что респондент, которому делается невыгодное предложение, ощущает конфликт между рациональным и эмоциональным поведением: рациональный подход подразумевает принятие предложения, тогда как эмоции требуют его отвержения из-за несправедливости распределения денег.

Используя метод *функциональной магнитно-резонансной томографии* (фМРТ), американский нейробиолог Алан Сенфи (Sanfey et al., 2003) показал, что у респондента, которому сделали несправедливое предложение в игре «Ультиматум», наблюдается активация островковой коры (anterior insula), которая, как известно из других нейробиологических исследований, вовлечена в обработку негативной эмоциональной информации и особенно активна при эмоции отвращения. Интересно, что в данном исследовании активность островковой коры была пропорциональна степени несправедливости предложения. Кроме того, наблюдалась активация верхних областей лобной коры (dorsolateral prefrontal cortex, DLPFC) правого полушария и поясной извилины (anterior cingulate cortex) – областей, активация которых наблюдается при запуске когнитивных процессов самоконтроля и при внутренних конфликтах, соответственно. Особого внимания заслуживает тот факт, что по соотношению активности островковой и лобной коры (DLPFC) можно предсказать, будет полученное игроком предложение принято или отвергнуто. Если несправедливое предложение отвергалось, то наблюдалась более сильная активация островковой коры по сравнению с областью DLPFC, тогда как в случае принятия несправедливого предложения, наоборот, большая активация наблюдалась в DLPFC и меньшая в островковой коре. Разумно предположить, что негативная реакция, вызываемая нарушением существующих норм поведения (несправедливость дележа), превалировала над рациональным желанием получе-

ния выгоды. В целом приведенный пример ярко демонстрирует взаимодействие эмоциональных и рациональных процессов в момент принятия решений.

Между тем наблюдавшееся взаимодействие DLPFC и островковой коры можно объяснить по-разному. Согласно первой гипотезе, несправедливые предложения дележа денег могут вызывать желание немедленно отвергнуть их, что выражается в «эмоциональной» активации островковой коры. В этом случае параллельная активность DLPFC отражает контроль за импульсивным эмоциональным поведением через механизм рационального обдумывания. Таким образом, в предложенной модели DLPFC выступает как некий управляющий механизм, обеспечивающий контроль импульсивного поведения в социальных ситуациях (Sanfey and Chang, 2008). Тогда гипотетическое временное отключение DLPFC должно привести к еще более выраженной аверсии к несправедливости. Альтернативная гипотеза предполагает, что врожденные инстинкты, к примеру, рациональное эгоистическое желание получить деньги, нуждаются в некоем контроле со стороны механизмов, участвующих в обеспечении взаимодействия индивидуальных и социальных мотивов, в том числе в тех случаях, когда индивидум оказывается перед лицом несправедливости (Henrich et al., 2001). Согласно второй гипотезе, активация DLPFC необходима для преодоления утилитарных, или эгоцентричных (self-interest), склонностей и желаний, тесно связанных с нашими эмоциями. Следуя этой логике, при временном отключении DLPFC должно наблюдаться усиление утилитарных стремлений: вероятность того, что игрок согласится принять деньги, увеличится, а влияние эмоции негодования на принятие решения уменьшится.

Итак, обе гипотезы по-разному предсказывают последствия нарушения работы DLPFC для участников игры: согласно первой гипотезе, в случае нарушений в работе DLPFC мы будем чаще отвергать несправедливые предложения, а согласно второй – мы, наоборот, будем их чаще принимать. Благодаря развитию методов неинвазивного нейромиджинга у нейробиологов появилась возможность проверить данные гипотезы. Одним из последних методических достижений когнитивных нейронаук стала *транскраниальная магнитная стимуляция* (transcranial magnetic stimulation, TMC). TMC-стимулятор – прибор, который при помощи электромагнитной катушки (конфигурации типа петли Мебиуса для оптимальной фокусировки стимулирующего импульса), располагающейся у поверхности головы, подавляет или активирует работу прилегающих зон мозга. Когда при помощи TMC временно подавлялась активность DLPFC в правом полушарии, то испытуемые принимали несправедливые предложения в игре «Ультиматум» гораздо чаще, чем испытуемые, у которых подавлялась активность DLPFC в левом полушарии (Knoch et al., 2006). Полученные данные подтверждают гипотезу о том, что DLPFC правого полушария играет ключевую роль в подавлении или ослаблении утилитарного (рационального) поведения. Таким образом, верной оказалась вторая гипотеза: после инактивации DLPFC испытуемые активнее начинают демонстрировать утилитарное поведение, приводящее к принятию невыгодных или несправедливых предложений. Интересно, что инактивация DLPFC влияла на поведение респондента, но не на субъективную оценку ситуации, т.е. на ощущение того, насколько справедливо или несправедливо предложение, сделанное ему в ходе игры. Испытуемые, у которых в правом полушарии была инактивирована область DLPFC, отмечали, что их возмущала несправедливость ситуации, но отвергнуть несправедливые предложения после TMC-стимуляции они не могли.

В дальнейшем исследователей заинтересовала возможность влиять на поведение испытуемых, не прибегая к помощи TMC-стимуляции. В одном из исследований (Sanfey and



Chang, 2008) перед началом игры «Ультиматум» у испытуемых индуцировали плохое настроение, так что в дальнейшем поведение в ответ на несправедливые предложения началось на фоне так называемого *негативного эмоционального прайминга*. Известно, что негативный эмоциональный прайминг (переживание грусти, злости или отвращения) активирует островковую кору, следовательно, можно предположить, что негативный прайминг должен приводить к еще более высокому уровню отказа от несправедливых предложений. Исследование подтвердило, что испытуемые, которые перед проведением эксперимента просмотрели драматический фильм, демонстрировали статистически более высокие показатели отклонения несправедливых предложений по сравнению с теми, кто смотрел эмоционально нейтральный кинофильм или комедию (Sanfey and Chang, 2008). Эти экспериментальные данные не просто подтверждают нейробиологическую теорию, объясняющую иррациональность нашего поведения, но и демонстрируют, что мимолетные и незначительные изменения эмоционального состояния, не имеющие прямого отношения к текущей задаче, влияют на экономические решения индивидуума.

Оказалось, что модель, согласно которой при принятии решения доступная информация оценивается с помощью разных нейронных сетей – «когнитивной» (DLPFC) и «эмоциональной» (островковая кора), достаточно хорошо предсказывает экспериментальные результаты: принятие окончательного решения осуществляется по результатам сравнения (вычитания) когнитивной и эмоциональной оценок, и если разница между двумя альтернативами достаточно велика, то принимается решение в пользу «лучшей» альтернативы. Важно отметить, что обширные поведенческие и нейробиологические исследования принятия решений показывают, что, как правило, эмоциональная реакция на несправедливость имеет больший вес по сравнению с рациональной (утилитарной) реакцией.

Вероятно, способность оценивать справедливость социального взаимодействия чрезвычайно важна для обеспечения долговременной кооперации внутри социальной группы. Немедленное произвольное подкрепление справедливых предложений и наказание за несправедливое поведение – это важный эволюционный механизм существования и сплочения социальных групп. Поэтому автоматическая произвольная реакция Системы 1 зачастую подавляет Систему 2. «Диффузный» характер принятия решений по механизму сравнения или вычитания допускает регулирование весового вклада эмоциональной и рациональной (когнитивной) компоненты в зависимости от контекста (например, настроения), что придает системе принятия решений необходимую пластичность.

### Межвременной выбор

Другим относительно подробно изученным аспектом теории принятия решений в нейроэкономике является парадигма межвременного выбора (*intertemporal choice*), основанная на том факте, что люди по-разному оценивают величину вознаграждения в зависимости от того, в какой момент времени они предполагают его получить.

Многочисленные исследования демонстрируют феномен *временного дисконтирования* (*temporal discounting*), т. е. тенденцию занижать субъективную полезность вознаграждения, если ее получение отложено во времени (*Choice over time*, 1992). В целом мы переоцениваем величину вознаграждения, если велика вероятность его получения в скором времени, т. е. предпочитаем получить вознаграждение «здесь и сейчас». Обычно испытуемому в подобных исследованиях предлагается выбрать между двумя альтернативами: получить 10€ сегодня или 11€ завтра. Показано, что в такой ситуации испытуемый скорее все-

го предпочтет получить 10€, но сегодня. Изменим условия эксперимента: сохраним те же суммы и тот же временной интервал между их получением, но предложим отсрочку выбора на один год: хотите ли вы получить 10€ через год или 11€ через год и один день? В таком случае оказывается, что испытуемые предпочитают большую сумму, пренебрегая разницей в один день. Кажется, что такое явное противоречие при выборе альтернатив не отличается рациональностью.

В рамках гипотезы параллельной обработки (dual-processing hypothesis) можно предположить существование конкуренции между импульсивной Системой 1 и рациональной Системой 2 при оценке долговременной перспективы последствий принятия решений: система быстрых, импульсивных предпочтений, очевидно, отдает предпочтение немедленному вознаграждению, так как она малочувствительна к долговременным перспективам. Система 2, наоборот, способна оценивать преимущества получения награды в будущем.

Возникает вопрос: возможно ли объяснить феномен временного дисконтирования нейробиологическими механизмами? В одном из наиболее известных нейроэкономических исследований (McClure et al., 2004) была предложена следующая модель: допустим, что наше поведение описывается двумя параметрами:  $\beta$  и  $\delta$ . Пусть  $\beta$  – величина удельного веса немедленного или эмоционального результата, а  $\delta$  – удельный вес рационального результата на всем промежутке времени ( $t$ ). Тогда предположение о том, что параметр  $\delta$  регулируется структурами мозга, отвечающими за эмоции (например, стриатум), а параметр  $\beta$  регулируется структурами, обеспечивающими наши когнитивные функции (например, DLPFC), можно подтвердить в случае, если в ситуации межвременного выбора наблюдается дифференцированная активность этих двух областей мозга. Данные нейроимиджинговых исследований указывают на существование областей мозга, вовлеченных в принятие решения о незамедлительном получении награды ( $\beta$ -зоны: области стриатума и орбитофронтальной коры, являющиеся компонентами эмоциональной системы мозга). Одновременно были обнаружены и другие области, названные  $\delta$ -зоны (в том числе DLPFC), нейрональная активность которых не зависит от длительности временной задержки вознаграждения.

### Нейроэкономика морали в принятии решений

Описанные выше нейробиологические исследования экспериментальной игры «Ультиматум» и временного дисконтирования хорошо иллюстрируют и обосновывают современную нейроэкономическую теорию, предполагающую, что взаимодействие систем произвольной и произвольной обработки информации во время принятия решений основывается на конкуренции определенных областей мозга. Данная теория оказалась приложима и к исследованиям процессов принятия решений, связанных с системой моральных норм. Может ли убийство другого человека считаться моральным поступком? «Конечно нет!» – таков будет ответ, который первым приходит в голову. В подавляющем большинстве случаев – не может. Но, как всегда, существуют исключения из правил. Если речь идет о полицейском, который совершил убийство агрессивного преступника, защищая невинных детей, тогда совершенное им деяние может быть оправданно. Так или иначе, каждый человек постоянно сравнивает результаты своего поведения с социальными нормами. Влияние морали на принятие решений активно изучается с использованием экспериментальных моделей – *моральных дилемм* (moral dilemma).

Популярной моральной дилеммой является так называемый «случай с трамваем» (*trolley dilemma*) (Thomson, 1986): по рельсам мчится трамвай, через мгновение он неминуе-



мо собьет пятерых зазевавшихся прохожих в том случае, если не изменит направления движения. Единственный выход – постараться переключить стрелку и направить трамвай на запасные пути, но в этом случае погибнет ничего не подозревающий, случайно оказавшийся именно в этом месте человек. Предположим, что вы стоите рядом со стрелкой и можете переключить рельсы с одного пути на другой. Что предпочесть: смерть пятерых или убийство одного? Большинство людей согласится, что предпочтительна гибель одного прохожего (Greene et al., 2001). Но давайте рассмотрим аналогичную ситуацию лишь в слегка измененных условиях – «историю с пешеходным мостом» (*footbridge dilemma*) (Thomson, 1986): представьте, что вы стоите рядом с дородным незнакомцем на пешеходном мосту над дорогой, как раз между приближающимся трамваем и теми пятью несчастными потенциальными жертвами; столкнув вниз дородного незнакомца, можно остановить трамвай и спасти жизнь пяти пешеходам. Что вы выберете: спихнуть незнакомца или наблюдать за смертью пяти человек? Оказывается, что в ситуации выбора в этом «персонифицированном» случае большинство выскажется против идеи принести в жертву невинного незнакомца (Greene et al., 2001). На пешеходном мостике (персонифицированная дилемма) испытуемый оказывается в эмоционально нагруженной ситуации и вынужден сделать гораздо более тяжелый выбор, чем в ситуации переключения стрелки (неперсонифицированная дилемма), где эмоциональный накал ниже и может быть предпринято рациональное осмысление происходящего. фМРТ-сканирование показало, что участки мозга, связанные с эмоциональными и социальными реакциями (например, поясная извилина коры), активируются сильнее, когда испытуемым приходится решать персонифицированную дилемму, аналогичную «истории с пешеходным мостом», в то время как когнитивные области, ответственные за рациональные аспекты принятия решений (например, DLPFC), активируются сильнее в задачах на решение неперсонифицированных утилитарных проблем, аналогичных дилемме с переключением стрелки («случай с трамваем»).

Дальнейшие исследования сфокусировались на сравнении моральных дилемм разной степени сложности. Было обнаружено (Greene et al., 2004), что при трудном выборе активируется передняя часть поясной извилины (*anterior cingulate cortex, ACC*) – участок мозга, вовлеченный в обработку внутренних конфликтов (Botvinick et al., 2001). В то же время была показана положительная корреляция активности «рациональной» области DLPFC с принятием утилитарных решений, таких, как рациональное решение спасти пять человек, пожертвовав одним.

Обе упомянутые области мозга – ACC и DLPFC – активируются в ходе решения и другого типа моральных дилемм. Например, испытуемому предлагалось оценить и принять решение в следующей воображаемой ситуации: агрессоры захватили населенный пункт и приказали уничтожить оставшихся в живых мирных жителей. Вы и небольшая группа ваших товарищей-беженцев прячетесь в подвале большого дома. Наверху оккупанты устроили обыск. Ваш новорожденный ребенок начинает громко плакать. Вы прикрываете его рот ладонью так, чтобы не было слышно его рыданий. Если вы уберете руку, то плач будет неминуемо услышан – солдаты придут и расстреляют вас, вашего ребенка и всех скрывающихся в подвале. Чтобы сохранить жизнь себе и товарищам, остается только лишить жизни ребенка. Каков будет ваш выбор? Возможно ли убить ребенка, чтобы спасти жизнь другим людям? Этот случай – пример наисложнейшей моральной дилеммы. Когда испытуемых просили разрешить эту дилемму, они собирались с ответом чрезвычайно долго, а их решения были самыми разными. Наблюдавшаяся в данной задаче повышенная активация ACC

и DLPFC указывала на высокий уровень конфликта, при этом именно активность DLPFC предсказывала утилитарное «когнитивное» решение (Greene et al., 2004).

Инфантицидная дилемма (*infanticide dilemma*) – пример еще одной моральной дилеммы. Испытуемым предлагается дать оценку действиям женщины, которая предпочла лишить жизни своего новорожденного ребенка, чтобы избежать позора, связанного с его незаконнорожденностью. Без колебаний испытуемые высказывают свое осуждение подобных действий, ведь никто не одобрит чудовищного детоубийства. Ответ в данной дилемме напрашивается сам собой: убийство – преступление! Результаты сканирования мозга подтвердили однозначность решения испытуемых (Greene et al., 2004), показав, что активность АСС – области, отвечающей за когнитивный конфликт, и DLPFC – области, отвечающей за когнитивный контроль, была ниже, чем в эксперименте, где испытуемым было предложено обдумать возможность пожертвовать жизнью ребенка ради жизни других людей.

Данные этих и других нейроэкономических исследований доказывают существование двух конкурирующих систем нормативного («не убий») и утилитарного подхода («выбирай, что лучше») к морали, которые обеспечиваются работой различных областей мозга (Greene et al., 2004). Система 1, которую мы, по всей видимости, унаследовали от наших предков, отвечает за следование морально-социально-экономически значимым непроизвольным табу, являющимся основой наших общественных законов. Моральный утилитаризм Системы 2, возможно, возник позже в процессе эволюции и связан с появлением эволюционно новых структур – дорзальных областей лобной коры – зон мозга, являющихся нейробиологическим субстратом таких когнитивных функций, как абстрактное мышление и когнитивный контроль более высокого порядка (*high-order cognitive control*). Две рассматриваемые нами системы могут конфликтовать, результатом чего является множество примеров, когда утилитарное мышление приводило к ужасающим последствиям и прямо противоречило нормам морали. Злодеяния фашизма во время Холокоста, бесчинства тоталитарного режима в ГУЛАГе и прочие преступления зачастую оправдывались с точки зрения утилитарных, рациональных причин, противореча нормам общепринятой морали. Возможно, поэтому свидетельства массовых убийств, насилия и террора старательно скрывались от широких масс во избежание возможности возникновения морального конфликта, ведь манипулирование массовым сознанием через прессу и медийные каналы позволяет вносить принципиальные изменения в баланс между утилитарным и моральным мировоззрениями.

### Эмоции и принятие решений

Идея о параллельных (конкурирующих или взаимодействующих) нейронных сетях, обеспечивающих процесс выбора, популярна в самых разных областях нейронаук – от исследований, изучающих категориальное восприятие, до исследований процессов памяти. Считается, что в процессе принятия решений параллельные эмоциональные и когнитивные пути обработки информации функционируют как единый механизм. Многие исследования показывают, что иррациональное поведение человека может быть объяснено большим вкладом эмоциональных зон мозга, имеющих ряд преимуществ перед когнитивными областями.

Как правило, автоматические эмоциональные реакции способствуют быстрому и эффективному процессу принятия решений. Работа эмоциональной системы позволяет делать выбор между «плохими» и «хорошими» альтернативами, т.е. совершать действие, ко-



торое в корне отличается от однозначного безусловного рефлекторного ответа «стимул–реакция». Сравнительно недавно в эволюции произошла интеграция эмоций (Система 1) с более сложными и поздними в филогенетическом плане механизмами принятия решений – «когнитивной» системой целенаправленного поведения (Система 2). Интегральная модель, способная объединить обе системы, была предложена американским нейрофизиологом Эдмундом Ролзом (Rolls, 2000). Согласно теории Ролза, информация о стимуле, попадая в базальные ганглии мозга и включая стриатум, вызывает непроизвольную поведенческую реакцию, основанную на эмоциональных характеристиках стимула. Если обработка и сравнение стимулов происходят с вовлечением лобной коры, то человек вербализирует действия, активируя систему многоуровневого синтаксического планирования. В последнем случае лобные области обеспечивают принятие произвольных решений.

Другой влиятельный американский исследователь – Жозеф Леду (LeDoux, 2000) предложил так называемую модель «короткого» пути принятия решений, согласно которой такие эмоциональные зоны, как область миндалины, получают еще не обработанный, «сырой» сигнал на самом начальном этапе переработки сенсорной информации. Данный «короткий» путь позволяет организму реагировать на потенциальную опасность в течение нескольких десятков миллисекунд. Вместе с тем, эмоциональные зоны мозга также получают информацию о стимуле от кортикальных областей по «углубленному» пути обработки информации, занимающему гораздо больше времени, – слишком длительный процесс принятия решения в экстремальной ситуации может стоить человеку жизни, несмотря на то, что в этом случае происходит более детальное декодирование информации. Кроме того, по мнению Леду (LeDoux, 2000), разрыв связей между двумя вышеупомянутыми путями может приводить к целому ряду заболеваний: «Что для нормального человека является благом (например, настороженное внимание к веткам деревьев и палкам в лесу, напоминающим прячущуюся змею), то для страдающего фобиями человека может превратиться в кошмар: такой человек начинает видеть змей везде и всегда». В целом предварительная (грубая) обработка сенсорной информации происходит в нейрональных сетях, ответственных за принятие быстрых решений, после чего возможна более детальная обработка с использованием произвольного внимания. Когда необходимо быстро принять решение, в действие вступает Система 1, опережая Систему 2. Таким образом, фактор ограниченного времени серьезно влияет на механизм решения проблемы.

Получается, что обе системы – произвольная и непроизвольная – одинаково важны, но установление баланса между ними играет существенную роль в гармонизации поведения. Нарушение в одной из систем может привести к выбору неоптимальных стратегий. Результаты нейропсихологических исследований показывают, что функционирование Системы 2 рационального принятия решений тесно связано с точностью обработки информации Системой 1. Пациенты с двухсторонними повреждениями нижних областей лобной коры (orbitofrontal cortex, OFC) страдают серьезными нарушениями принятия самых разнообразных решений (Bechara and Damasio, 2005). Наблюдая за своими OFC-пациентами, американский невролог Антонио Дамасио выявил у них определенные сложности при планировании рабочего дня, взаимодействии с друзьями и супругами и т. д. Нередко совершаемые OFC-пациентами поступки приводят к финансовым неудачам (например, к проигрышам на бирже), разрыву с родственниками или потере друзей, как будто бы пациенты действуют не в своих интересах, а себе во вред. Оказалось, такие люди страдают определенными нарушениями способности к обучению и в результате совершают ошибки снова и снова

ва. Интересно, что их способность к решению задач, не требующих эмоциональной оценки, остается неизменной: обладая нормальным либо высоким интеллектом, они, тем не менее, остаются неспособными избежать ошибок в более «эмоциональных» ситуациях (Bechara et al., 1997). Например, выбор места свидания для обычного человека не составляет труда: доверившись собственным чувствам, он выберет любимое кафе, устраивающее обе стороны. В аналогичной ситуации пациент Антонио Дамасио затеет нескончаемый рациональный перебор всех подходящих мест с тщательным сравнением цен, местоположения, качества... При этом он скорее всего будет игнорировать эмоции партнера, проявляющего удивление и раздражение, уставшего от бесконечного выбора. Парадоксально, но чрезмерная рациональность может в результате привести к нерациональному поведению.

Обычно пациентов с эмоциональными расстройствами и неспособностью к принятию решений исследуют с помощью специального теста-игры (Iowa Gambling Task). В данной игре используются четыре колоды карт: А, В, С и D. Испытуемому предлагается попытаться выиграть как можно больше денег. В игре 100 раундов, во время которых игрок может выбирать по одной карте из любой колоды. Открытие карты в двух колодах (А и В) приносит крупный и незамедлительный выигрыш, но при дальнейшем использовании ведет к долгосрочным потерям (назовем их – *долговременные потери*), существенно превышающим потери в других колодах (С и D), использование которых приведет к более скромному сиюминутному выигрышу, но зато и к отсутствию потерь в будущем. Испытуемый сам выбирает, из какой колоды он берет карту, и после нескольких раундов здоровые испытуемые начинают оказывать предпочтение «хорошим» колодам С и D, приносящим большую суммарную прибыль, и избегать несчастливых колод А и В, приносящих суммарный убыток. Пациенты Дамасио вели себя иначе: они не только не избегали «плохих» колод, но даже порой предпочитали их.

Как было показано в одном из исследований (Bechara and Damasio, 2005) с использованием метода кожно-гальванической реакции (КГР) для измерения эмоциональной реакции испытуемых, у OFC-пациентов отсутствует способность предсказывать последствия, связанные с возможными долговременными потерями. Выраженных эмоциональных реакций у OFC-пациентов в ситуациях выбора карт из колод А и В, в отличие от здоровых испытуемых, выявлено не было. Таким образом, в отсутствие способности к предвидению последствий своего поведения пациенты не способны избежать болезненных проигрышей. Они продолжали выбирать несчастливые колоды точно так же, как они вели бы себя в реальной жизни – нерационально. Более того, пациенты продолжали снова и снова выбирать несчастливые колоды карт даже после того, как они полностью осознавали последствия своих действий. Таким образом, становится понятным, что деятельность эмоциональных зон мозга вовлечена в рациональное поведение. Отсутствие эмоций ведет к своеобразной диссоциации между тем, что мы говорим или знаем, и тем, как мы себя ведем. В обычной жизни эмоциональные реакции позволяют нам избегать тех действий, которые могут привести к неудачным последствиям. Причем подобный интуитивный выбор происходит еще до осознания последствий нашего поведения (Bechara et al., 1997), в то время как у OFC-пациентов коррекции поведения не происходит даже после осознания возможных последствий неоптимального выбора.

Эlegantное исследование (Shiv et al., 2005) OFC-пациентов показало, что в особых случаях их поведение, напротив, более соответствует экономическим моделям рационального поведения, чем поведение здоровых испытуемых, по той причине, что у первых приня-



тие решений никак не связано с эмоциональным опытом предыдущих решений. В данном исследовании при решении специально сконструированной задачи, где рискованная тактика приводила к большей выгоде, здоровые испытуемые проигрывали OFC-пациентам, так как в результате чередующихся проигрышей и выигрышей вырабатывали консервативную стратегию, становясь все более осторожными от раунда к раунду, чего не происходило с OFC-пациентами, принимавшими более рискованные решения.

В целом хорошо известно, что склонность к риску опосредована эмоциональным опытом и сиюминутными переживаниями, связанными, например, с текущими результатами поведения. Так, Р. Талер и Е. Джонсон (Thaler and Johnson, 1990) продемонстрировали, что человек, проигравший существенную сумму денег, обнаруживал повышенную склонность к риску при принятии дальнейших решений. Талер и Джонсон назвали этот феномен *break even effect*. После крупного выигрыша, который невозможно полностью проиграть в следующих раундах игры, также наблюдался эмоциональный подъем, который сопровождался повышенным риском, возникавшим из-за гарантированной возможности «остаться при своих» (*house money effect*). Аналогичные эффекты можно наблюдать и за пределами экспериментальной лаборатории, например, в телевизионной игре «Сделка» (Deal or No Deal) (Post et al., 2008). В этом популярном западном телевизионном шоу главным призом является колоссальная сумма – зачастую более миллиона долларов. Участники игры располагаются перед 26 запечатанными чемоданами, в которых находятся разные суммы денег – от 1 доллара до 5 млн долларов. Разумеется, участники игры не знают, какая сумма находится в конкретном плотно закрытом чемодане, но в каждом раунде игры участник избавляется от одного из 26 чемоданов, с тем чтобы к концу игры оставить всего один чемодан с максимальным выигрышем. Каждый чемодан, который игрок убирает из игры, немедленно открывается, и его содержимое оглашается: при удачном выборе игрок избавляется от чемодана с минимальным выигрышем (скажем, в 1 доллар), а при неудаче он может вывести из игры и максимальный выигрыш в 5 млн долларов. Ключевой момент игры заключается в том, что перед тем, как игрок делает очередной выбор, какой чемодан следует вывести из игры и открыть, ведущий предлагает ему некоторую сумму денег в обмен на прекращение игры. Торговля сопровождается вопросом: «Deal or No Deal?» («Договорились?»). У участника игры, таким образом, есть выбор: рискнуть и продолжить игру, открыв (удалив) следующий чемодан, или не рисковать и взять деньги. Структура телевизионной игры позволяет экономистам изучать влияние, которое оказывает только что принятое игроком решение на его дальнейший выбор: рисковать или нет в следующем раунде. Анализ игр показал, что испытуемые действительно склонны рисковать, демонстрируя *break even effect*, если случайно выводят из игры максимальный выигрыш, и *house money effect*, удачно удалив минимальные выигрыши (Post et al., 2008).

В своем недавнем нейроэкономическом исследовании Кайса Хютонен и соавторы (Hytönen et al., 2010) имитировали условия TV-шоу для изучения нейрональных процессов, отвечающих за вышеперечисленные эффекты, которые приводят к рискованным решениям. Компьютерная версия игры позволила исследователям вызывать и сравнивать активацию мозга во время принятия решений как после того, как удалялся чемодан с маленькой суммой денег (тем самым увеличивая вероятность того, что в игре остались чемоданы с большей суммой денег), так и после того, как из игры удалялся чемодан с большой суммой денег, а также в промежуточных ситуациях. В исследовании было показано, что усиление тенденции к риску (как после проигрыша, так и после выигрыша) ассоциируется

с усилением активности эмоциональных зон мозга, таких, как островковая кора, и уменьшением активности тех зон, которые отвечают за рациональные аспекты принятия решений (DLPFC). Таким образом, узнать о том, будет ли рисковать человек, можно по соотношению активации нейрональных сетей «автоматической» Системы 1 и «когнитивной» Системы 2. Исследование Кайсы Хютонен и коллег (Hytönen et al., 2010) показывает, что работа автоматической системы может являться источником иррациональной склонности к риску, основанной в том числе на прошлом эмоциональном опыте, влияние которого сказывается, как подтверждают многочисленные нейроэкономические исследования, на принятии человеком самых разных решений.

### Социальное влияние

В предыдущих разделах мы обсуждали механизмы принятия решений независимыми индивидуумами, т. е. автономными испытуемыми, абсолютно свободными от влияния окружающих. Однако такая идеальная ситуация в реальной жизни представляется маловероятной. Помимо физического присутствия рядом других людей мы испытываем всевозможные формы давления со стороны общества – от рекламы и семейных традиций до художественной или публицистической прозы. Важно отметить, что наше поведение эффективно регулируется моральными категориями, находящимися в прямой зависимости от общепринятых социальных норм поведения. Влияние других людей на поведение индивидуума принято называть *социальным влиянием*, а общепринятые нормативные паттерны поведения – *социальными нормами* (Cialdini and Goldstein, 2004).

Одним из наиболее известных феноменов социального бытия является конформизм, определяемый как изменение поведения или мнения в направлении, заданном поведением большинства (Cialdini and Goldstein, 2004). При анализе конформного поведения социальные психологи выделяют следующие мотивационные основы: стремление правильно интерпретировать окружающую действительность, равно как и желание вести себя адекватно ситуации (1), стремление получить одобрение окружающих (2), а также стремление сохранить позитивное представление о самом себе (3), поддерживая так называемую «Я-концепцию» (Cialdini and Goldstein, 2004).

Пожалуй, самым известным исследованием конформизма явилось новаторское исследование Соломона Аша (Asch, 1951) – американского психолога, показавшего в начале 60-х годов прошлого столетия, что большинство из исследованных им испытуемых изменяло свое мнение в сторону очевидно ложного мнения большинства, стремились согласовать свое поведение с принятыми в группе нормами.

В настоящее время принято разделять социальные (групповые) нормы на следующие типы (Cialdini and Goldstein, 2004):

- 1) *иньюктивные*, или запретительные, нормы (injunctive norms), имеющие моральную окраску и являющиеся предписанием того, как, собственно, следует поступать;
- 2) *дескриптивные*, или описательные, нормы (descriptive norm), определяющие, как большинство людей поступает в данной ситуации независимо от приемлемости и неприемлемости данного действия.

Действие иньюктивных норм («не убий», «не укради» и т. д.) обычно подкрепляется различными формами наказаний. В дескриптивных нормах содержится информация о поведении, преобладающем в определенной группе лиц. Несмотря на то, что дескриптивные нормы несут скорее информативный характер, они обладают необычайной эффективностью.



ностью. Именно дескриптивные нормы в значительной степени определяют процент недобросовестных налогоплательщиков, уровень преступности, заботу об экологии, частоту семейных измен и многие другие формы поведения (Kenrick et al., 2004). Интересно, что в обычной жизни мы почти не задумываемся над тем, какое влияние на нас оказывают социальные нормы.

Несмотря на то, что год от года в научной литературе появляется все больше данных о том, какие нейрональные механизмы лежат в основе процесса принятия решений, мы до сих пор относительно мало знаем о нейробиологических механизмах социального влияния на наши решения. Каким может быть механизм социального влияния дескриптивных норм? В то время как многие социально-психологические теории подчеркивают роль социального одобрения или принадлежности к группе как своеобразного социального подкрепления (Cialdini and Goldstein, 2004), микроэкономическая теория в большей степени сосредоточена на изучении влияния наказания за нарушение норм (Fehr and Fischbacher, 2004). С точки зрения когнитивно-поведенческого подхода из обеих концепций следует, что конформность по отношению к социальным нормам обуславливается обучением с подкреплением: социальные нормы избирательно усиливают (подкрепляют) определенное поведение. Можно ли проверить данную гипотезу в нейроэкономическом исследовании?

Целенаправленное поведение требует контроля за результатами и, следовательно, постоянной обратной связи (Montague et al., 2006): оптимальные поведенческие акты закрепляются, а ошибочные исправляются. Иными словами, ошибки указывают на необходимость изменения поведения и таким образом приводят к научению. Многие модели обучения с подкреплением включают звено *ошибки предсказания вознаграждения* (подкрепления) – разницу между ожидаемым и получаемым результатом (Sutton and Barto, 1998). Считается, что ошибка предсказания вознаграждения корректирует наши действия через механизм сигнализации о необходимости изменить поведение. К настоящему моменту экспериментально подтверждена и охарактеризована функция поясной извилины (в особенности ее центральной части – rostral cingulate zone, RCZ) в генерации данного сигнала (Gehring et al., 1993). Есть основания полагать, что помимо поясной извилины в распределенную нейронную сеть, осуществляющую контроль за результатами принятия решений, вовлечены и другие структуры, такие, как нижние области стриатума (ventral striatum), в особенности прилежащее ядро (nucleus accumbens), которое в популярной литературе зачастую называют *центром удовольствия*.

Размышляя над источниками социального влияния, исследователи (Klucharev et al., 2009) предположили, что в ответ на рассогласование поведенческого акта с общепринятыми нормами нейроны нашего мозга генерируют сигнал, сходный по механизму с сигналом ошибки предсказания вознаграждения (как его определяет теория обучения с подкреплением) и являющийся автоматическим сигналом к тому, чтобы поведение индивида изменилось и стало похожим на поведение большинства. Допустим, что конформность действительно имеет нейробиологическое происхождение, сходное по механизмам с обучением с подкреплением, тогда логично было бы предположить, что на активность поясной извилины и прилежащего ядра будет оказывать влияние наличие конфликта с социальными нормами. Если активность в этих зонах будет сходна с ошибкой предсказания вознаграждения, то по изменению степени активации оказывается возможным предсказывать, изменит испытуемый свое мнение или нет. Чтобы проверить эту гипотезу, исследователи предложи-

ли студенткам, активность мозга которых оценивалась с помощью фМРТ, оценить привлекательность женских лиц. После того как испытуемая оценивала привлекательность изображенного лица, ей предъявлялась групповая оценка, отражавшая «мнение большинства», якобы полученная в результате аналогичного тестирования 200 студентов. В действительности групповые оценки генерировались компьютером по специальному псевдослучайному алгоритму: а) оценка группы более позитивна, чем оценка испытуемого; б) оценка группы более негативна, чем оценка испытуемого; в) оценки совпадают. Таким образом, в пробах, когда мнение испытуемых отличалось от нормативной оценки большинства, экспериментаторы вызывали конфликт мнений. Для оценки степени конформности каждого из участников эксперимента через полчаса после выполнения основной задачи испытуемым неожиданно предлагалось оценить те же самые фотографии, предъявленные в новом (случайном) порядке.

Оказалось, что испытуемые систематически изменяли свое мнение о привлекательности лиц в соответствии с мнением большинства, т. е. проявляли конформность. Как и ожидалось, рассогласование собственного мнения с мнением большинства приводило к активации центральной части поясной извилины (RCZ) – области, отвечающей за детекцию внутренних конфликтов и ошибок предсказания вознаграждения; в то же время наблюдалось торможение активации прилежащего ядра, т. е. осуществлялось генерирование сигнала ошибки, означающего необходимость изменения своего мнения и согласования его с социальной нормой. Интересно, что активность прилежащего ядра коррелировала с индивидуальным уровнем конформности испытуемых (рис. 2).

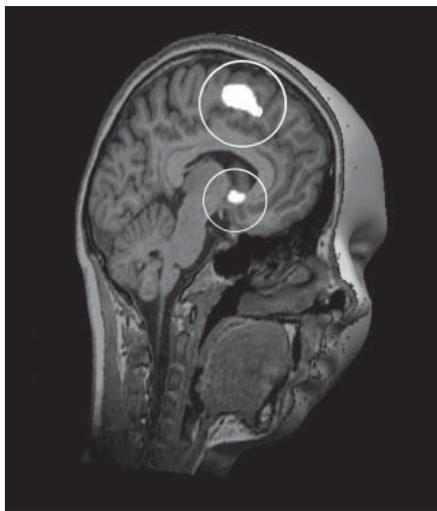


Рис. 2. Активация поясной извилины (выделено белым цветом и окружностью вверху изображения) и дезактивация прилежащего ядра (выделено белым цветом и окружностью внизу изображения) в ходе конфликта с мнением большинства. Показатели активности этих же зон коррелируют со степенью конформности испытуемых (адаптировано из: Klucharev et al., 2009)

видят похожих на себя полных людей, у них не возникает желания похудеть – «зачем?» – ведь все вокруг такие же полные, как и я.

В этом и последующих исследованиях были получены доказательства сходства нейробиологических механизмов социального влияния с механизмами процесса обучения с подкреплением (Campbell-Meiklejohn et al., 2010; Klucharev et al., 2009). Результаты исследований указывают на то, что конформность представляет собой автоматическую реакцию (сходную с Системой 1). Автоматическое корректирование поведения в соответствии с мнением большинства может основываться на более точном моделировании реальности группой людей по сравнению с отдельными индивидами (Yaniv, 2004) и, возможно, имеет важное приспособительное значение в эволюции. Однако у такой автоматической системы могут быть и свои недостатки. Сбой в ней может привести к нерациональному поведению. Например, в США обнаруживается необычайно большое количество людей с избыточным весом на весьма ограниченной территории. Следуя вышеизложенной логике, автоматическая система корректировки поведения в данном случае дает сбой: когда люди с избыточным весом



Ричард Докинз (Dawkins, 1976) выдвинул предположение о существовании *эволюционно стабильного* поведения большинства, постоянно тестируемого эволюцией, а значит, оптимального в данных условиях окружающей среды. Эволюционно стабильная стратегия – это такая стратегия, которой руководствуются большинство особей популяции и которая является предпочтительнее всех иных альтернативных стратегий, т. е. с эволюционной точки зрения поведение может быть усвоено большинством лишь в том случае, если оно лучше своих альтернатив. В результате рациональным решением может считаться решение следовать за большинством. Таким образом, согласно Докинзу, каждый индивидуум старается максимизировать свою личную выгоду и «конформизм» – единственно верная стратегия, позволяющая выжить, ведь отклонение от оптимальной стратегии большинства наказывается в ходе естественного отбора. Получается, что автоматическая конформность, являясь эффективной стратегией на этапе естественного отбора, может сыграть злую шутку и привести к неожиданным последствиям функционирования человека в современном обществе.

### Заключение и выводы

Нейробиологические исследования продемонстрировали существование нескольких параллельно действующих нейрональных систем, обеспечивающих процесс принятия решений и специализирующихся на решении тех или иных типов поведенческих задач (рис. 3), что само по себе не является удивительным фактом, ведь естественный отбор постоянно тестирует эффективность работы нейрональных сетей. В то же время именно из-за высокой специализации возможности таких нейрональных систем могут оказаться весьма ограниченными, особенно в быстро меняющейся современной среде. Социальная история современного человека несоизмеримо короче, чем вся предыдущая биологическая история вида *Homo sapiens*. Неудивительно, что функциональные возможности нервной системы человека неограниченны и не всегда могут соответствовать постоянно трансформирующейся современной действительности, поскольку большинство адаптивных особенностей и способностей человеческого мозга было отобрано в процессе эволюции в более «консервативной» среде. Например, средняя продолжительность жизни англичанок сегодня 82 года, а английских мужчин 71 год. В начале XX века, в 1901 году, средний возраст англичан равнялся 49 и 45 годам соответственно, а 10 тысяч лет назад люди жили примерно 20 лет. Логично предположить, что обсуждавшийся ранее феномен временного дисконтирования (temporal discounting) (когда человек хочет получить «все и сразу», не дожидаясь материальных благ отдаленного порядка) и увеличение продолжительности жизни могут быть драматическим образом связаны друг с другом, – современный человек попросту недооценивает значение отложенного во времени вознаграждения, поскольку никогда в своей эволюционной истории не жил так долго.

В целом результаты нейроэкономических исследований убедительно указывают на то, что осуществление выбора в процессе принятия решения обусловлено работой параллельных нейрональных систем. Автоматическая произвольная система обеспечивает быструю реакцию на меняющиеся условия, но порой дает сбой и приводит к экономически неоптимальным решениям. Эта система возникла раньше всех других систем принятия решений и зачастую плохо адаптирована к реалиям современной экономической действительности. Возможно, именно поэтому возникла необходимость в формировании произвольной системы, которая корректировала бы деятельность произвольных механизмов. Согласно современной нейроэкономической теории, для понимания механизмов поведения челове-



ка следует принимать во внимание психический автоматизм человеческой иррациональности, с одной стороны, не преувеличивая роль рациональных механизмов в принятии решений – с другой.

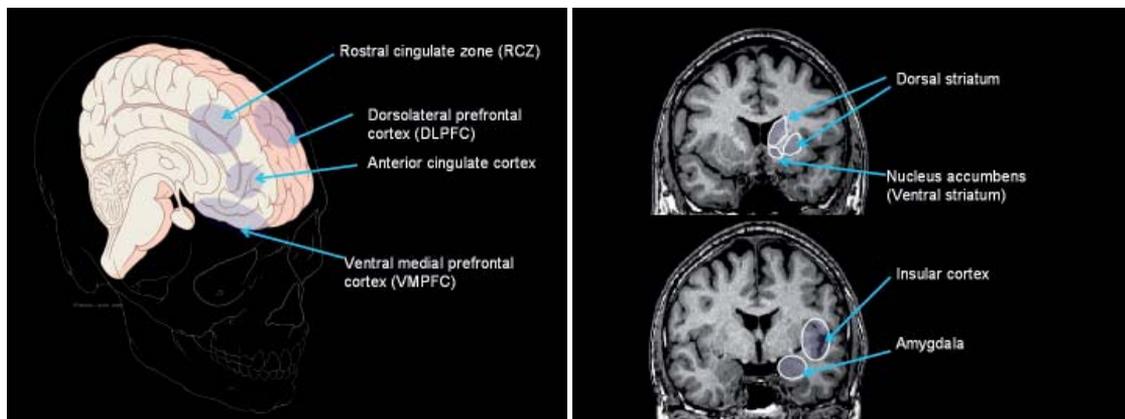


Рис.3. Области мозга, играющие ключевую роль в принятии решений:

- Rostral cingulate zone (RCZ) – ростальная область **поясной коры**.
- Dorsolateral prefrontal cortex (DLPFC) – верхняя **лобная кора**.
- Anterior cingulate cortex – передняя область **поясной коры**.
- Ventral medial prefrontal cortex (VMPFC) – нижняя медиальная **лобная кора**.
- Dorsal striatum – верхняя часть **стриатума** (полосатого тела).
- Nucleus accumbens (ventral striatum) – **прилежащее ядро** (нижний стриатум).
- Insular cortex – **островковая кора**.
- Amygdala – **миндалины**.

### Литература

- Asch S. Effects of group pressure upon the modification and distortion of judgments // Groups, Leadership and Men Research in Human Relations / Ed. H. Guetzkow, Pittsburgh: Carnegie Press, 1951. P. 177–190.
- Bechara A., Damasio A. The Somatic Marker Hypothesis: A Neural Theory of Economic Decision // Games and Economic Behavior. 2005. № 52. P. 336–372.
- Bechara A., Damasio H., Tranel D., Damasio A.R. Deciding advantageously before knowing the advantageous strategy // Science. 1997. № 275. P. 1293–1295.
- Bogacz R. Optimal decision-making theories: linking neurobiology with behaviour // Trends in Cognitive Sciences. 2007. № 11. P. 118–125.
- Botvinick M.M., Braver, T.S., Barch D. M., Carter C.S., Cohen J.D. Conflict monitoring and cognitive control // Psychological Review. 2001. № 108. P. 624–652.
- Campbell-Meiklejohn D., Bach D., Roepstorff A., Dolan R., C.F. How the Opinion of Others Affects Our Valuation of Objects // Current Biology. 2010. № 20. P. 1165–1170.
- Cialdini R.B., Goldstein N.J. Social influence: compliance and conformity // Annual Review of Psychology. 2004. № 55. P. 591–621.
- Dawkins R. The Selfish Gene. New York: Oxford University Press, 1976.
- Fehr E., Fischbacher U. Third-party punishment and social norms // Evolution and Human Behavior. 2004. № 25. P. 63–87.
- Gazzaniga M. S. Forty-five years of split-brain research and still going strong // Nature Reviews Neuroscience. 2005. № 6. P. 653–659.



- Gehring W.J., Goss B., Coles M.G.H., Meyer D.E. and Donchin E.* A neural system for error detection and compensation // *Psychological Sciences*. 1993. № 4. P. 385–390.
- Choice over time/ Eds. G. Loewenstein and J. Elster. New York: Russell Sage Foundation, 1992.
- Glimcher P.W.* The neurobiology of visual-saccadic decision making // *Annual Review of Neuroscience*. 2003. № 26. P. 133–179.
- Glimcher P.W. and Rustichini A.* Neuroeconomics: the consilience of brain and decision // *Science*. 2004. № 306. P. 447–452.
- Greene J.D., Nystrom L.E., Engell A.D., Darley J.M. and Cohen J.D.* The neural bases of cognitive conflict and control in moral judgment // *Neuron*. 2004. № 44. P. 389–400.
- Greene J.D., Sommerville R.B., Nystrom L.E., Darley J.M. and Cohen J.D.* An fMRI investigation of emotional engagement in moral judgment // *Science*. 2001. № 293. P. 2105–2108.
- Henrich J., Boyd R., Bowles S., Camerer C., Fehr E., Gintis H. and McElreath R.* In search of Homo economicus: Behavioral experiments in 15 small-scale societies // *American Economic Review*. 2001. № 91. P. 73–78.
- Hytönen K., Baltussen G., van den Assem, M.J., Klucharev V., Smidts A. and Post G.T.* How Prior Gains and Losses Influence Our Decisions Under Risk: An fMRI Study (submitted for publication) [2010].
- Kahneman D.* A perspective on judgment and choice: mapping bounded rationality // *American Psychologist*. 2003. № 58. P. 697–720.
- Kenrick D.T., Neuberg S.L. and Cialdini, R.B.* *Social Psychology: Unraveling the Mystery*, 3 ed. Boston: Allyn & Bacon. 2004.
- Klucharev V., Hytonen K., Rijpkema M., Smidts A. and Fernandez G.* Reinforcement learning signal predicts social conformity // *Neuron*. 2009. № 61. P. 140–151.
- Klucharev V., Smidts A., Fernandez G.* Brain mechanisms of persuasion: how 'expert power' modulates memory and attitudes // *Soc. Cogn. Affect Neurosci.* 2008. № 3. P. 353–366.
- Kuhnen C.M., Knutson B.* The neural basis of financial risk taking // *Neuron*. 2005. V. 47. № 5. P. 763–770.
- Knutson B., Taylor J., Kaufman M., Peterson R., Glover G.* Distributed neural representation of expected value // *J. Neurosci.* 2005. V. 25. № 19. P. 4806–4812.
- Knoch D., Pascual-Leone A., Meyer K., Treyer V. and Fehr E.* Diminishing reciprocal fairness by disrupting the right prefrontal cortex // *Science*. 2006. № 314. P. 829–832.
- LeDoux J.E.* Emotion circuits in the brain // *Annual Review of Neuroscience*. 2000. № 23. P. 155–184.
- Leiguarda R., Starkstein S., Noguez M., Berthier M. and Arbeláiz R.* Paroxysmal alien hand syndrome // *Journal of Neurology, Neurosurgery & Psychiatry*. 1993. № 56. P. 788–792.
- McClure S.M., Laibson, D.I., Loewenstein G. and Cohen J.D.* Separate neural systems value immediate and delayed monetary rewards // *Science*. 2004. № 306. P. 503–507.
- Montague P.R., King-Casas B. and Cohen J.D.* Imaging Valuation Models in Human Choice // *Annual Review of Neuroscience*. 2006. № 29. P. 417–448.
- Plassmann H., O'Doherty J., Shiv B., Rangel A.* Marketing actions can modulate neural representations of experienced pleasantness // *Proc. Natl. Acad. Sci. U S A*. 2008. V. 105. № 3. P. 1050–1054.
- Post T., van den Assem, M., Baltussen, G., and Thaler, R.* Deal or No Deal? Decision Making under Risk in a Large-Payoff Game Show // *American Economic Review*. 2008. № 98. P. 38–71.
- Rangel A., Camerer C. and Montague P.R.* A framework for studying the neurobiology of value-based decision making // *Nature Reviews Neuroscience*. 2008. № 9. P. 545–556.
- Rilling J.K., King-Casas B. and Sanfey A.G.* The neurobiology of social decision-making // *Current Opinion in Neurobiology*. 2008. № 18. P. 159–165.
- Rolls E.T.* *Precis of The brain and emotion* // *Behavioral and Brain Sciences*. 2000. № 23. P. 177–191; P. 192–233.
- Sanfey A.G. and Chang L.J.* Multiple systems in decision making // *Annals of the New York Academy of Sciences*. 2008. № 1128. P. 53–62.
- Sanfey A.G., Rilling J.K., Aronson J.A., Nystrom L.E. and Cohen J.D.* The neural basis of economic decision-making in the Ultimatum Game // *Science*. 2003. № 300. P. 1755–1758.

*Shiv B., Loewenstein G., Bechaka A., Damasio H., Damasio A.R.* Investment behavior and the negative side of emotion // *Psychological Science*. 2005. № 16. P. 435–439.

*Sugrue L.P., Corrado G.S., and Newsome W.T.* Choosing the greater of two goods: neural currencies for valuation and decision making // *Nature Reviews Neuroscience*. 2005. № 6. P. 363–375.

*Sutton R.S. and Barto A.G.* Reinforcement Learning: An introduction. MIT Press, 1998.

*Thaler R.H. and Johnson E.J.* Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice // *Management Science*. 1990. № 36. P. 643–660.

*Thomson J.J.* Rights, Restitution, and Risk: Essays, in Moral Theory. Cambridge, MA: Harvard University Press, 1986.

*Tom S.M., Fox C.R., Trepel C. and Poldrack R.A.* The neural basis of loss aversion in decision-making under risk // *Science*. 2007. № 315. P. 515–518.

*Yaniv I.* The benefit of additional opinions. *Current Directions in Psychological // Science*. № 13. P. 75–78.

## NEUROECONOMICS: THE NEUROBIOLOGY OF DECISION-MAKING

*KLUCHAREV V.A., University of Basel, Switzerland, St. Petersburg State University, St. Petersburg*

*SMIDTS A., Erasmus University, Rotterdam, Netherlands*

*SHESTAKOVA A.N., Moscow City University of Psychology and Education, Moscow, St. Petersburg State University, St. Petersburg*

The ongoing synthesis of economic sciences, psychology and neuroscience in the new discipline of neuroeconomics, as well as development of new neuroimaging methods promise the appearance of a new interdisciplinary model, which explains decision-making mechanisms. The results of the neuroeconomics studies show that decision-making is determined by the work of relatively independent neural systems in the brain, chosen through natural selection. Interaction of «cognitive» and «emotional» neural mechanisms of decision making determines the degree of the rationality of human behavior.

**Keywords:** neuroeconomics, neurobiology, decision making, social influence, temporal discounting, moral dilemma, fMRI.