

Принятие решения в условиях неопределенности: стратегии исследования и использования

Сайфулина К.Э.

*Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ),
г. Москва, Российская Федерация
ORCID: <https://orcid.org/0000-0002-2017-0811>, e-mail: kseniasayfulina@gmail.com*

Козунова Г.Л.

*Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ),
г. Москва, Российская Федерация
ORCID: <https://orcid.org/0000-0002-1286-8654>, e-mail: chukhutova@gmail.com*

Медведев В.А.

*Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ),
г. Москва, Российская Федерация
ORCID: <https://orcid.org/0000-0002-3252-8809>, e-mail: ixdon@yandex.ru*

Рытикова А.М.

*Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ),
г. Москва, Российская Федерация
ORCID: <https://orcid.org/0000-0003-0153-9457>, e-mail: ann.zelener@mail.ru*

Чернышев Б.В.

*Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ),
Московский государственный университет имени М.В. Ломоносова (МГУ имени М.В. Ломоносова); Национальный
исследовательский университет «Высшая школа экономики» (НИУ ВШЭ), г. Москва, Российская Федерация
ORCID: <https://orcid.org/0000-0002-8267-3916>, e-mail: b_chernysh@mail.ru*

Принятие человеком решений в условиях дефицита информации сопряжено с построением, проверкой и уточнением гипотез. В новой среде субъект сталкивается с высоким уровнем неопределенности, поэтому поведение должно быть вариабельным: это позволяет собирать информацию о закономерностях среды и находить наиболее выгодные опции. Такое поведение соответствует стратегии исследования. После формирования внутренней модели среды становится оправданной стратегия использования — т. е. применение выгодных опций, уже известных субъекту. В меняющейся или сложной среде оптимально применять обе стратегии попеременно. Баланс этих двух стратегий активно изучается в психологии, нейробиологии, нейроэкономике. В данном обзоре мы рассмотрим факторы, влияющие на баланс между стратегиями исследования и использования, механизмы принятия решения в условиях неопределенности, нейрофизиологические основы поддержания стратегий исследования/использования и переключения между ними, осветим роль основных задействованных в этих процессах областей мозга и нейромедиаторов.

Ключевые слова: неопределенность, принятие решения, стратегии исследования и использования, норадреналин, дофамин, ацетилхолин.

Финансирование: Исследование выполнено при финансовой поддержке Российского научного фонда (РНФ) в рамках научного проекта №20-18-00252.

Благодарности: Авторы благодарят Т.А. Строганову за неоценимый вклад в инициацию и продвижение исследования нейрокогнитивных механизмов принятия решений на базе МЭГ-центра.

Для цитаты: Принятие решения в условиях неопределенности: стратегии исследования и использования [Электронный ресурс] / К.Э. Сайфулина, Г.Л. Козунова, В.А. Медведев, А.М. Рытикова, Б.В. Чернышев // Современная зарубежная психология. 2020. Том 9. № 2. С. 93–106. DOI: <https://doi.org/10.17759/jmfp.2020090208>

Decision making under uncertainty: exploration and exploitation

Ksenia E. Sayfulina

*Moscow State University of Psychology & Education, Moscow, Russia,
ORCID: <https://orcid.org/0000-0002-2017-0811>, e-mail: kseniasayfulina@gmail.com*

Galina L. Kozunova

Moscow State University of Psychology & Education, Moscow, Russia,
ORCID: <https://orcid.org/0000-0002-1286-8654>, e-mail: chukhutova@gmail.com

Vladimir A. Medvedev

Moscow State University of Psychology & Education, Moscow, Russia,
ORCID: <https://orcid.org/0000-0002-3252-8809>, e-mail: ixdon@yandex.ru

Anna M. Rytikova

Moscow State University of Psychology & Education, Moscow, Russia,
ORCID: <https://orcid.org/0000-0003-0153-9457>, e-mail: ann.zelener@mail.ru

Boris V. Chernyshev

Moscow State University of Psychology & Education; Lomonosov Moscow State University;
National Research University «Higher School of Economics», Moscow, Russia,
ORCID: <https://orcid.org/0000-0002-8267-3916>, e-mail: b_chernysh@mail.ru

Decision-making under conditions of the lack of sufficient information is associated with hypotheses construction, verification and refinement. In a novel environment subjects encounter high uncertainty; thus their behavior needs to be variable and aimed at testing the range of multiple options available; such variability allows acquiring information about the environment and finding the most beneficial options. This type of behavior is referred to as exploration. As soon as the internal model of the environment has been formed, the other strategy known as exploitation becomes preferential; exploitation presupposes using profitable options that have already been discovered by the subject. In a changing or complex (probabilistic) environment, it is important to combine these two strategies: research strategies to detect changes in the environment and utilization strategies to benefit from the familiar options. The exploration-exploitation balance is a hot topic in psychology, neurobiology, and neuroeconomics. In this review, we discuss factors that influence exploration-exploitation balance and its neurophysiological basis, decision-making mechanisms under uncertainty, and switching between them. We address the roles of major brain areas involved in these processes such as locus coeruleus, anterior cingulate cortex, frontopolar cortex, and we describe functions of some important neurotransmitters involved in these processes — dopamine, norepinephrine, and acetylcholine.

Keywords: uncertainty, decision-making, exploration-exploitation trade-off, norepinephrine, dopamine, acetylcholine.

Funding: The reported study was funded by Russian Science Foundation (RSF), project number 14-06-14029.

Acknowledgements: The authors are grateful to Stroganova T.A. for her great contribution to research on neurocognitive mechanisms of decision-making in the Moscow MEG center.

For citation: Sayfulina K.E., Kozunova G.L., Medvedev V.A., Rytikova A.M., Chernyshev B.V. Decision making under uncertainty: exploration and exploitation [Elektronnyi resurs]. *Sovremennaya zarubezhnaya psikhologiya* = *Journal of Modern Foreign Psychology*, 2020. Vol. 9, no. 2, pp. 93—106. DOI: <https://doi.org/10.17759/jmfp.2020090208> (In Russ.; Abstr. in Engl.).

Три способа принятия решения в условиях дефицита информации

Многие ситуации реальной жизни вынуждают человека делать выбор, несмотря на недостаточную осведомленность о его возможных последствиях.

Ограниченный индивидуальный опыт позволяет субъекту учесть лишь часть факторов, которые влияют на результат. В этом случае решение принимается на основе субъективного вероятностного прогноза — иначе говоря, интуитивно [42]. Поведенческий ответ в условиях неопределенности может быть сформирован тремя способами: логическим, статистическим и эвристическим [16]. Каждый из этих способов имеет свои ограничения и допускает возможность различных ошибок.

Логический поиск решения оптимален лишь при наличии исчерпывающей информации. Он линеен,

схематичен, длителен и требует высокой концентрации внимания. При формировании суждений некоторые критически важные обстоятельства могут быть упущены из виду.

Недостатки логического мышления становятся очевидными при столкновении человека с известными экономическими парадоксами [1]. Например, два товара стоят вместе 1 доллар и 10 центов. Один из них на 1 доллар дороже другого. Какова стоимость каждого из них по отдельности? Более половины успешно обучающихся студентов престижных американских университетов с уверенностью отвечали, что первый стоил 1 доллар, а второй — 10 центов. Они буквально не замечали, что разница между 1 долларом и 10 центами принципиально не может равняться 1 доллару. Эта типичная ошибка объясняется склонностью большинства людей к упрощению информации, окру-

глению чисел, следованию привычным стереотипам мышления.

Другой способ принятия решения — статистический [5]. В этом случае субъект делает выбор, основываясь только на собственном, по определению ограниченном, опыте проб и ошибок. Каждый текущий исход обновляет внутреннюю модель ситуации, на основе которой будет сделан следующий выбор. При этом действуют базовые механизмы обучения с подкреплением: часто подкрепляемое поведение повторяется, а редко вознаграждаемое — угасает [28]. Экспериментально показано, что к успешному вероятностному прогнозированию на основе частотности событий способны очень маленькие дети на предречевом этапе онтогенеза и животные с высоким уровнем развития нервной системы [16].

Еще менее надежен эвристический способ принятия решений, в основе которого лежит формирование эвристик — ассоциаций между похожими или совпадающими по времени событиями. Случайное сходство единичного события с предупреждающим сигналом о награде или наказании из прошлого опыта субъекта сообщает ему субъективную ценность, хотя такая аналогия может быть неправомерна.

При всей своей ненадежности эвристики имеют одно ключевое преимущество перед логическим и статистическим способами принятия решения: они не требуют длительного накопления опыта. При удачном стечении обстоятельств они позволяют человеку или животному быстро сформировать оптимальный поведенческий ответ в незнакомой среде [53].

Эвристический способ принятия решений может лежать в основе рискованного поведения, как человека, так и животных. В качестве примера можно привести феномен склонности голубей к риску [63]. При выборе из двух альтернатив птицы отдавали предпочтение той из них, которая была сопряжена с низкими (20%) шансами получить большую награду, отказываясь от гарантированного, но скромного подкрепления.

Интересно, что содержание голубей в обогащенной среде вместе с другими особями уменьшает у них число неоптимальных выборов. Можно предполагать, что избыток разнообразных ассоциаций в обогащенной среде естественным образом способствует элиминации неэффективных связей.

Склонность к выборам с низкими шансами получить большую награду также характерна для азартных игроков [64].

Неопределенность как характеристика ситуаций реальной жизни

В отличие от заведомо безопасной (во всяком случае, для человека) экспериментальной комнаты, в реальной жизни невыгодные решения могут приводить к значительным потерям. Самые драматические последствия человеческих ошибок наблюдаются в

медицинской диагностике, политике, судебной экспертизе.

Также неопределенность исхода является неотъемлемой частью спортивных соревнований и творческих конкурсов. С ней повсеместно сталкиваются профессионалы в бизнесе и специалисты по подбору кадров.

Неопределенность крайне важна в контексте нейроэкономики. Часто она является неотъемлемой характеристикой среды для человека как потребителя. Изобилие взаимозаменяемых товаров, продуктов и изделий ставит современного потребителя перед беспрецедентным количеством вариантов выбора. Человек вынужден принимать решение о том, покупать ли уже знакомый продукт или попробовать что-нибудь новое; он должен выбирать между несколькими вариантами, делать прогнозы на основе множества противоречивых данных. Доля неопределенности остается и после окончательного принятия решения, поскольку, выбирая одно, человек неизбежно лишается другого. Был ли сделанный выбор оптимальным — неизвестно. Эта дилемма не устраняется и в случае приобретения сразу нескольких сопоставимых товаров, поскольку это подразумевает отказ от денег, потраченных на дополнительную покупку. Избыточная когнитивная нагрузка делает потребителей уязвимыми для разнообразных маркетинговых технологий.

Активное построение гипотез на основе индивидуального опыта

Принятие решений в условиях неопределенности стало предметом систематических научных исследований с начала 1970-х гг. Начало этому направлению положили фундаментальные работы американского психолога Д. Канемана, в 2002 г. удостоенного Нобелевской премии за вклад в развитие экономических наук.

Согласно классической концепции Канемана, неопределенность включает в себя два компонента: внешний и внутренний [27]. Внешнюю неопределенность обуславливают случайные условия окружающей обстановки, которые субъект не может контролировать или влиять на них. Внутренняя неопределенность устанавливается субъективной неосведомленностью индивида об обстоятельствах, в которых он находится.

Для оптимального поведения важна оценка субъектом собственной компетентности в решении задачи. Способность к такой оценке проявляют даже животные.

Например, в одном из экспериментов [62] крысы делали выбор между легкодоступной, но небольшой наградой и шансом получить большую награду за решение задачи, уровень сложности которой варьировался. Грызуны предпочитали трудную, но «высокооплачиваемую» задачу лишь до тех пор, пока нагрузка не превышала объем их рабочей памяти. Сталкиваясь с непосильными требованиями, они переключались на

простую задачу с гарантированным маленьким подкреплением.

То есть, крысы продемонстрировали способность регулировать уровень неопределенности, учитывая собственные когнитивные возможности. Такие способности формируют особый тип поведения, целью которого является уменьшение внутренней неопределенности [44].

Живые организмы развиваются в сложной вероятностной среде, адаптироваться к которой только на основании индивидуального опыта невозможно.

Базовым механизмом психики, преодолевающим ограниченность обучения с подкреплением, является *активное построение внутренней модели среды*.

Это понятие значительно шире, чем сознательный логический вывод. Внутренняя модель является отражением разнообразия возможных состояний среды и Байесовских прогнозов вероятности для каждого из них [44].

Например, в зрительном восприятии часто встречающиеся комбинации признаков формируют в системе распознавания образов некий прототип, который оказывает сильное влияние на последующий процесс обработки зрительной информации. Воспринимая изображение, наблюдатель как бы ищет подтверждения уже существующей у него догадке, производя «бессознательное умозаключение», как назвал его Г. Гельмгольд [31].

Аналогичные процессы прогностического кодирования сопровождают распознавание устной и письменной речи [48]. В определенном смысле репрезентация внешних условий присуща любым самоорганизующимся системам, в частности, одноклеточным организмам.

Для объяснения саморегуляции поведения организмов нейробиология позаимствовала из кибернетики «теорему о хорошем регуляторе» [13]. Согласно ей, эффективная управляющая система должна являться адекватной моделью среды, с которой она взаимодействует.

Благодаря наличию прогностической модели, восприятие высоковероятных событий облегчается, а маловероятных затрудняется, что и порождает внутренний конфликт. Расхождения текущих ощущений с предварительным прогнозом ставит субъекта перед необходимостью пересмотра внутренней модели среды. Число и величина таких несоответствий задают степень неопределенности ситуации.

Считается, что всем организмам имманентно присуще стремление минимизировать эту разницу, поскольку неожиданность несет в себе потенциальную угрозу для сохранения целостности живой системы. Поэтому организмы активно отбирают (или с тем же результатом игнорируют) информацию из окружающей среды, для того чтобы уменьшить для себя ее неопределенность, или «свободную энергию» в терминологии К. Фристана [44].

В свете этой концепции любой выбор субъекта несет в себе не только прагматическую ценность

(немедленную награду или наказание), но и эпистемическую (повышение осведомленности).

Принятие решения в рамках стратегии исследования и стратегии использования (exploration-exploitation trade-off)

В целенаправленном поведении человека и животных можно выделить две диаметрально противоположных стратегии: использование (exploitation) и исследование (exploration) [4]. Стратегия использования — это применение знакомой выгодной опции. В рамках такой стратегии субъект раз за разом совершает один и тот же выбор, который по опыту оказался лучше всех остальных альтернатив. Этот выбор сопряжен с минимальной степенью неопределенности — т. е. гарантированно или с высокой вероятностью приносит вознаграждение [45].

Вторая стратегия — стратегия исследования — это поиск новых возможностей. В рамках такой стратегии субъект переключается между разными вариантами выбора, пробует незнакомые опции и имеет дело с высоким уровнем неопределенности. Тем самым стратегия исследования, в отличие от стратегии использования, позволяет получить новую информацию.

Экспериментальные парадигмы для изучения стратегий исследования и использования

В реальных ситуациях стратегии исследования и использования попеременно сменяют друг друга.

Стратегия использования обеспечивает приток положительных подкреплений в краткосрочной перспективе, а поиск новых возможностей открывает путь к отсроченной награде [8]. Например, копытным животным важно исследовать новые территории в летний сезон, когда доступно много разнообразной пищи. Даже не обнаружив ее в новом месте, они получают прогностически важную информацию. Зимой, когда цена ошибочного решения несопоставимо возрастает, животное будет опираться на накопленный за лето опыт. Повышенная склонность к риску и поиску новых впечатлений, характерные для подростков, расширяет их жизненный опыт, который может быть использован в зрелом возрасте [49]. Взаимные переходы между стратегиями использования и исследования наблюдаются в процессе решения практически любых новых задач.

В экспериментальных исследованиях принятия решения человеком часто используют модель тотализатора с двумя или несколькими «рычагами» («N-armed bandit task», «задача с многоруким бандитом») [55]. Испытуемому предлагается делать ставки, выбирая любой из двух или более «рычагов» (вместо «рычагов» могут быть стимулы на экране). Величина выигрыша при выборе каждой альтернативы варьируется в соответствии с кривой нормального распределения неза-

висимо от предыдущего выбора. Один рычаг систематически приносит больше прибыли, чем другие. Испытуемым необходимо эмпирически его определить, чтобы выиграть как можно больше очков или денег за фиксированное время.

В другом варианте задачи («restless bandit task») приоритетность рычагов постепенно меняется в течение игры. Наиболее выгодная в начале опция теряет свою ценность, уступая первенство другой. В этом случае испытуемым необходимо в подходящий момент перестроить стратегию выбора.

Типичный игрок в первых нескольких пробах хаотично переключается с одного рычага на другой, а затем у него вырабатывается устойчивое предпочтение наиболее выгодной альтернативы. Лишь иногда он возвращается к другим рычагам [4]. Обучившись — т. е. поняв, какая из опций выгоднее, — игрок переключается от стратегии исследования к стратегии использования и большую часть времени следует ей. Но возвращения к другим рычагам — случаи исследования — происходят и после обучения. В целом, выбор стратегии большинством испытуемых подтверждает философско-экономические размышления выдающегося исследователя XVIII века Даниэля Бернулли [1]: человек с готовностью рискует, пока ему практически нечего терять, однако по мере накопления дохода он все больше избегает неопределенности. Интересно, что в этой парадигме эксперимента люди и животные (макаки) вели себя сходным образом. У тех и других пропорция «исследовательских» выборов составляла около 25% [39].

Еще одной экспериментальной моделью для изучения стратегий исследования и использования является парадигма «наблюдай или делай ставку» (observe-or-bet) [47]. У испытуемого есть выбор: наблюдать исход, ассоциированный с выбранным стимулом, при этом не получая ни награды, ни штрафа, или выбрать стимул (сделать ставку), получив награду в случае правильного выбора. Первый вариант считается исследованием, второй — использованием. В одном из вариантов такого эксперимента перед испытуемым две лампочки; в каждой пробе он может либо просто наблюдать, какая лампочка загорится, в результате получая информацию о вероятностях включения каждой из двух лампочек, либо попробовать угадать, какая из лампочек загорится, и в случае правильного предположения он получает деньги, а в случае неправильного — теряет [9].

Третья экспериментальная парадигма для изучения баланса исследования/использования — это «задача с часами» (clock task) [21]. Испытуемым показывают циферблат часов, по которому движется стрелка; в одной пробе стрелка преодолевает пятисекундный интервал. Стрелку можно остановить в любом положении внутри интервала и получить награду. Величина награды и ее вероятность — функции от времени реакции. Например, чем дальше стрелка, т. е. чем больше время реакции, тем выше награда и ниже вероятность ее получить. В этом случае использованием считаются ответы с маленьким временем реакции, приносящие

маленькую гарантированную награду, а исследованием — ответы с увеличенным временем реакции, приносящие большую награду, но редко.

В целом, для экспериментального изучения стратегий исследования можно использовать практически любые задачи с вероятностной структурой обучения, любые «гемблинг-задачи», где вероятность выигрыша при выборе разных опций отличается.

Факторы, влияющие на баланс стратегий исследования и использования

Факторы, играющие роль в переключении между стратегиями исследования и использования, можно разделить на три группы: средовые, индивидуальные (личностные характеристики) и социальные [60].

К *факторам среды* можно отнести истощение ресурсов: если выбираемая ранее опция перестала быть выгодной (например, у крысы в кормушке закончилась еда, на поле кони съели всю траву), это толкает субъекта на то, чтобы переключиться со стратегии использования на стратегию исследования.

Кроме того, важным фактором является соотношение выгод и рисков исследования и использования (а также самого переключения). Так, животные меньше исследуют территории в поисках пищи, если высок риск нападения хищников [61].

Кроме того, важную роль играет стабильность/предсказуемость среды: если условия быстро меняются, это толкает к исследованию [23]. Влияет также доступная информация о разных вариантах: ситуация, когда выбор в рамках стратегии исследования несет только информацию (например, в парадигме «наблюдай или делай ставку»), отличается от ситуации, когда такой выбор несет информацию, но при этом сопряжен с рисками проигрыша (например, в парадигме многорукого бандита).

К *индивидуальным факторам* относятся когнитивные способности, текущее состояние организма, предыдущий опыт, психические характеристики; есть исследования, согласно которым склонность к исследованию снижается с возрастом [34], а также что девочки менее склонны к *стратегии исследования*, чем мальчики [52]. Кроме того, на баланс стратегий исследования—использования влияет уровень нейромедиаторов: у животных и у людей высокий уровень дофамина связан со сниженной исследовательской активностью, а низкий — с повышенной [24]. Индивиды с дефицитом навыков самоконтроля и планирования, нетерпимые к отсрочке награды (по результатам опросников) совершали меньше «исследовательских» выборов в задаче с многоруким бандитом [4]. Дети, которые в младенческом возрасте воспитывались в социальных учреждениях, также продемонстрировали дефицит поискового поведения [19]. Аналогичные тенденции описаны у испытуемых с никотиновой и алкогольной зависимостью [54].

Социальные факторы, влияющие на баланс исследования—использования у человека, рассматриваются достаточно редко — вероятно, отчасти по причине сложности изучения в условиях лаборатории.

Среди социальных животных склонность к исследованию либо использованию часто определяется социальной ролью: например, у муравьев и птиц есть особи, специализирующиеся на исследовании и использовании [14]: первые ищут новые источники пищи, а вторые пользуются уже найденными.

Физиологические механизмы стратегий исследования и использования

Нейроанатомия и теории, описывающие возможные механизмы исследования и использования

Одной из ключевых областей мозга, связанных с балансом исследования и использования, считается голубое пятно — основной источник норадреналина в головном мозге.

Эстон-Джонс с соавторами [7; 32] в экспериментах на обезьянах показали, что у голубого пятна есть два режима работы: фазический и тонический. Первый характеризуется умеренным уровнем активности нейронов голубого пятна и мощным ответом на целевые стимулы; на поведенческом уровне этот режим проявляется как отсутствие пропусков и очень низкий процент ложных тревог. В тоническом режиме фоновая активность голубого пятна выше, но фазический ответ на целевой стимул снижается или подавляется полностью. При этом повышается доля ложных тревог и увеличивается время реакции на целевые стимулы.

Предполагают, что фазический режим работы голубого пятна соответствует использованию, а тонический — исследованию [57]. В фазическом режиме в голубом пятне происходит выделение норадреналина только в ответ на релевантные задачи события (например, предъявление целевого стимула), что способствует обработке этих событий. При таком режиме работы голубого пятна субъект не «отвлекается» на посторонние стимулы и эффективно использует привычную опцию, т. е. реализует стратегию использования. В тоническом режиме норадреналин высвобождается постоянно и поддерживает обработку всех событий, не только релевантных — это дает возможность субъекту переключаться с базовой опции и пробовать новые возможности, т. е. применять стратегию исследования.

В пользу этой модели говорят результаты двух работ, где использовали фМРТ: в одной показали, что активность голубого пятна выше при исследовании, чем при использовании [59], в другой обнаружили повышенную активность в области ствола головного мозга [54] и предположили, что источником этой активности может быть голубое пятно [4], однако нужно учитывать, что точность локализации активности стволовых структур может быть невысокой.

Возникает вопрос: что заставляет голубое пятно переключаться из одного режима в другой? Что служит сигналом для перехода от стратегии использования к стратегии исследования? Эстон-Джонс и Коэн [6] предположили, что сигналы к переключению поступают к голубому пятну из вентральных и медиальных фронтальных структур — в частности, передней поясной коры (anterior cingulate cortex, ACC), активность которой связана с обработкой ошибок, конфликтов, негативной обратной связи и т. д. [11].

При применении субъектом стратегии использования одна из альтернатив получает статус «выгодной», другая — «невыгодной». В этом случае принятие решения, по сути, уже не включает в себя рассмотрение альтернатив: происходит лишь выбор «выгодной» альтернативы. В отличие от стратегии использования, при переходе к стратегии исследования возникает конфликт между решениями: выбирать либо заведомо «выгодную» альтернативу, либо «невыгодную» в расчете на получение информации, потенциально полезной в будущем.

В рамках дальнейшего развития модели МакКлёр с соавторами [35] предположили, что сигналы о переключении приходят в голубое пятно из передней поясной коры (ACC) и орбитофронтальной коры (OFC) — двух основных входов голубого пятна. OFC оценивает результат выбора, величину полученной награды, и в зависимости от этих параметров регулирует работу голубого пятна.

Известно, что активность орбитофронтальной коры связана с вариабельностью ответов [43]. Повреждение OFC приводит к ухудшению выполнения задач на переучивание («reversal learning task») [29], в которых требуется переключаться с усвоенного правила на новое; это также дает основания предполагать участие OFC в переходе от использования к исследованию. Передняя поясная кора активна при конфликте реакций. Авторы считают, что перед переходом к стратегии исследования конфликт между вариантами выбора, т. е. между реакциями, нарастает, и при продолжительном конфликте ACC отправляет голубому пятну сигнал о переключении на стратегию исследования.

Если ACC действительно отправляет сигналы о переходе к исследованию в голубое пятно, то не вполне ясно, как именно ее активность соотносится во времени с переключением голубого пятна из одного режима в другой. Возможно, активность этой области связана не с переключением (или не только с ним), а с поддержанием одного из режимов. Так, Бланчард и Гершман в фМРТ-исследовании в парадигме «наблюдай или делай ставку» показали, что активность передней поясной коры (а также островковой доли) выше во время режима исследования — именно во время, а не перед переключением [9].

Важная область, участвующая в регуляции баланса исследования—использования — фронтополярная кора (поле 10, FPC).

В фМРТ-исследовании с применением поведенческой методики «четырёхрукий бандит» Доу с соавтора-

ми показали, что фронтальная кора (FPC) и внутрименная борозда (IPS) более активны при исследовании [15], а стриатум и вентромедиальная префронтальная кора (vmPFC) — при использовании.

Бочин с соавторами показали, что у макак при разрушении фронтальной коры нарушаются быстрое научение с одного повтора и исследовательское поведение [10]. Авторы предполагают, что FPC не просто связана с «исследовательскими» выборами, а отвечает за первоначальную оценку выгод разных новых альтернатив. Даже если о других альтернативах ничего не известно, FPC аккумулирует историю выборов: при получении обратной связи туда «записывается», насколько успешным было решение.

Авторы считают, что FPC накапливает свидетельства в пользу перехода к стратегии исследования, а затем отправляет сигнал о переключении другим областям коры.

Мансури с соавторами в подробном обзоре, посвященном функциям фронтальной коры, тоже предполагают, что у людей она играет ключевую роль в осуществлении стратегии исследования, причем медиальная фронтальная кора отвечает за случайное исследование, а латеральная — за целенаправленное [33]; а за использование отвечает задняя часть префронтальной коры.

Лаурейро-Мартинез с соавторами [31] реплицировали эксперимент Дуу на большей выборке, также используя фМРТ, и тоже обнаружили активацию медиальной префронтальной коры во время использования, а IPS и FPC — во время исследования. Но помимо этого в случае использования обнаружили активацию гиппокампа, а в случае стратегии исследования — целого ряда зон, связанных с вниманием: в теменной области — височно-теменное соединение (temporo-parietal junction, TPJ), верхняя теменная доля (superior parietal lobule, SPL); во фронтальных областях — переднее глазодвигательное поле (FEF), среднефронтальная извилина (MFG), передняя поясная кора (dACC), часть дополнительной моторной области (pre-SMA). Кроме того, авторы наблюдали, что при исследовании сильнее активируется голубое пятно, что согласуется с теорией Эстон-Джонса и Коэна [6]. Помимо этого, в режиме исследования наблюдалась активация вентральной фронтальной коры (VFC). Авторы предположили, что стратегия использования связана с большей активацией областей мозга, входящих в систему подкрепления (в том числе гиппокампа), а стратегия исследования — с активацией областей, отвечающих за когнитивный контроль (в частности, фронтальной коры — FPC) [36]).

Это предположение логично, поскольку для переключения на исследование необходимо подавить преобладающую реакцию, привычную и более выгодную, которая соответствует стратегии использования. Контроль внимания отчасти обеспечивается активностью голубого пятна (LC) [51], а активность голубого пятна, в свою очередь, регулируется входами от дофаминэргических мезокортиколимбических областей.

Еще одна зона, потенциально участвующая в балансе исследования—использования — это ростолатеральная префронтальная кора, которая сильнее активируется при переходе от стратегии использования к стратегии исследования под влиянием неопределенности [50].

Таким образом, в регуляции баланса исследования и использования задействована сеть из многих областей, основные из которых — голубое пятно, передняя поясная кора, фронтальная кора, орбитофронтальная кора, префронтальная кора.

Электрофизиологические методы изучения баланса исследования—использования

Из областей мозга, задействованных в балансе исследования—использования, наиболее проста и перспективна для изучения методами электрофизиологии передняя поясная кора. В изучении роли этой области можно опираться на данные анализа частотно-временной активности и потенциалов, связанных с событиями. Одним из подходов является изучение тета-осцилляций фронтальной средней линии: их главным источником является передняя поясная кора, и они имеют максимальную выраженность во фронтальных отведениях ЭЭГ. Тета-осцилляции фронтальной средней линии считают отражением работы механизма детекции необходимости включения (или усиления) когнитивного контроля [12]. Этот механизм может включаться в нескольких случаях: в условиях новизны или неопределенности, в условиях конфликта, после ошибок, «наказания» или «проигрыша». Все эти ситуации требуют адаптивных изменений поведения. Вариантов таких адаптивных изменений может быть несколько: можно «активировать когнитивные ресурсы», усилить когнитивный контроль (например, внимательнее выполнять задачу), оставшись на уже выбранной стратегии, а можно уйти от выбранной стратегии и попробовать новую, т. е. перейти в режим исследования.

Мощность тета-осцилляций фронтальной средней линии можно рассматривать как предиктор перехода к стратегии исследования: Чжан с соавторами показали, что мощность этого ритма увеличивается в интервале 250—500 мс после обратной связи и возрастает сильнее в реализациях, предшествующих смене выбираемого стимула [30]. Каванах с соавторами также показали, что повышение мощности фронтальных тета-осцилляций предшествует переходу к стратегии исследования [17].

Помимо тета-осцилляций фронтальной средней линии, в качестве коррелятов активности передней поясной коры можно изучать ряд потенциалов, связанных с событиями, таких как N2, ERN и FRN — генератором всех этих компонентов считается передняя поясная кора [22; 25]. Компонент N2 — это негативная волна с латентностью 200—300 мс относительно момента предъявления стимула, которая имеет локализацию на фронтальной средней линии; N2 считают отражением конфликта между репрезентациями — чем сильнее конфликт, тем больше амплитуда N2 [38; 46].

Компонент ERN (англ. error-related negativity, негативность, связанная с ошибкой) — негативная волна, возникающая после совершения ошибки, с пиком в интервале 50—100 мс и с локализацией на фронтальной средней линии [3]. Данный компонент является отражением внутренней детекции ошибки.

Компонент FRN (англ. feedback-related negativity, негативность в ответ на обратную связь, также известная как MFN, medial frontal negativity) — это негативная волна, возникающая в ответ на обратную связь с пиком около 250 мс с момента ее предъявления [37]. FRN максимально выражена в отведениях фронтальной средней линии и отражает ошибку предсказания — расхождение предполагаемого исхода (например, ожидаемого знака обратной связи) с фактическим [25].

Другая группа методов изучения баланса исследования—использования фокусируется на нейромедиаторах. В следующих разделах мы рассмотрим такие методы и результаты исследований в этом направлении.

Дофамин связан со стратегией использования

В современных теориях, описывающих механизмы научения и принятия решений, важнейшая роль отводится дофамину, но его функции намного шире — в частности, он участвует в регуляции баланса использования—исследования.

Выделяют два основных вида активности дофаминэргических нейронов: тоническая активность — т. е. фоновая, постоянная; фазическая активность — быстрая, в ответ на те или иные события. В соответствии с этими видами активности выделяют два пула дофамина — тонический, концентрация которого меняется медленно (за минуты), и фазический, который выделяется в результате фазической активности дофаминэргических нейронов; концентрация в нем меняется быстро (меньше чем за секунду) [42].

Одной из ключевых частей механизма научения считается ошибка предсказания награды (reward prediction error) — разница между ожидаемой и полученной наградой, которая на нейрофизиологическом уровне отражается в фазическом выделении дофамина в вентральной области покрышки (VTA).

Помимо участия в научении есть свидетельства в пользу связи дофамина с избеганием риска [42], а также его участия в регуляции баланса исследования и использования. Считается, что чем выше уровень тонического дофамина, тем больше склонность к стратегии использования, чем ниже — тем больше склонность к стратегии исследования. Так, при блокаде дофаминовых рецепторов D1 и D2, т. е. при снижении чувствительности к дофамину в префронтальной коре, у крыс повышается процент «исследовательских» выборов в задаче с трехруким бандитом [17], при том что скорость обучения не меняется. Однако, судя по всему, рецепторы D1 и D2 в этом смысле работают по-разному: у крыс предпочтение риска, что грубо можно приравнять к предпочтению исследования, при блокаде D1-рецепторов снижалось, а при блокаде D2-рецепторов повышалось [41].

У мышей с повышенной внеклеточной концентрацией «тонического» дофамина наблюдается снижение гибкости поведения; кроме того, они готовы выполнять большую, чем обычные мыши, работу за то же самое вознаграждение [2]. Разницу в поведении, связанную с концентрацией внеклеточного дофамина, исследовали и на людях; правда, результаты противоречивы. Так, Франк с соавторами показали, что обладатели менее активной формы COMT (фермента, разрушающего дофамин в синаптической щели и действующего в префронтальной коре), гомозиготы Met/Met, т. е. с повышенной концентрацией дофамина в префронтальной коре, при выполнении «задачи с часами» были более склонны к исследованию по сравнению с теми, у кого активность COMT была выше (Met/Val и Val/Val) [21]. В другом исследовании на людях этот результат не воспроизвели, зато обнаружили, что введение ингибитора COMT, толкапона, приводит к повышению доли «исследовательских» выборов у гомозигот Met/Met, а у остальных генотипов не приводит [18]. То есть, видимо, и так низкая активность COMT у таких гомозигот подавляется настолько, что это проявляется на уровне поведения, а у других генотипов подавляется недостаточно сильно и на уровне поведения не проявляется.

Норадреналин: противоречивые данные о роли в балансе исследования-использования

Норадреналин играет ключевую роль в модели Эстона—Джонса, согласно которой исследование и использование обеспечиваются двумя режимами работы голубого пятна.

Под действием норадреналина увеличивается диаметр зрачка, и при постоянном уровне освещения его размер можно использовать для непрямой оценки тонической активности голубого пятна [26]: чем больше зрачок, тем выше активность. Джемма и Нивенхейс показали, что перед «исследовательскими» выборами в задаче с четырехруким бандитом диаметр зрачка у людей увеличен по сравнению с выборами, относящимися к использованию [26]. Однако введение ребоксетина, селективного ингибитора обратного захвата норадреналина, который повышает концентрацию внеклеточного норадреналина, в такой же задаче не повлияло на баланс исследования—использования [58]. А Уоррен с соавторами получили, что, в противоположность их первоначальной гипотезе, при введении атомоксетина (блокатора транспортера норадреналина, который должен повышать внеклеточную концентрацию норадреналина в коре) склонность к исследованию у людей снижалась [56], а не повышалась, как можно было бы предположить на основании модели Эстона—Джонса.

Ацетилхолин оказывает на баланс исследования-использования эффекты, сходные с дофамином

Роль этого нейромедиатора в балансе исследования—использования изучена хуже других, но некоторые данные

позволяют предположить, что он тоже вовлечен в данные процессы. Так, никотиновые ацетилхолиновые рецепторы (nAChRs) в вентральной области покрышки (VTA) могут участвовать в осуществлении стратегии исследования: в задаче с многоруким бандитом мыши, у которых работа этого рецептора была нарушена, демонстрировали большую по сравнению с диким типом склонность к рискованным выборам с высокой неопределенностью, что фактически эквивалентно склонности к исследованию.

Влияние ацетилхолина на принятие решений сходно с дофамином [20]; это может объясняться тем, что ацетилхолин воздействует на дофаминэргические нейроны VTA через никотиновые ацетилхолиновые рецепторы [40].

Заключение

Мы обсудили вопрос о том, что представляет собой неопределенность, и рассмотрели способы принятия решения в условиях неопределенности.

Одним из таких способов является стратегия исследования — поиск и проверка новых возможностей. Противоположная стратегия — стратегия использования — становится оправданной, когда неопределенность снижается.

В реальной жизни эти стратегии сменяют друг друга, адаптивное поведение предполагает переключение с одной на другую.

На переключение влияет ряд факторов — индивидуальные, социальные, факторы среды.

Баланс этих двух стратегий — исследования и использования — является предметом активного изучения в различных областях наук, и к настоящему времени накоплен массивный объем данных о нейрофизиологии процессов, лежащих в основе этого баланса. Стратегиям исследования и использования соответствуют два режима работы голубого пятна — соответственно тонический и фазический.

Стратегия исследования сопряжена с повышенным фоновым уровнем норадреналина, а стратегия использования — с пониженным, однако выбросы норадреналина происходят локально в ответ на релевантные события (например, наиболее выгодные стимулы). Переход голубого пятна из фазического режима в тонический инициируется сигналами от передней поясной коры, фронтополярной коры, орбитофронтальной коры, которые, по всей видимости, оценивают разные аспекты ситуации и «приводят разные доводы» в пользу перехода от стратегии использования к стратегии исследования. Помимо норадренэргической системы, в регуляцию баланса исследования и использования вовлечены дофаминэргическая и ацетилхолинэргическая системы. Повышенный уровень дофамина, как и ацетилхолина, приводит к повышению склонности к использованию.

В целом, несмотря на значительный массив накопленных данных, в теме «баланс стратегий исследования и использования» остается еще много вопросов: в частности, не выяснена временная динамика активности вовлеченных в регуляцию баланса областей мозга.

Литература

1. Канеман Д., Тверски А. Рациональный выбор, ценности и фреймы // Психологический журнал. 2003. Т. 24. № 4. С. 31—43.
2. A kinder, gentler dopamine... highlighting dopamine's role in behavioral flexibility / J.A. Beeler [et al.] // *Frontiers in Neuroscience*. 2014. Vol. 8. Article ID 4. 2 p. DOI:10.3389/fnins.2014.00004
3. A neural system for error detection and compensation / W.J. Gehring [et al.] // *Psychological Science*. 1993. Vol. 4. № 6. P. 385—390. DOI:10.1111/j.1467-9280.1993.tb00586.x
4. A primer on foraging and the explore/exploit trade-off for psychiatry research / M.A. Addicott [et al.] // *Neuropsychopharmacology*. 2017. Vol. 42. P. 1931—1939. DOI:10.1038/npp.2017.108
5. *Aspers P.* Forms of uncertainty reduction: decision, valuation, and contest // *Theory and Society*. 2018. Vol. 47. P. 133—149. DOI:10.1007/s11186-018-9311-0
6. *Aston-Jones G., Cohen J.D.* An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance // *Annual Review of Neuroscience*. 2005. Vol. 28. P. 403—450. DOI:10.1146/annurev.neuro.28.061604.135709
7. *Aston-Jones G., Rajkowski J., Kubiak P.* Conditioned responses of monkey locus coeruleus neurons anticipate acquisition of discriminative behavior in a vigilance task // *Neuroscience*. 1997. Vol. 80. № 3. P. 697—715. DOI:10.1016/S0306-4522(97)00060-2
8. *Barack D.L., Gold J.I.* Temporal trade-offs in psychophysics // *Current Opinion in Neurobiology*. 2016. Vol. 37. P. 121—125. DOI:10.1016/j.conb.2016.01.015
9. *Blanchard V.C., Gershman S.J.* Pure correlates of exploration and exploitation in the human brain // *Cognitive, Affective, & Behavioral Neuroscience*. 2018. Vol. 18. № 1. P. 117—126. DOI:10.3758/s13415-017-0556-2
10. *Boschin E.A., Piekema C., Buckley M.J.* Essential functions of primate frontopolar cortex in cognition // *Proceedings of the National Academy of Sciences*. 2015. Vol. 112. № 9. P. E1020—E1027. DOI:10.1073/pnas.1419649112
11. *Botvinick M.M., Cohen J.D., Carter C.S.* Conflict monitoring and anterior cingulate cortex: an update // *Trends in Cognitive Sciences*. 2004. Vol. 8. № 12. P. 539—546. DOI:10.1016/j.tics.2004.10.003
12. *Cavanagh J.F., Frank M.J.* Frontal theta as a mechanism for cognitive control // *Trends in Cognitive Sciences*. 2014. Vol. 18. № 8. P. 414—421. DOI:10.1016/j.tics.2014.04.012
13. *Conant R.C., Ross Ashby W.* Every good regulator of a system must be a model of that system // *International Journal of Systems Science*. 1970. Vol. 1. № 2. P. 89—97. DOI:10.1080/00207727008920220

14. Cook Z., Franks D.W., Robinson E.J.H. Exploration versus exploitation in polydomous ant colonies // *Journal of Theoretical Biology*. 2013. Vol. 323. P. 49—56. DOI:10.1016/j.jtbi.2013.01.022
15. Cortical substrates for exploratory decisions in humans / N.D. Daw [et al.] // *Nature*. 2006. Vol. 441. P. 876—879. DOI:10.1038/nature04766
16. Denison S., Xu F. Infant statisticians: The origins of reasoning under uncertainty // *Perspectives on Psychological Science*. 2019. Vol. 14. № 4. P. 499—509. DOI:10.1177/1745691619847201
17. Dopamine blockade impairs the exploration-exploitation trade-off in rats / F. Cinotti [et al.] // *Scientific Reports*. 2019. Vol. 9. № 1. P. 1—14. DOI:10.1038/s41598-019-43245-z
18. Dopamine, locus of control, and the exploration-exploitation tradeoff / A.S. Kayser [et al.] // *Neuropsychopharmacology*. 2015. Vol. 40. № 2. P. 454—462. DOI:10.1038/npp.2014.193
19. Exploration—Exploitation strategy is dependent on early experience / K.L. Humphreys [et al.] // *Developmental Psychobiology*. 2015. Vol. 57. № 3. P. 313—321. DOI:10.1002/dev.21293
20. Fobbs W.C., Mizumori S.J.Y. Cost—Benefit Decision Circuitry: Proposed Modulatory Role for Acetylcholine // *Progress in Molecular Biology and Translational Science*. 2014. Vol. 122. P. 233—261. DOI:10.1016/B978-0-12-420170-5.00009-X
21. Frank M.J., Hutchison K. Genetic contributions to avoidance-based decisions: striatal D2 receptor polymorphisms // *Neuroscience*. 2009. Vol. 164. № 1. P. 131—140. DOI:10.1016/j.neuroscience.2009.04.048
22. Gehring W.J., Willoughby A.R. The medial frontal cortex and the rapid processing of monetary gains and losses // *Science*. 2002. Vol. 295. № 5563. P. 2279—2282. DOI:10.1126/science.1066893
23. Gold J.I., Shadlen M.N. The neural basis of decision making // *Annual review of neuroscience*. 2007. Vol. 30. P. 535—574. DOI:10.1146/annurev.neuro.29.051605.113038
24. Hills V.V. Animal foraging and the evolution of goal-directed cognition // *Cognitive Science*. 2006. Vol. 30. № 1. P. 3—41. DOI:10.1207/s15516709cog0000_50
25. Huang Y., Yu R. The feedback-related negativity reflects «more or less» prediction error in appetitive and aversive conditions // *Frontiers in Neuroscience*. 2014. Vol. 8. Article ID 108. 6 p. DOI:10.3389/fnins.2014.00108
26. Jepma M., Nieuwenhuis S. Pupil diameter predicts changes in the exploration—exploitation trade-off: Evidence for the adaptive gain theory // *Journal of Cognitive Neuroscience*. 2011. Vol. 23. № 7. P. 1587—1596. DOI:10.1162/jocn.2010.21548
27. Kahneman D., Tversky A. Variants of uncertainty // *Cognition*. 1982. Vol. 11. № 2. P. 143—157. DOI:10.1016/0010-0277(82)90023-3
28. Killeen P.R. Pavlov + Skinner = Premack [Электронный ресурс] // *International Journal of Comparative Psychology*. 2014. Vol. 27. № 4. P. 544—568. URL: https://www.researchgate.net/profile/Peter_Killeen2/publication/269873794_Pavlov_Skinner_Premack/links/549861d30cf2c5a7e342bdca.pdf (дата обращения: 05.06.2020).
29. Learning theory: a driving force in understanding orbitofrontal function / M.A. McDannald [et al.] // *Neurobiology of Learning and Memory*. 2014. Vol. 108. P. 22—27. DOI:10.1016/j.nlm.2013.06.003
30. Linking brain electrical signals elicited by current outcomes with future risk decision-making / D. Zhang [et al.] // *Frontiers in Behavioral Neuroscience*. 2014. Vol. 8. Article ID 34. 15 p. DOI:10.3389/fnbeh.2014.00084
31. Linson A., Parr V., Friston K.J. Active inference, stressors, and psychological trauma: A neuroethological model of (mal) adaptive explore-exploit dynamics in ecological context // *Behavioural Brain Research*. 2020. Vol. 380. P. 112—421. DOI:10.1016/j.bbr.2019.112421
32. Locus coeruleus neurons in monkey are selectively activated by attended cues in a vigilance task / G. Aston-Jones [et al.] // *Journal of Neuroscience*. 1994. Vol. 14. № 7. P. 4467—4480. DOI:10.1523/JNEUROSCI.14-07-04467.1994
33. Managing competing goals — a key role for the frontopolar cortex / F.A. Mansouri [et al.] // *Nature Reviews Neuroscience*. 2017. Vol. 18. № 11. P. 645—657. DOI:10.1038/nrn.2017.111
34. Mata R., Wilke A., Czienskowski U. Foraging across the life span: is there a reduction in exploration with aging? // *Frontiers in Neuroscience*. 2013. Vol. 7. Article ID 53. 7 p. DOI:10.3389/fnins.2013.00053
35. McClure S.M., Gilzenrat M.S., Cohen J.D. An exploration-exploitation model based on norepinephrine and dopamine activity [Электронный ресурс] // *Advances in neural information processing systems: proceedings from the conference “Neural Information Processing Systems 2005”* / Eds. Y. Weiss, B. Schölkopf, J.C. Platt. Neural Information Processing Systems Foundation, 2006. P. 867—874. URL: <https://papers.nips.cc/paper/2950-an-exploration-exploitation-model-based-on-norepinephrine-and-dopamine-activity.pdf> (дата обращения: 05.06.2020).
36. Miller E.K., Cohen J.D. An integrative theory of prefrontal cortex function // *Annual Review of Neuroscience*. 2001. Vol. 24. P. 167—202. DOI:10.1146/annurev.neuro.24.1.167
37. Miltner W.H.R., Braun C.H., Coles M.G.H. Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a “generic” neural system for error detection // *Journal of Cognitive Neuroscience*. 1997. Vol. 9. № 6. P. 788—798. DOI:10.1162/jocn.1997.9.6.788
38. N200 in the Eriksen-task: Inhibitory executive process? / M. Heil [et al.] // *Journal of Psychophysiology*. 2000. Vol. 14. № 4. P. 218—225. DOI:10.1027//0269-8803.14.4.218
39. Neurons in posterior cingulate cortex signal exploratory decisions in a dynamic multioption choice task / J.M. Pearson [et al.] // *Current biology*. 2009. Vol. 19. № 18. P. 1532—1537. DOI:10.1016/j.cub.2009.07.048

40. Nicotinic receptors in the ventral tegmental area promote uncertainty-seeking / J. Naudé [et al.] // *Nature Neuroscience*. 2016. Vol. 19. № 3. P. 471—478. DOI:10.1038/nn.4223
41. Onge J.R.S., Abhari H., Floresco S.B. Dissociable contributions by prefrontal D1 and D2 receptors to risk-based decision making // *Journal of Neuroscience*. 2011. Vol. 31. № 23. P. 8625—8633. DOI:10.1523/JNEUROSCI.1020-11.2011
42. Overriding phasic dopamine signals redirects action selection during risk/reward decision making / C.M. Stopper [et al.] // *Neuron*. 2014. Vol. 84. № 1. P. 177—189. DOI:10.1016/j.neuron.2014.08.033
43. Padoa-Schioppa C., Conen K.E. Orbitofrontal cortex: a neural circuit for economic decisions // *Neuron*. 2017. Vol. 96. № 4. P. 736—754. DOI:10.1016/j.neuron.2017.09.031
44. Parr V., Friston K.J. Uncertainty, epistemics and active inference // *Journal of The Royal Society Interface*. 2017. Vol. 14. № 136. 10 p. DOI:10.1098/rsif.2017.0376
45. Psychological models of human and optimal performance in bandit problems / M.D. Lee [et al.] // *Cognitive Systems Research*. 2011. Vol. 12. № 2. P. 164—174. DOI:10.1016/j.cogsys.2010.07.007
46. Psychophysiological evidence of response conflict and strategic control of responses in affective priming / B.D. Bartholow [et al.] // *Journal of Experimental Social Psychology*. 2009. Vol. 45. № 4. P. 655—666. DOI:10.1016/j.jesp.2009.02.015
47. Rakow V., Newell B.R., Zougkou K. The role of working memory in information acquisition and decision making: Lessons from the binary prediction task // *The Quarterly Journal of Experimental Psychology*. 2010. Vol. 63. № 7. P. 1335—1360. DOI:10.1080/17470210903357945
48. Recognizing sequences of sequences / S.J. Kiebel [et al.] // *PLoS Computational Biology*. 2009. Vol. 5. № 8. 14 p. DOI:10.1371/journal.pcbi.1000464
49. Risk-taking behavior in adolescent mice: psychobiological determinants and early epigenetic influence / G. Laviola [et al.] // *Neuroscience & Biobehavioral Reviews*. 2003. Vol. 27. № 1—2. P. 19—31. DOI:10.1016/S0149-7634(03)00006-X
50. Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration / D. Badre [et al.] // *Neuron*. 2012. Vol. 73. № 3. P. 595—607. DOI:10.1016/j.neuron.2011.12.025
51. Sara S.J. The locus coeruleus and noradrenergic modulation of cognition // *Nature Reviews Neuroscience*. 2009. Vol. 10. № 3. P. 211—223. DOI:10.1038/nrn2573
52. Slovic P. Risk-taking in children: Age and sex differences // *Child Development*. 1966. Vol. 37. № 1. P. 169—176. DOI:10.2307/1126437
53. Smith A.P., Beckmann J.S., Zentall V.R. Gambling-like behavior in pigeons: 'jackpot' signals promote maladaptive risky choice // *Scientific Reports*. 2017. Vol. 7. № 1. P. 1—11. DOI:10.1038/s41598-017-06641-x
54. Smoking and the bandit: A preliminary study of smoker and nonsmoker differences in exploratory behavior measured with a multiarmed bandit task / M.A. Addicott [et al.] // *Experimental and Clinical Psychopharmacology*. 2013. Vol. 21. № 1. P. 66—73. DOI:10.1037/a0030843
55. Steyvers M., Lee M.D., Wagenmakers E.J. A Bayesian analysis of human decision-making on bandit problems // *Journal of Mathematical Psychology*. 2009. Vol. 53. № 3. P. 168—179. DOI:10.1016/j.jmp.2008.11.002
56. The effect of atomoxetine on random and directed exploration in humans / C.M. Warren [et al.] // *PloS One*. 2017. Vol. 12. № 4. 17 p. DOI:10.1371/journal.pone.0176034
57. The role of locus coeruleus in the regulation of cognitive performance / M. Usher [et al.] // *Science*. 1999. Vol. 283. № 5401. P. 549—554. DOI:10.1126/science.283.5401.549
58. The role of the noradrenergic system in the exploration-exploitation trade-off: a pharmacological study / M. Jepma [et al.] // *Frontiers in Human Neuroscience*. 2010. Vol. 4. Article ID 170. 13 p. DOI:10.3389/fnhum.2010.00170
59. Understanding the exploration—exploitation dilemma: An fMRI study of attention control and decision-making performance / D. Laureiro-Martínez [et al.] // *Strategic Management Journal*. 2015. Vol. 36. № 3. P. 319—338. DOI:10.1002/smj.2221
60. Unpacking the exploration—exploitation tradeoff: A synthesis of human and animal literatures / K. Mehlhorn [et al.] // *Decision*. 2015. Vol. 2. № 3. P. 191—215. DOI:10.1037/dec0000033
61. Verdolin J.L. Meta-analysis of foraging and predation risk trade-offs in terrestrial systems // *Behavioral Ecology and Sociobiology*. 2006. Vol. 60. № 4. P. 457—464. DOI:10.1007/s00265-006-0172-6
62. Yuki S., Okanoya K. Rats show adaptive choice in a metacognitive task with high uncertainty // *Journal of Experimental Psychology: Animal Learning and Cognition*. 2017. Vol. 43. № 1. P. 109—118. DOI:10.1037/xan0000130
63. Zentall V.R. An animal model of human gambling based on pigeon suboptimal choice [Электронный ресурс] // *Research & Reviews: Neuroscience*. 2017. Vol. 1. № 2. P. 27—37. URL: <https://pdfs.semanticscholar.org/f4ba/8ebce42ca058e780c9afb1322b7440bc8649.pdf> (дата обращения: 05.06.2020).
64. Zentall V.R. Suboptimal choice by pigeons: An analog of human gambling behavior // *Behavioural Processes*. 2014. Vol. 103. P. 156—164. DOI:10.1016/j.beproc.2013.11.004

References

1. Kaneman D., Tverski A. Ratsional'nyi vybor, tsennosti i freimy [Rational choice, values and frames]. *Psikhologicheskii zhurnal [Psychological journal]*, 2003. Vol. 24, no. 4, pp. 31—43.

2. Beeler J.A. et al. A kinder, gentler dopamine... highlighting dopamine's role in behavioral flexibility. *Frontiers in neuroscience*, 2014. Vol. 8, article ID 4, 2 p. DOI:10.3389/fnins.2014.00004
3. Gehring W.J. et al. A neural system for error detection and compensation. *Psychological science*, 1993. Vol. 4, no. 6, pp. 385—390. DOI:10.1111/j.1467-9280.1993.tb00586.x
4. Addicott M.A. et al. A primer on foraging and the explore/exploit trade-off for psychiatry research. *Neuropsychopharmacology*, 2017. Vol. 42, pp. 1931—1939. DOI:10.1038/npp.2017.108
5. Aspers P. Forms of uncertainty reduction: decision, valuation, and contest. *Theory and society*, 2018. Vol. 47, pp. 133—149. DOI:10.1007/s11186-018-9311-0
6. Aston-Jones G., Cohen J.D. An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annual Review of Neuroscience*, 2005. Vol. 28, pp. 403—450. DOI:10.1146/annurev.neuro.28.061604.135709
7. Aston-Jones G., Rajkowski J., Kubiak P. Conditioned responses of monkey locus coeruleus neurons anticipate acquisition of discriminative behavior in a vigilance task. *Neuroscience*, 1997. Vol. 80, no. 3, pp. 697—715. DOI:10.1016/S0306-4522(97)00060-2
8. Barack D.L., Gold J.I. Temporal trade-offs in psychophysics. *Current opinion in neurobiology*, 2016. Vol. 37, pp. 121—125. DOI:10.1016/j.conb.2016.01.015
9. Blanchard V.C., Gershman S.J. Pure correlates of exploration and exploitation in the human brain. *Cognitive, Affective, & Behavioral Neuroscience*, 2018. Vol. 18, no. 1, pp. 117—126. DOI:10.3758/s13415-017-0556-2
10. Boschini E.A., Piekema C., Buckley M.J. Essential functions of primate frontopolar cortex in cognition. *Proceedings of the National Academy of Sciences*, 2015. Vol. 112, no. 9, pp. E1020—E1027. DOI:10.1073/pnas.1419649112
11. Botvinick M.M., Cohen J.D., Carter C.S. Conflict monitoring and anterior cingulate cortex: an update. *Trends in cognitive sciences*, 2004. Vol. 8, no. 12, pp. 539—546. DOI:10.1016/j.tics.2004.10.003
12. Cavanagh J.F., Frank M.J. Frontal theta as a mechanism for cognitive control. *Trends in cognitive sciences*, 2014. Vol. 18, no. 8, pp. 414—421. DOI:10.1016/j.tics.2014.04.012
13. Conant R.C., Ross Ashby W. Every good regulator of a system must be a model of that system. *International journal of systems science*, 1970. Vol. 1, no. 2, pp. 89—97. DOI:10.1080/00207727008920220
14. Cook Z., Franks D.W., Robinson E.J.H. Exploration versus exploitation in polydomous ant colonies. *Journal of theoretical biology*, 2013. Vol. 323, pp. 49—56. DOI:10.1016/j.jtbi.2013.01.022
15. Daw N.D. et al. Cortical substrates for exploratory decisions in humans. *Nature*, 2006. Vol. 441, pp. 876—879. DOI:10.1038/nature04766
16. Denison S., Xu F. Infant statisticians: The origins of reasoning under uncertainty. *Perspectives on Psychological Science*, 2019. Vol. 14, no. 4, pp. 499—509. DOI:10.1177/1745691619847201
17. Cinotti F. et al. Dopamine blockade impairs the exploration-exploitation trade-off in rats. *Scientific reports*, 2019. Vol. 9, no. 1, pp. 1—14. DOI:10.1038/s41598-019-43245-z
18. Kayser A.S. et al. Dopamine, locus of control, and the exploration-exploitation tradeoff. *Neuropsychopharmacology*, 2015. Vol. 40, no. 2, pp. 454—462. DOI:10.1038/npp.2014.193
19. Humphreys K.L. et al. Exploration—Exploitation strategy is dependent on early experience. *Developmental Psychobiology*, 2015. Vol. 57, no. 3, pp. 313—321. DOI:10.1002/dev.21293
20. Fobbs W.C., Mizumori S.J.Y. Cost—Benefit Decision Circuitry: Proposed Modulatory Role for Acetylcholine. *Progress in molecular biology and translational science*, 2014. Vol. 122, pp. 233—261. DOI:10.1016/B978-0-12-420170-5.00009-X
21. Frank M.J., Hutchison K. Genetic contributions to avoidance-based decisions: striatal D2 receptor polymorphisms. *Neuroscience*, 2009. Vol. 164, no. 1, pp. 131—140. DOI:10.1016/j.neuroscience.2009.04.048
22. Gehring W.J., Willoughby A.R. The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, 2002. Vol. 295, no. 5563, pp. 2279—2282. DOI:10.1126/science.1066893
23. Gold J.I., Shadlen M.N. The neural basis of decision making. *Annual review of neuroscience*, 2007. Vol. 30, pp. 535—574. DOI:10.1146/annurev.neuro.29.051605.113038
24. Hills V.V. Animal foraging and the evolution of goal-directed cognition. *Cognitive science*, 2006. Vol. 30, no. 1, pp. 3—41. DOI:10.1207/s15516709cog0000_50
25. Huang Y., Yu R. The feedback-related negativity reflects “more or less” prediction error in appetitive and aversive conditions. *Frontiers in neuroscience*, 2014. Vol. 8, article ID 108, 6 p. DOI:10.3389/fnins.2014.00108
26. Jepma M., Nieuwenhuis S. Pupil diameter predicts changes in the exploration—exploitation trade-off: Evidence for the adaptive gain theory. *Journal of cognitive neuroscience*, 2011. Vol. 23, no. 7, pp. 1587—1596. DOI:10.1162/jocn.2010.21548
27. Kahneman D., Tversky A. Variants of uncertainty. *Cognition*, 1982. Vol. 11, no. 2, pp. 143—157. DOI:10.1016/0010-0277(82)90023-3
28. Killeen P.R. Pavlov + Skinner = Premack [Elektronnyi resurs]. *International Journal of Comparative Psychology*, 2014. Vol. 27, no. 4, pp. 544—568. URL: https://www.researchgate.net/profile/Peter_Killeen2/publication/269873794_Pavlov_Skinner_Premack/links/549861d30cf2c5a7e342bdca.pdf (Accessed 05.06.2020).
29. McDannald M.A. et al. Learning theory: a driving force in understanding orbitofrontal function. *Neurobiology of learning and memory*, 2014. Vol. 108, pp. 22—27. DOI:10.1016/j.nlm.2013.06.003
30. Zhang D. et al. Linking brain electrical signals elicited by current outcomes with future risk decision-making. *Frontiers in behavioral neuroscience*, 2014. Vol. 8, article ID 34, 15 p. DOI:10.3389/fnbeh.2014.00084

31. Linson A., Parr V., Friston K.J. Active inference, stressors, and psychological trauma: A neuroethological model of (mal) adaptive explore-exploit dynamics in ecological context. *Behavioural Brain Research*, 2020. Vol. 380, pp. 112—421. DOI:10.1016/j.bbr.2019.112421
32. Aston-Jones G. et al. Locus coeruleus neurons in monkey are selectively activated by attended cues in a vigilance task. *Journal of Neuroscience*, 1994. Vol. 14, no. 7, pp. 4467—4480. DOI:10.1523/JNEUROSCI.14-07-04467.1994
33. Mansouri F.A. et al. Managing competing goals — a key role for the frontopolar cortex. *Nature Reviews Neuroscience*, 2017. Vol. 18, no. 11, pp. 645—657. DOI:10.1038/nrn.2017.111
34. Mata R., Wilke A., Czienskowski U. Foraging across the life span: is there a reduction in exploration with aging? *Frontiers in neuroscience*, 2013. Vol. 7, article ID 34, 7 p. DOI:10.3389/fnins.2013.00053
35. McClure S.M., Gilzenrat M.S., Cohen J.D. An exploration-exploitation model based on norepinephrine and dopamine activity [Elektronnyi resurs]. In Weiss Y., Schölkopf B., Platt J.C. (eds.). *Advances in neural information processing systems: proceedings from the conference “Neural Information Processing Systems 2005”*, 2006, pp. 867—874. URL: <https://papers.nips.cc/paper/2950-an-exploration-exploitation-model-based-on-norepinephrine-and-dopamine-activity.pdf> (Accessed 05.06.2020).
36. Miller E.K., Cohen J.D. An integrative theory of prefrontal cortex function. *Annual review of neuroscience*, 2001. Vol. 24, pp. 167—202. DOI:10.1146/annurev.neuro.24.1.167
37. Miltner W.H.R., Braun C.H., Coles M.G.H. Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a “generic” neural system for error detection. *Journal of cognitive neuroscience*, 1997. Vol. 9, no. 6, pp. 788—798. DOI:10.1162/jocn.1997.9.6.788
38. Heil M. et al. N200 in the Eriksen-task: Inhibitory executive process? *Journal of Psychophysiology*, 2000. Vol. 14, no. 4, pp. 218—225. DOI:10.1027//0269-8803.14.4.218
39. Pearson J.M. et al. Neurons in posterior cingulate cortex signal exploratory decisions in a dynamic multioption choice task. *Current biology*, 2009. Vol. 19, no. 18, pp. 1532—1537. DOI:10.1016/j.cub.2009.07.048
40. Naudé J. et al. Nicotinic receptors in the ventral tegmental area promote uncertainty-seeking. *Nature neuroscience*, 2016. Vol. 19, no. 3, pp. 471—478. DOI:10.1038/nn.4223
41. Onge J.R.S., Abhari H., Floresco S.B. Dissociable contributions by prefrontal D1 and D2 receptors to risk-based decision making. *Journal of Neuroscience*, 2011. Vol. 31, no. 23, pp. 8625—8633. DOI:10.1523/JNEUROSCI.1020-11.2011
42. Stopper C.M. et al. Overriding phasic dopamine signals redirects action selection during risk/reward decision making. *Neuron*, 2014. Vol. 84, no. 1, pp. 177—189. DOI:10.1016/j.neuron.2014.08.033
43. Padoa-Schioppa C., Conen K.E. Orbitofrontal cortex: a neural circuit for economic decisions. *Neuron*, 2017. Vol. 96, no. 4, pp. 736—754. DOI:10.1016/j.neuron.2017.09.031
44. Parr V., Friston K.J. Uncertainty, epistemics and active inference. *Journal of The Royal Society Interface*, 2017. Vol. 14, no. 136, 10 p. DOI:10.1098/rsif.2017.0376
45. Lee M.D. et al. Psychological models of human and optimal performance in bandit problems. *Cognitive Systems Research*, 2011. Vol. 12, no. 2, pp. 164—174. DOI:10.1016/j.cogsys.2010.07.007
46. Bartholow B.D. et al. Psychophysiological evidence of response conflict and strategic control of responses in affective priming. *Journal of Experimental Social Psychology*, 2009. Vol. 45, no. 4, pp. 655—666. DOI:10.1016/j.jesp.2009.02.015
47. Rakow V., Newell B.R., Zougkou K. The role of working memory in information acquisition and decision making: Lessons from the binary prediction task. *The Quarterly Journal of Experimental Psychology*, 2010. Vol. 63, no. 7, pp. 1335—1360. DOI:10.1080/17470210903357945
48. Kiebel S.J. et al. Recognizing sequences of sequences. *PLoS computational biology*, 2009. Vol. 5, no. 8, 14 p. DOI:10.1371/journal.pcbi.1000464
49. Laviola G. et al. Risk-taking behavior in adolescent mice: psychobiological determinants and early epigenetic influence. *Neuroscience & Biobehavioral Reviews*, 2003. Vol. 27, no. 1—2, pp. 19—31. DOI:10.1016/S0149-7634(03)00006-X
50. Badre D. et al. Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration. *Neuron*, 2012. Vol. 73, no. 3, pp. 595—607. DOI:10.1016/j.neuron.2011.12.025
51. Sara S.J. The locus coeruleus and noradrenergic modulation of cognition. *Nature reviews neuroscience*, 2009. Vol. 10, no. 3, pp. 211—223. DOI:10.1038/nrn2573
52. Slovic P. Risk-taking in children: Age and sex differences. *Child Development*, 1966. Vol. 37, no. 1, pp. 169—176. DOI:10.2307/1126437
53. Smith A.P., Beckmann J.S., Zentall V.R. Gambling-like behavior in pigeons: ‘jackpot’ signals promote maladaptive risky choice. *Scientific reports*, 2017. Vol. 7, no. 1, pp. 1—11. DOI:10.1038/s41598-017-06641-x
54. Addicott M.A. et al. Smoking and the bandit: A preliminary study of smoker and nonsmoker differences in exploratory behavior measured with a multiarmed bandit task. *Experimental and clinical psychopharmacology*, 2013. Vol. 21, no. 1, pp. 66—73. DOI:10.1037/a0030843
55. Steyvers M., Lee M.D., Wagenmakers E.J. A Bayesian analysis of human decision-making on bandit problems. *Journal of Mathematical Psychology*, 2009. Vol. 53, no. 3, pp. 168—179. DOI:10.1016/j.jmp.2008.11.002
56. Warren C.M. et al. The effect of atomoxetine on random and directed exploration in humans. *PloS one*, 2017. Vol. 12, no. 4, 17 p. DOI:10.1371/journal.pone.0176034

57. Usher M. et al. The role of locus coeruleus in the regulation of cognitive performance. *Science*, 1999. Vol. 283, no. 5401, pp. 549—554. DOI:10.1126/science.283.5401.549
58. Jepma M. et al. The role of the noradrenergic system in the exploration-exploitation trade-off: a pharmacological study. *Frontiers in human neuroscience*, 2010. Vol. 4, article ID 170, 13 p. DOI:10.3389/fnhum.2010.00170
59. Laureiro-Martínez D. et al. Understanding the exploration—exploitation dilemma: An fMRI study of attention control and decision-making performance. *Strategic Management Journal*, 2015. Vol. 36, no. 3, pp. 319—338. DOI:10.1002/smj.2221
60. Mehlhorn K. et al. Unpacking the exploration—exploitation tradeoff: A synthesis of human and animal literatures. *Decision*, 2015. Vol. 2, no. 3, pp. 191—215. DOI:10.1037/dec0000033
61. Verdolin J.L. Meta-analysis of foraging and predation risk trade-offs in terrestrial systems. *Behavioral Ecology and Sociobiology*, 2006. Vol. 60, no. 4, pp. 457—464. DOI:10.1007/s00265-006-0172-6
62. Yuki S., Okanoya K. Rats show adaptive choice in a metacognitive task with high uncertainty. *Journal of Experimental Psychology: Animal Learning and Cognition*, 2017. Vol. 43, no. 1, pp. 109—118. DOI:10.1037/xan0000130
63. Zentall V.R. An animal model of human gambling based on pigeon suboptimal choice [Elektronnyi resurs]. *Research & Reviews: Neuroscience*, 2017. Vol. 1, no. 2, pp. 27—37. URL: <https://pdfs.semanticscholar.org/f4ba/8ebce42ca058e780c9afb1322b7440bc8649.pdf> (Accessed 05.06.2020).
64. Zentall V.R. Suboptimal choice by pigeons: An analog of human gambling behavior. *Behavioural processes*, 2014. Vol. 103, pp. 156—164. DOI:10.1016/j.beproc.2013.11.004

Информация об авторах

Сайфулина Ксения Эльдусовна, младший научный сотрудник, Центр нейрокогнитивных исследований (МЭГ-центр), Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ), г. Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0002-2017-0811>, e-mail: kseniasayfulina@gmail.com

Козунова Галина Леонидовна, кандидат психологических наук, старший научный сотрудник, Центр нейрокогнитивных исследований (МЭГ-центр), Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ), г. Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0002-1286-8654>, e-mail: chukhutova@gmail.com

Медведев Владимир Александрович, младший научный сотрудник, Центр нейрокогнитивных исследований (МЭГ-центр), Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ), г. Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0002-3252-8809>, e-mail: ixdon@yandex.ru

Рытикова Анна Менашевна, младший научный сотрудник, Центр нейрокогнитивных исследований (МЭГ-центр), Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ), г. Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0003-0153-9457>, e-mail: ann.zelener@mail.ru

Чернышев Борис Владимирович, кандидат биологических наук, руководитель, Центр нейрокогнитивных исследований (МЭГ-центр), Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ); доцент, кафедра высшей нервной деятельности, биологический факультет, Московский государственный университет имени М.В. Ломоносова (МГУ имени М.В. Ломоносова); доцент, департамент психологии, факультет социальных наук, Национальный исследовательский университет «Высшая школа экономики» (НИУ ВШЭ), г. Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0002-8267-3916>, e-mail: b_chernysh@mail.ru

Information about the authors

Ksenia E. Sayfulina, Junior Researcher, Center for Neurocognitive Research (MEG Center), Moscow State University of Psychology & Education, Moscow, Russia, ORCID: <https://orcid.org/0000-0002-2017-0811>, e-mail: kseniasayfulina@gmail.com

Galina L. Kozunova, PhD in Psychology, Senior Researcher, Center for Neurocognitive Research (MEG Center), Moscow State University of Psychology & Education, Moscow, Russia, ORCID: <https://orcid.org/0000-0002-1286-8654>, e-mail: chukhutova@gmail.com

Vladimir A. Medvedev, Junior Researcher, Center for Neurocognitive Research (MEG Center), Moscow State University of Psychology & Education, Moscow, Russia, ORCID: <https://orcid.org/0000-0002-3252-8809>, e-mail: ixdon@yandex.ru

Anna M. Rytikova, Junior Researcher, Center for Neurocognitive Research (MEG Center), Moscow State University of Psychology & Education, Moscow, Russia, ORCID: <https://orcid.org/0000-0003-0153-9457>, e-mail: ann.zelener@mail.ru

Boris V. Chernyshev, PhD in Biology, Associate Professor, Head of Center for Neurocognitive Research (MEG Center), Moscow State University of Psychology & Education; Associate Professor, Department of Higher Nervous Activity, Faculty of Biology, Lomonosov Moscow State University; Associate Professor, Department of Psychology, National Research University «Higher School of Economics», Moscow, Russia, ORCID: <https://orcid.org/0000-0002-8267-3916>, e-mail: b_chernysh@mail.ru