

**СРАВНИТЕЛЬНЫЙ АНАЛИЗ АКТИВНОСТИ НЕЙРОНОВ
ПРЕФРОНТАЛЬНОЙ КОРЫ МОЗГА КРЫС В СОСТОЯНИИ
ПРОДУКТИВНОЙ И НЕПРОДУКТИВНОЙ РАБОТЫ**

© 2021 г. Е. В. Филатова¹, *, А. А. Орлов¹, С. В. Афанасьев¹, Н. И. Горбунов¹

¹Институт эволюционной физиологии и биохимии им. И.М. Сеченова РАН,
Санкт-Петербург, Россия

*E-mail: filena17@gmail.com

Поступила в редакцию 06.08.2020 г.

После доработки 28.09.2020 г.

Принята к публикации 25.10.2020 г.

Задачей работы было исследование нейронной активности клеток префронтальной коры при разном нервно-психическом состоянии животного. Проводилась внеклеточная регистрация импульсной активности нейронов префронтальной коры мозга крыс при выполнении поведенческой задачи выбора в соответствии с условными сигналами в двухкольцевом лабиринте. В каждый опытный день в последовательности выполняемых реализаций выделяли периоды продуктивной и непродуктивной работы, различающиеся преобладанием правильного или неправильного выбора. Сравнительный анализ нейронной активности в шести различных вариантах выполнения программы (правильные побежки налево и направо в продуктивном и непродуктивном состоянии и ошибочные побежки), обнаружил существенное отличие профиля коллективной нейронной активности, сопутствующего подкрепляемым побежкам на одну сторону в непродуктивном состоянии. Это связывается с активацией сформированного в начале обучения профиля активности, сопутствующего подкрепляемым побежкам именно на эту сторону. Предполагается динамический механизм смены стратегий принятия решения в зависимости от состояния животного, проявляющийся в вытормаживании старых моделей при обучении в продуктивном состоянии и ослабление такого механизма контроля в непродуктивном.

Ключевые слова: префронтальная кора, поведение, крысы, нейронная активность, принятие решение, поведение с подкреплением

DOI: 10.31857/S0869813921010040

Реализация когнитивных функций находится в тесной связи с нервно-психическим состоянием животного. Широко исследуется роль префронтальной коры в когнитивных процессах, включая принятие решения, детекцию ошибок, контроль исполнения и память [1]. Префронтальная кора играет существенную роль в цепях фильтрации внимания и сенсорного выбора [2]. Эксперименты с временной инактивацией различных регионов префронтальной коры обнаруживают ее роль в тор-можении уже инициированного ответа, опосредуемую дофаминергическими и норадренергическими системами [3]. Исследования механизмов участия префронтальной коры в реализации инструментального поведения показало различную специфику в формировании ассоциаций действие—результат и результат—действие. Разрушение прелимбической префронтальной коры приводит к трудностям с из-

влечением и сохранением информации о результатах действий, необходимой для планирования последующей активности [4]. Эксперименты с повреждением медиальной орбитофронтальной коры показывают ее роль в облегчении целенаправленного тормозного ответа в обстоятельствах, которые требуют принятия новых стратегий реагирования [5].

Известна роль медиальной префронтальной коры (mPFC) в реализации рискованного поведения. Показано, что при повреждении mPFC у людей возникают дефекты в принятии решений, которые выражаются в нарушениях, связанных с планированием и игнорированием долгосрочных перспектив и последствием своих действий [6]. В экспериментах на животных показано, что модуляция активности дофаминовой системы или инактивация mPFC приводят к нарушениям в принятии решений о рисках и вознаграждениях [7]. В исследованиях на людях с использованием антагониста D2-рецепторов показано динамическое отслеживание в префронтальной коре значимости выбираемого ответа [8]. Для исследования механизмов принятия решения в разных состояниях на животных используют распространенную парадигму RDT (Risky Decision making Task), когда крысы выбирают между небольшим, “безопасным” вознаграждением и большим, но “рискованным”, которое сопровождается различной вероятностью умеренного наказания [9]. Показано, что инактивация mPFC увеличивает выбор большой рискованной награды (ascending RDT), когда вероятность наказания растет в течение сессии, но при этом уменьшается, когда вероятность наказания снижается в течение сеанса (descending RDT) [10].

Однако эксперименты, использующие методы повреждения или временной инактивации мозговых структур, имеют свои ограничения. Целью данной работы было исследование функционирования префронтальной коры в условиях естественной модуляции поведения. Различная степень успешности выполнения задачи влияет на нервно-психическое состояние животного и приводит к смене стратегий выбора, а одновременная регистрация импульсной активности позволяет проводить сравнительный анализ активности одних и тех же нейронов в условиях одной сессии, но в разном состоянии животного.

МЕТОДЫ ИССЛЕДОВАНИЯ

Животные. Исследование нейронной активности проводили на половозрелых самцах крыс линии Вистар ($n = 3$) после имплантации микроэлектродной системы. Все манипуляции осуществлялись в соответствии с протоколом обращения с лабораторными животными, утвержденным биоэтическим комитетом Института эволюционной физиологии и биохимии РАН и основанном на директиве Европейского Сообщества по гуманному обращению с экспериментальными животными (European Communities Council Directive #86/609 for the Care of Laboratory Animals).

Поведенческая модель. Тестирование поведения осуществлялось в двухкольцевом лабиринте [11]. Лабиринт состоит из двух замкнутых на стартовую камеру рукавов с отсеками, разделенными дверками, открывающимися только вперед (рис. 1). Данная поведенческая модель позволяет животному автономно перемещаться, осуществляя в каждой реализации выбор траектории перемещения. Выбор направления происходит в сигнальном отсеке, где расположены светодиоды. Толкание дверки, ведущей в один из рукавов в соответствии с горящим условным сигналом, сопровождается щелчком кормушки. Побежка по любому выбранному рукаву приводит в общий пищевой отсек, где животное, в зависимости от правильности выбора, получает или не получает подкрепление. Выход из пищевого отсека имеется только в стартовую камеру, откуда начнется следующая реализация.

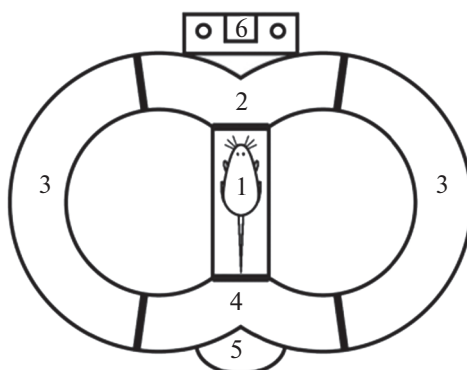


Рис. 1. Схема поведенческого двухкольцевого лабиринта. Отсеки: 1 – стартовый отсек, 2 – сигнальный отсек, 3 – правый и левый рукава лабиринта, 4 – пищевой отсек, 5 – кормушка, 6 – условные сигналы.
Fig. 1. Scheme of the two-ring maze. Compartments: 1 – start compartment, 2 – signal compartment, 3 – right and left arms, 4 – food compartment, 5 – feeder, 6 – signals.

Обучение. Обучение животных поведению в лабиринте осуществлялось поэтапно. На первом этапе животных обучали получать подкрепление в пищевом отсеке. В качестве подкрепления использовались пшеничные шарики. Появление каждого шарика сопровождалось щелчком кормушки. Далее животных обучали толкать дверку выхода из пищевого отсека в стартовый, потом, после предупреждающего сигнала и снятия стопора, животное обучалось проходить в правый рукав и далее в пищевой отсек, получая там подкрепление. Левая часть лабиринта была загорожена заслонкой, так что при выходе из стартового отсека на этом этапе животному была доступна только правая дверка. После успешного обучения прохождению всего цикла по правому рукаву проводилась операция по имплантации микроэлектродной системы. Такое предварительное обучение, проводимое до имплантации системы, было необходимо для того, чтобы отобрать животных для дальнейшего эксперимента. Отдельные особи были склонны часто разворачиваться в отсеках лабиринта или проявлять страх и агрессию по отношению к экспериментатору. Регистрация активности клеток требует многочисленных манипуляций с животным, поэтому в период предварительного обучения производился отбор животных. Дальнейшее обучение с допуском к левому рукаву лабиринта и обучением работе на условные сигналы осуществлялось уже одновременно с регистрацией нейронной активности. Сигналы представляли собой светодиоды, расположенные в сигнальном отсеке. Включение лампочки производилось после открытия крысой дверки в сигнальный отсек. Выбор стороны побежки, совпадающий с горячей лампой, сопровождался щелчком кормушки и появлением подкрепления в кормушке. Щелчок раздавался в момент закрытия за крысой дверки в выбранный рукав лабиринта: правый (R) или левый (L). Такая побежка считалась правильной. В дальнейшем мы обозначаем их знаком плюс (+). При несовпадении стороны сигнала и выбранного рукава лабиринта щелчка не раздавалось и в кормушке не появлялось подкрепления. Такая побежка считалась ошибочной. В дальнейшем мы обозначаем их знаком минус (–).

Имплантация системы и регистрация нейронной активности. Микроэлектродная система из 14 электродов представляет из себя два блока, каждый по 7 платиново-иридиевых заточенных микроэлектродов в кварцевой изоляции ($d = 50$ мкм, оголенный кончик жилы $d = 5-7$ мкм, сопротивление около 1 МОм), где каждый электрод имеет самостоятельный микродрайвер, позволяющий передвигать его не-

зависимо от остальных. Каждый блок из 7 микроэлектродов представляет собой конус, в котором направляющие каналы расположены под углом к вертикальной оси системы. Направляющие расположены по окружности [12]. Полный оборот микродрайвера погружает микроэлектрод на 250 мкм, небольшой угол поворота позволяет подводить кончик каждого электрода близко к мембране нейрона. Имплантация системы проводилась во время хирургической операции под наркозом (золетил 60 мг/кг, внутримышечно), это суммарная доза, которую вводили дробно по потребности по ходу операции. Крепление к черепу системы осуществлялось стоматологической пластмассой. Также пластмассой крепились два индифферентных серебряных электрода, расположенных в отверстиях черепа в симметричных точках теменной кости двух полушарий. Внеклеточная регистрация нейронной активности производилась одновременно в двух симметричных зонах префронтальной коры мозга крысы: координаты по атласу (AP:–1, L:1.2–1.5), угол оси погружения – 15° относительно сагиттальной плоскости [13]. Диаметр кончика направляющей, располагающейся непосредственно над поверхностью коры – 0.5 мм, расстояние между отдельными электродами зависит от разницы в глубине погружения, но в большинстве случаев не превышает 200 мкм. Непосредственно на черепе также располагался предусилитель, выполненный на 8 двухканальных микросхемах (коэффициент усиления 10), соединенный кабелем с усилителем (коэффициент усиления 100, полоса пропускания от 300 до 10000 Гц), сигналы от которого поступали на аналого-цифровой преобразователь компьютера. Синхронизация регистрируемой импульсной активности и временных интервалов выполняемого поведения осуществлялась через контроллер Nasional Instruments, управление установкой автоматизировано с использованием среды LabView.

Через неделю после операции и полного заживления приступали к поведенческим экспериментам и регистрации активности нейронов. Регистрация осуществлялась через 40–60 мин после поиска клеток, при устойчивой активности каждой клетки. В каждый опытный день крысы выполняли 50–80 реализаций с предъявлением условных сигналов. Сигналы предъявлялись в псевдослучайной последовательности, так, чтобы альтернатив и персевераций было сравнимое количество и их появление было равновероятным. Таким образом, регистрировались правильные (подкрепляемые) и ошибочные (неподкрепляемые) побежки налево и направо. В данной работе для анализа взяты две последовательных опытных серии для каждой крысы, в которых правильных реализаций было около 50%. У первой крысы это были 12-я и 13-я серии от начала эксперимента, у второй – 13-я и 14-я, у третьей 12-я и 13-я.

Анализ нейронной активности. Из последовательности выполняемых реализаций выделялись периоды, когда крыса совершала больше ошибочных или больше правильных выборов, и относили эти периоды к продуктивному или непродуктивному состоянию. Реализации отбирались по критерию количества ошибок и правильных выполнений в скользящем окне из последовательности в 6 реализаций. Если правильная реализация попадала в окно, где из 6 последовательных реализаций было не больше 1 ошибки, то есть, когда ошибок в текущем интервале было не более 16%, то она относилась к продуктивному состоянию, остальные – к непродуктивному. Здесь приведен пример последовательности выполняемых испытаний, где индекс обозначает последовательный номер испытания: R₁⁺, L₂⁺, L₃⁺, R₄[–], L₅⁺, L₆⁺, R₇⁺, R₈[–], R₉[–], R₁₀⁺, L₁₁⁺, L₁₂[–], R₁₃⁺, L₁₄[–], R₁₅⁺. В приведенном примере 1-е, 2-е, 3-е, 5-е, 6-е и 7-е правильные испытания относятся к продуктивному периоду, а 10-е, 11-е, 13-е и 15-е – к непродуктивному. Разделяя таким образом успешные правые и левые реализации, производилось сравнение успешных выполнений непродуктивных периодов (R₁⁺) и (L₁⁺), мы их маркируем цифрой 1, с успешными выполнениями продуктивных периодов (R₂⁺) и (L₂⁺), мы их марки-

Таблица 1. Число реализаций программы для каждого варианта поведения
Table 1. The number of the trials for each variant of the behavior

Поведение Behavior	Крыса № 1 Rat № 1	Крыса № 2 Rat № 2	Крыса № 3 Rat № 3
R2+	40	13	28
R1+	8	10	18
L2+	39	11	21
L1+	6	14	10
R–	25	41	23
L–	18	11	31
Всего All	136	100	131

руем цифрой 2, а также с ошибочными выполнениями (R–) и (L–). Ошибочные выполнения не разделяли на попавшие в продуктивный и непродуктивный интервал, предполагая, что ошибка сама по себе является свидетельством непродуктивности. В табл. 1 представлено число попавших в анализ реализаций каждого поведенческого варианта для каждой крысы.

Все электроды после начала регистрации оставались неподвижными, и для верификации нейронов использовался сравнительный анализ формы спайков, зарегистрированных в начале и в конце эксперимента. Таким образом, активность каждого из нейронов, вошедших в анализ, была зарегистрирована во всех вариантах выполнения поведенческой программы. Анализ формы спайков, их фильтрация от электрических наводок, накопление и синхронизация с поведенческими метками производились в специально написанной на заказ программе shevelev.pro. Регистрируемые временные метки прохождения отсеков лабиринта позволяют выделить 6 периодов поведенческой программы: (1) фоновая активность, регистрируемая в стартовом отсеке, от входа в отсек до появления звукового предупреждающего сигнала, (2) стартовая активность – от предупреждающего сигнала до открытия дверок в сигнальный отсек, (3) от открытия дверок в сигнальный отсек до открытия дверок в выбранный рукав лабиринта, (4) время прохождения дверки входа в рукав лабиринта – от ее открытия до закрытия, (5) время прохождения рукава – от закрытия дверки входа до открытия дверки выхода из рукава, (6) период пребывания в пищевом отсеке. Оценивалась средняя частота импульсации на каждом этапе выполнения поведенческой программы. Таким образом, анализ активности каждого нейрона включал в себя 36 измерений: усредненную частоту, зарегистрированную на 6 последовательных этапах в 6 поведенческих вариантах – подкрепляемые побежки направо и налево в непродуктивный период (R1+), (L1+), подкрепляемые побежки направо и налево в продуктивный период (R2+), (L2+) и ошибочный выбор (R–) и (L–). Полученные данные обрабатывались стандартными методами дискриминантного анализа (Statistica 6.0), который позволяет оценить степень сходства/различия групп (в данном случае группой являлся один из 6 поведенческих вариантов) путем вычисления расстояний Махаланобиса между их центроидами. Матрица для анализа состояла из 36 строк, в каждой из которых была представлена средняя частота импульсной активности на 6 поведенческих этапах в 6 вариантах выполнения поведенческой программы. Поведенческие варианты в использовании дискриминантного анализа мы рассматриваем как отдельные группы, каждый нейрон – как отдельную переменную. Достоверность межгрупповых различий определялась F-статистикой и ее значимостью. Степень сходства оценивалась по расстоянию Махаланобиса между центроидами сравниваемых облаков соответствующих активности групп нейронов в определенной поведенче-

ской ситуации. Для визуализации сходства и отличий профилей нейронной активности, сопутствующих разным поведенческим ситуациям, использовали кластерный анализ (древовидная кластеризация по методу ближайшего соседа) полученных квадратов расстояний Махаланобиса.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Анализ времени прохождения различных отсеков лабиринта в разных поведенческих ситуациях показал высокую вариативность у трех крыс. Так, время принятия решения (этап 3) у крысы № 1 колеблется в интервале от 800 до 3000 мс, при этом минимальное время наблюдается при верном выборе левой стороны в продуктивный период (L2+), а максимальное – при выборе той же левой стороны в непродуктивный (L1+). У крысы № 2 – от 1300 до 20000 мс, при этом минимальное время наблюдается при выборе правой стороны в непродуктивном состоянии (R1+), а максимальное – при ошибочном выборе правой стороны (R–). У крысы № 3 – от 1300 до 18000 мс, при этом минимальное время наблюдается при верном выборе левой стороны в непродуктивный период (L1+), а максимальное – при ошибочном выборе правой стороны (R–). Менее вариативно время побежки по рукаву (этап 5), так, для крыс № 2 и № 3 это время колеблется в пределах 1300–4500 мс без существенных отличий между вариантами поведения. У крысы № 1, наоборот, время побежки по рукаву существенно короче – от 350 до 1300 мс при правильных выполнениях (R1+, R2+ и L2+) и значительно длиннее при ошибочных (R– и L–) и правильном левом выборе в непродуктивном состоянии (L1+).

Анализ нейронной активности проводился для каждой крысы отдельно и включал в себя записи активности 58 клеток префронтальной коры ($n = 20$, $n = 19$, $n = 19$). На рис. 2 представлены паттерны изменения частоты импульсной активности по этапам программы при разных вариантах ее выполнения. У разных животных эти паттерны отличны. Так, у первой крысы частота импульсной активности выше практически на всех этапах программы при ошибочном выборе правой стороны (R–) и правильном выборе левой в непродуктивном состоянии (L1+) по сравнению с другими вариантами выполнения программы. У крысы 2 наблюдается существенное превышение частоты импульсной активности при ошибочном выборе как левого, так и правого рукава (R–, L–), в отличие от крысы 3, где минимальная частота выявлялась при ошибочной побежке направо (R–), а максимальная – при ошибочной побежке налево (L–).

На рис. 3 представлен результат кластерного анализа квадратов расстояний Махаланобиса, отражающий степень сходства профилей коллективной активности нейронов, проявляемой в разных вариантах выполнения поведенческой программы.

Древовидные диаграммы демонстрируют кластеры, объединяющие отдельные группы. На диаграммах видно выделение в отдельный кластер активности, соответствующей успешным побежкам направо в непродуктивном состоянии (R1+). В табл. 2 представлен результат F-статистики и вероятность попарного сравнения, которая позволяет оценить значимость отличий активности в разных вариантах поведения. Во всех случаях активность, сопутствующая (R1+) статистически значимо отличается от всех прочих ($p < 0.05$). При этом различия между правыми и левыми успешными реализациями в продуктивном состоянии во всех случаях (R2+ и L2+) нет.

Анализ коэффициентов дискриминантных функций позволяет оценить вклад отдельных нейронов в дискриминацию. Для каждой группы были отобраны клетки с максимальными коэффициентами, имеющие статистически значимые ($p < 0.05$) значения критерия Фишера. Из 58 нейронов таких оказалось 12. Анализ не обнару-

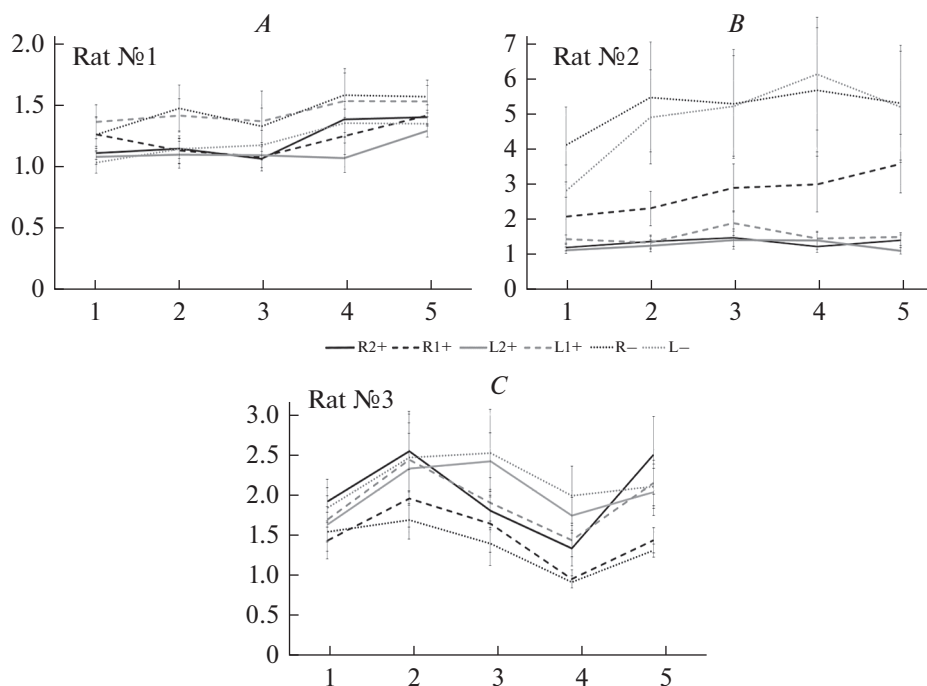


Рис. 2. Средняя частота импульсной активности по отношению к фоновой на разных этапах поведенческой программы при разных вариантах ее выполнения (R1+, R2+, L1+, L2+, R-, L-). Все правые побежки – черные, все левые – серые, сплошная линия – продуктивное состояние, длинный штрих – непродуктивное, точки – ошибочные выполнения. По оси абсцисс этапы программы: 1 – старт, 2 – принятие решения, 3 – прохождение выбранной двери в рукав, 4 – прохождение рукава, 5 – подкрепление. По оси ординат – частота импульсной активности $\text{имп/с} \pm$ ошибка среднего. *A*, *B* и *C* – разные животные.

Fig. 2. The average frequency of the impulse activity (to the background) at the different stages of the behavioral program for different behavioral states (R1+, R2+, L1+, L2+, R-, L-). All (R) trials are black, all (L) trials are gray, a solid line is an efficient state, a long stroke is ineffective state, dots are error trials. On the abscissa axis the stages of the program: 1 – start, 2 – decision making, 3 – passage across the choosed door, 4 – passage across the arm, 5 – reinforcement. The ordinate is the frequency of impulse activity, $\text{имп/с} \pm$ SEM. *A*, *B* and *C* are different animals

жил никаких специфических признаков для этой выборки. Средняя фоновая частота этих клеток варьирует в диапазоне от 0.2 до 18 имп/с , а также эти нейроны представлены в обоих полушариях. Изменение средней частоты по этапам программы этих нейронов для каждой группы в отдельности повторяет паттерн такого изменения, наблюдаемого для всей группы, представленного на рис. 2. Это может свидетельствовать о неспецифическом характере вовлекаемых в формирование профиля ответа клеток и соответствует данным актограмм, показывающих индивидуальный характер поведенческой модели.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Обнаруженная разница для разных животных в соотношении длительностей прохождения отсеков лабиринта при разных вариантах выполнения программы, а также разное распределение средней частоты импульсной активности, свидетельствует о специфически персональном характере сформированного поведенческого паттерна и отражает индивидуальные особенности исследованных животных. Многомерная статистика дискриминантного анализа учитывает разницу в актив-

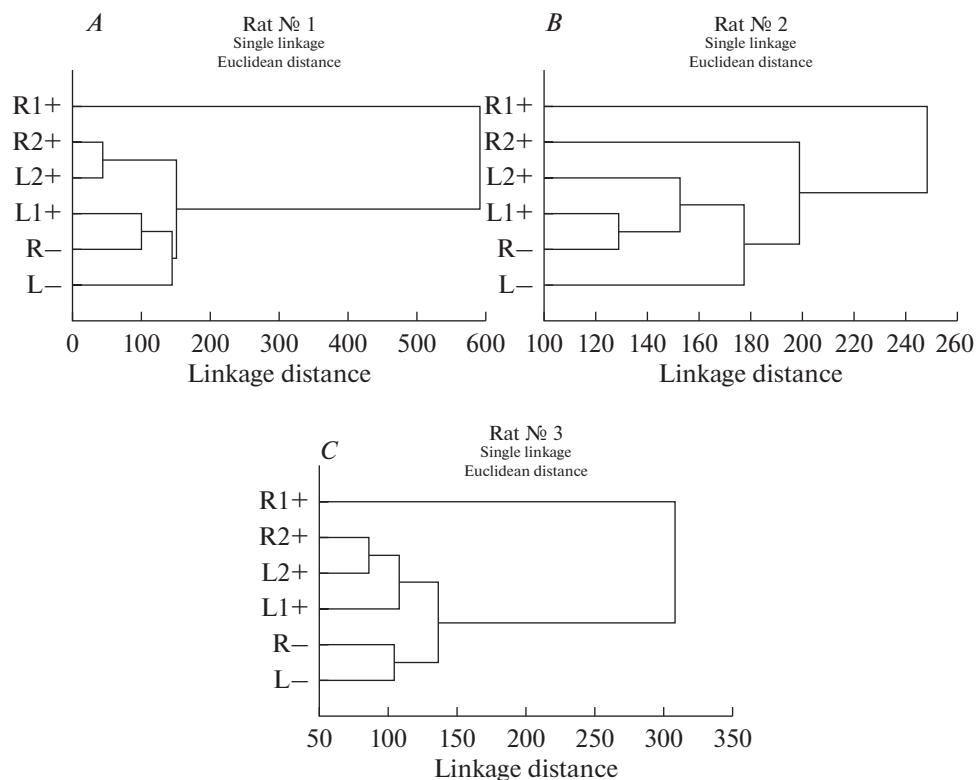


Рис. 3. Кластерный анализ профилей импульсной активности, сопутствующих разным поведенческим состояниям: успешные побежки в непродуктивном состоянии (R1+) и (L1+), успешные побежки в продуктивном состоянии (R2+) и (L2+) и ошибочные реализации (R-) и (L-). *A, B и C* – разные животные.

Fig. 3. Cluster analysis of the impulse activity associated with different behavioral states: successful motions in an ineffective state (R1+) and (L1+), successful motions in a efficient state (R2+) and (L2+) and error trials (R-) and (L-). *A, B and C* are different animals.

ности каждого нейрона при выполнении разных вариантов поведенческой программы. Результатом является метрика сходства или различия коллективной активности нейронов, сопутствующей разному поведению. Дискриминантный анализ коллективной активности нейронов, несмотря на индивидуальный профиль изменения средней частоты импульсации по этапам программы, обнаружил сходный признак у разных животных. Сравнение импульсной активности нейронов префронтальной коры в продуктивные и непродуктивные периоды работы выявило у всех исследованных животных общую закономерность – отличие профиля активности (R1+) от всех прочих вариантов выполнения поведенческой программы. Обучение всех животных проводилось с правой стороны. До имплантации микроэлектродной системы крыс обучали выполнению побегов по правому кольцу лабиринта с подкреплением каждой побежки. Дальнейшее обучение включало перенос навыка на другую сторону, и после успешного освоения левого кольца производилось обучение работе на условные сигналы. Таким образом, индивидуальный опыт животных включает побежки направо в двух разных режимах подкрепления. В первом случае, при формировании оперантной части, подкреплялись все побежки, а во втором, при предъявлении условных сигналов, подкреплялись только соответствующие сигналам. При обучении задаче билатерального выбора с выполне-

Таблица 2. Оценка достоверности (*p*-level) групповых различий по данным дискриминантного анализа. А, В и С – разные животные
Table 2. *p*-level of group differences (distances between groups) based on Discriminant function analysis. A, B and C are different animals

	p-level A						p-level B						p-level C					
	R1+	R2+	L1+	L2+	R-	L-	R1+	R2+	L1+	L2+	R-	L-	R1+	R2+	L1+	L2+	R-	L-
R1+	0.000	0.000	0.003	0.000	0.005	0.001	R1+	0.001	0.004	0.006	0.022	0.013	R1+	0.005	0.014	0.002	0.017	0.002
R2+	0.000	0.034	0.092	0.092	0.092	0.250	R2+	0.001	0.015	0.476	0.001	0.002	R2+	0.005	0.152	0.313	0.465	0.759
L1+	0.003	0.034	0.028	0.028	0.596	0.477	L1+	0.004	0.015	0.162	0.267	0.193	L1+	0.014	0.152	0.056	0.287	0.096
L2+	0.000	0.957	0.028	0.028	0.119	0.406	L2+	0.006	0.162	0.022	0.022	0.015	L2+	0.313	0.056	0.014	0.014	0.294
R-	0.005	0.092	0.596	0.119	0.736	0.736	R-	0.006	0.006	0.022	0.022	0.110	R-	0.465	0.287	0.014	0.082	0.082
L-	0.001	0.477	0.406	0.406	0.736	0.736	L-	0.013	0.193	0.015	0.110	L-	0.759	0.096	0.294	0.082	0.082	0.082

нием последовательности действий невозможно избежать этого, так как обучение моторной части задачи всегда заставляет экспериментатора выбрать первой для обучения одну из сторон. Такое обучение занимает не более двух экспериментальных серий, состоящих из не более 30–40 реализаций. Имплантация системы и обучение побежкам на другую сторону проводилось сразу после первичного обучения, но, по-видимому, этого достаточно для формирования навыка. Таким образом, при первичном обучении подкрепляемых побежек направо формируется устойчивый профиль, который, по-видимому, мы и наблюдаем при последующем анализе. Такие подкрепляемые побежки направо можно рассматривать как классическое оперантное обуславливание. В основе такого поведения лежит устойчивая связь между выполняемым действием и побегом, формируемая на основе мотивации голода. В терминах Конорского это можно рассматривать как подготовительную деятельность или драйв-рефлекс, в отличие от исполнительской [14]. Развивая концепцию двух разных мотивационных систем, Dickinson и Balleine разделяют Павловскую мотивацию, приводящую к стимул-реакция ответам и мотивацию, приводящую к выбору определенного действия, направленного на достижение определенной цели. Эта концепция полагает, что целенаправленные действия включают в себя некий когнитивный процесс, в котором целенаправленные действия вызываются взаимодействием убеждения, что реакция приводит к результату, с желанием этого конкретного результата [15]. В теории reinforcement learning предполагается, что эти два разных механизма имеют разную нейрофизиологическую природу. Классические представления о дофамине как предикторе ошибок, лежащем в основе ассоциации между ситуациями, действиями и наградами за счет модуляции силы синаптических связей между нейронами, хорошо описывают связи стимул-реакции [16, 17]. Считается, что дофаминовые нейроны отвечают за процессы подкрепления в обучении с вознаграждением без использования моделей (the model-free reward learning). Напротив, обучение с вознаграждением на основе модели (the model-based reward learning) задействует участие различных областей префронтальной коры, отвечающих за знание специфических для задачи правил [18]. С этим связан обнаруживаемый гибкий ответ префронтальной коры в условиях модуляции задач [19].

В нашей поведенческой модели можно предположить включение model-based reward learning в продуктивные периоды, а model-free reward learning в непродуктивные. Появление первичного R1+ профиля проявляется именно в непродуктивные периоды работы, когда животное совершает больше ошибочных, чем правильных реализаций. Видимо, в этот период происходит смена стратегии выбора и животное выбирает не собственное действие, а пытается предсказать сторону появления подкрепления, то есть, выбирает стратегию, при которой надеется получить небольшое подкрепление, не ориентируясь на сигналы, а довольствуется тем, которое получает при случайном выборе. По-видимому, в непродуктивные периоды работы животное осуществляет выбор случайным образом, и при успешном выборе правого рукава происходит активация профиля первичного обучения. Можно предположить, что такая активация будет тормозить обучение работе на сигналы, так как подкрепляется именно непродуктивная работа.

Важным тут является то, что положительное подкрепление подкрепляет именно выбранную стратегию принятия решения, но поскольку в большинстве случаев мы не знаем на основании чего совершен выбор, то для успешности обучения надо минимизировать такие подкрепления.

В экспериментах исследования выбора рискованного и нерискованного поведения показано, что инактивация префронтальной коры увеличивает выбор, связанный с повышенным риском [10]. Это может быть связано с тем, что животное переходит в режим вероятностного анализа получения подкрепления, игнорируя другие факторы, что согласуется с нарушениями в оценке результата выполняемых

действий [4]. В нашей модели условия эксперимента так же, как и вероятность получения подкрепления, остаются неизменными, однако животное в зависимости от выбранной стратегии сталкивается с разной частотой подкрепления и соответственно с модуляцией эмоциональной сферы.

Во всех трех случаях профили активности, соответствующие правильному выбору направо и налево (R2+ и L2+) и ошибочным побежкам (L– и R–), между собой перекрываются, не показывая существенной разницы. Проведенный анализ включает в себя все этапы поведенческой программы, как латерализованные так и нет, в анализ при этом включались все этапы поведенческой программы. Отсутствие дифференцировки между правыми и левыми реализациями не свидетельствует о ее истинном отсутствии, а говорит лишь о том, что разница между состояниями (продуктивным и непродуктивным) является более существенной для результатов дискриминантного анализа.

Очевидно, что при обучении мы имеем дело не только с формированием новых профилей, но и с отказом от непродуктивных старых. Обнаруженный в префронтальной коре устойчивый профиль нейронной активности, по-видимому, связанный с предшествующим опытом первичного обучения подкрепляемых побежек направо, свидетельствует о роли префронтальной коры в обучении, проявляющейся в конфликте старых моделей поведения с необходимостью формирования новых. Можно предположить динамический механизм смены стратегий принятия решения в зависимости от состояния животного, проявляемый в вытормаживании старых моделей при обучении в продуктивном состоянии и ослабление такого механизма контроля в непродуктивном.

ИСТОЧНИКИ ФИНАНСИРОВАНИЯ

Работа выполнена в рамках государственного задания (AAAA-A18-118012290373-7).

СПИСОК ЛИТЕРАТУРЫ

1. Euston D.R., Gruber A.J., McNaughton B.L. The role of medial prefrontal cortex in memory and decision making. *Neuron*. 76(6): 1057–1070. 2012.
<https://doi.org/10.1016/j.neuron.2012.12.002>
2. Nakajima M., Schmitt L.I., Halassa M.M. Prefrontal Cortex Regulates Sensory Filtering through a Basal Ganglia-to-Thalamus Pathway. *Neuron*. 103(3): 445–458. 2019.
<https://doi.org/10.1016/j.neuron.2019.05.026>
3. Bari A., Mar A.C., Theobald D.E., Elands S.A., Oganya K.C., Eagle D.M., Robbins T.W. Prefrontal and monoaminergic contributions to stop-signal task performance in rats. *J. Neurosci*. 31(25): 9254–9263. 2011.
<https://doi.org/10.1523/JNEUROSCI.1543-11.2011>
4. Corbit L.H., Balleine B.W. The role of prelimbic cortex in instrumental conditioning. *Behav. Brain Res*. 146: 145–157. 2003.
<https://doi.org/10.1016/j.bbr.2003.09.023>
5. Gourley S.L., Lee A.S., Howell J.L., Pittenger C., Taylor J.R. Dissociable regulation of instrumental action within mouse prefrontal cortex. *Eur. J. Neurosci*. 32(10): 1726–1734. 2010.
<https://doi.org/10.1111/j.1460-9568.2010.07438.x>
6. Bechara A., Damasio A.R., Damasio H., Anderson S.W. Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*. 50: 7–15. 1994.
[https://doi.org/10.1016/0010-0277\(94\)90018-3](https://doi.org/10.1016/0010-0277(94)90018-3)
7. St Onge J.R., Abhari H., Floresco S.B. Dissociable contributions by prefrontal D1 and D2 receptors to risk-based decision making. *J. Neurosci*. 31: 8625–8633. 2011.
<https://doi.org/10.1523/JNEUROSCI.1020-11.2011>
8. Jocham G., Klein T.A., Ullsperger M. Dopamine-mediated reinforcement learning signals in the striatum and ventromedial prefrontal cortex underlie value-based choices. *J. Neurosci*. 31(5): 1606–1613. 2011.
<https://doi.org/10.1523/JNEUROSCI.3904-10.2011>
9. Yates J.R. Examining the neurochemical underpinnings of animal models of risky choice: Methodological and analytic considerations. *Exp. Clin. Psychopharmacol*. 27(2): 178–201. 2019.
<https://doi.org/10.1037/pha0000239>
10. Orsini C.A., Heshmati S.C., Garman T.S., Wall S.C., Bizon J.L., Setlow B. Contributions of medial prefrontal cortex to decision making involving risk of punishment. *Neuropharmacology*.

- 139: 205–216. 2018.
<https://doi.org/10.1016/j.neuropharm.2018.07.018>
11. *Filatova E.V., Orlov A.A., Afanas'ev S.V.* A two-ring maze for studies of the behavior of laboratory animals. *Neurosci. Behav. Physiol.* 45(7): 765–770. 2015.
 12. *Orlov A.A., Filatova E.V.* A miniature system of independent micromanipulators for recording neuron activity in freely moving animals. *Neurosci. Behav. Physiol.* 42(1): 97–99. 2012.
 13. *Paxinos G., Watson Ch.* The Rat Brain in Stereotaxic Coordinates, 6th Edition Hard Cover Edition. Acad. Press. 2006.
 14. *Konorski J.* Conditioned Reflexes and Neuron Organization. Cambridge. England. Cambridge Univer. Press. 1948.
 15. *Dickinson A., Balleine B.* Motivational Control of Instrumental Action. *Current Directions Psychol. Sci.* 4(5): 162–167. 1995.
 16. *Schultz W., Dayan P., Montague P.R.* A neural substrate of prediction and reward. *Science.* 275: 1593–1599. 1997.
 17. *Schultz W.* Getting formal with dopamine and reward. *Neuron.* 36: 241–263. 2002.
 18. *Daw N.D., Niv Y., Dayan P.* Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8(12): 1704–1711. 2005.
 19. *Kaefer K., Nardin M., Blahna K., Csicsvari J.* Replay of Behavioral Sequences in the Medial Prefrontal Cortex during Rule Switching. *Neuron.* 106(1): 154–165. 2020.

Comparative Analysis of the Prefrontal Cortex Rat Brain Neuron Activity in the Efficient and Ineffective Behavior

E. V. Filatova^{a, *}, A. A. Orlov^a, S. V. Afanasyev^a, and N. I. Gorbunov^a

^a*Sechenov Institute of Evolutionary Physiology and Biochemistry Russian Academy of Sciences,
St. Petersburg, Russia*

^{*}*e-mail: filena17@gmail.com*

The goal of the work was to study the prefrontal cortex neural activity in the different neuropsychological conditions of the animal. Extracellular neuron activity of the prefrontal cortex rat brain was recorded in rats performing the behavioral task of choice in the two-ring maze. On each experimental day, in the sequence of trials, periods of efficient and ineffective work were distinguished, differing in the prevalence of correct or incorrect choice. A comparative analysis of the neuron activity in six different variants of behavior revealed a significant difference in one pattern of neural activity accompanying reinforced trials to the one side in the ineffective condition. It is assumed that this is due to the activation of a previously formed pattern of behavior such leads to a probabilistic choice strategy. A dynamic mechanism of changing some decision-making strategies depending on the neuropsychological conditions of the animal is assumed, manifested in the inhibition of old models during training in a efficient state and a weakening of such a control mechanism in an ineffective state.

Keywords: prefrontal cortex, efficient behavior, rat, unit activity, reinforcement learning, decision-making

ЦИТИРОВАТЬ:

Филатова Е.В., Орлов А.А., Афанасьев С.В., Горбунов Н.И. Сравнительный анализ активности нейронов префронтальной коры мозга крыс в состоянии продуктивной и непродуктивной работы. *Рос. физиол. журн. им. И.М. Сеченова.* 107(1): 43–54.

DOI: 10.31857/S0869813921010040

TO CITE THIS ARTICLE:

Filatova E.V., Orlov A.A., Afanasyev S.V., Gorbunov N.I. Comparative analysis of the prefrontal cortex rat brain neuron activity in the efficient and ineffective behavior. *Russian Journal of Physiology.* 107(1): 43–54.

DOI: 10.31857/S0869813921010040