

ФИЗИОЛОГИЯ ВЫСШЕЙ НЕРВНОЙ (КОГНИТИВНОЙ)  
ДЕЯТЕЛЬНОСТИ ЧЕЛОВЕКА

УДК 159.91

ИСПОЛЬЗОВАНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ  
ДЛЯ АНАЛИЗА ПАТТЕРНОВ АКТИВНОСТИ МОЗГА В ПРОЦЕССЕ  
ОЦЕНКИ СЕБЯ И ДРУГИХ ЛЮДЕЙ

© 2023 г. Г. Г. Князев<sup>1, \*</sup>, А. Н. Савостьянов<sup>1, 2, 3</sup>, П. Д. Рудыч<sup>1</sup>, А. В. Бочаров<sup>1, 2</sup>

<sup>1</sup>Федеральное государственное бюджетное научное учреждение

“Научно-исследовательский институт нейронаук и медицины”, Новосибирск, Россия

<sup>2</sup>Новосибирский государственный университет, Новосибирск, Россия

<sup>3</sup>Федеральный исследовательский центр Институт цитологии и генетики СО РАН,  
Новосибирск, Россия

\*e-mail: knyazev@physiol.ru

Поступила в редакцию 05.06.2022 г.

После доработки 28.06.2022 г.

Принята к публикации 31.10.2022 г.

Изучение активности мозга в процессе обработки самореферентной информации, в сравнении с обработкой информации, относящейся к другим людям, базируется на применении массового одномерного анализа, основанного на предположении, что активность в одном регионе не зависит от активности в других регионах. В последнее время в нейровизуализации растет интерес к исследованию пространственно распределенной информации с помощью многомерных подходов, таких как многовоксельный анализ паттернов (МВАП). В данной работе мы использовали МВАП для анализа фМРТ-данных, записанных при выполнении задания по оцениванию себя и других людей разной степени близости. Тестирование выявленных в процессе машинного обучения паттернов показало, что они позволяют в 75–88% случаев предсказать по активности мозга, оценивает ли испытуемый себя или другого человека. Прогностически значимые структуры были широко распределены по разным областям мозга и, помимо корковых срединных структур, дававших наибольший вклад, включали области зрительной, латеральной префронтальной и многих других корковых зон. При классификации “Я”–“Другие” наиболее информативными для выбора варианта “Я” являлись вентральные области медиальной префронтальной и поясной коры, а для выбора варианта “Другие” – теменные и затылочные срединные области. Анализ главных компонент позволил выявить комбинацию структур мозга, включающую переднюю часть поясной извилины и обе миндалины, факторные оценки которой коррелировали положительно с психометрической шкалой чувствительности к награде и отрицательно со шкалами нейротизма. Активность этой комбинации структур может рассматриваться в качестве фактора защиты от аффективных расстройств. В целом полученные результаты показывают плодотворность использования методов машинного обучения для анализа данных такого рода экспериментов.

*Ключевые слова:* “Я”, самооценка, оценка других людей, фМРТ, машинное обучение, многовоксельный анализ паттернов

DOI: 10.31857/S0044467723020065, EDN: ILHCWX

ВВЕДЕНИЕ

Природа “Я” – это один из самых сложных вопросов на протяжении всей истории философии и науки. Различные подходы подчеркивают разные аспекты этого конструкта, включая эмоциональное (Fossati et al., 2003), когнитивное (Turk et al., 2003) и социальное

(Frith, Frith, 2003) “Я”. Более того, некоторые авторы вообще отрицают его реальность, утверждая, что “Я” – это лишь иллюзия нашего восприятия (Hood, 2012). Несмотря на эти концептуальные противоречия, эмпирические данные свидетельствуют о наличии ряда областей мозга, которые специфически связаны с обработкой самореферентной ин-

формации (Northoff et al., 2006). В частности, в качестве системы, лежащей в основе человеческого “Я”, были предложены кортикальные срединные структуры (КСС), включающие различные части медиальной префронтальной коры (МПК), передней поясной коры (ППК), медиальной теменной коры, задней поясной коры (ЗПК) и ретроспленальной коры (Northoff, Vermpohl, 2004; Qin, Northoff, 2011). Однако многие исследования показывают, что обработка самореферентной информации связана не только с активностью КСС, но также задействует многие другие области мозга, такие как вентро- и дорсолатеральная префронтальная кора, височные области, височно-теменная кора, инсула и ряд подкорковых областей (Damasio, 2003; Gallagher, Frith, 2003; Gillihan, Farah, 2005; Finlayson-Short et al., 2020; Morin, Michaud, 2007; Northoff et al., 2006; Vanderwal et al., 2008).

Одним из наиболее популярных экспериментальных подходов к изучению активности мозга в процессе обработки самореферентной информации, в сравнении с обработкой информации, относящейся к другим людям, является задание, в котором участников просят решить, правильно ли предъявляемые оценочные прилагательные описывают либо самого участника, либо какого-то другого человека. Нейровизуализационные исследования с использованием этой парадигмы показали, что при вынесении суждений о себе, относительно суждений о других людях, наблюдается селективная активация области МПК (Heatherton et al., 2006; Kelley et al., 2002). На основе метаанализов фМРТ-исследований было также высказано предположение, что, в то время как вентральный МПК отвечает за обработку информации, относящейся к себе, задние срединные структуры, такие как ЗПК, могут преимущественно обрабатывать информацию, относящуюся к другим людям (Denny et al., 2012; Qin, Northoff, 2011). Использование показателей связности позволило выявить ключевую роль дефолтной сети в процессах обработки самореферентной информации, но также и участие сетей внимания и правой височно-теменной связки (Grigg, Grady, 2010; Knyazev et al., 2020, 2021; Soch et al., 2017).

Важным вопросом является наличие индивидуальных различий в процессах обработки самореферентной информации и информации о других людях. Эти различия могут быть связаны с личностными особенностями

(Knyazev et al., 2021) и могут быть факторами, предрасполагающими к аффективным расстройствам (Finlayson-Short et al., 2021; Hamilton et al., 2011; Knyazev et al., 2016, 2018). Общеизвестно, что в основе этиологии аффективных расстройств лежит взаимодействие индивидуальной предрасположенности с действием стрессовых факторов внешней среды (Ingram, Luxton, 2005). Из предрасполагающих факторов первостепенная роль отводится устойчивым личностным особенностям (Watson et al., 1994). Данные метаанализов показывают, что чрезмерное фокусирование на собственной персоне (ФСП) коррелирует с негативными эмоциями, а клиническая депрессия связана с его усилением (Grimm et al., 2009). Экспериментальная парадигма обработки самореферентной информации, в сравнении с обработкой информации, относящейся к другим людям, потенциально позволяет выявить мозговые корреляты ФСП и связать их с личностными особенностями, однако такие данные в мировой литературе отсутствуют.

Типичный подход к анализу данных в нейровизуализационных исследованиях заключается в применении так называемого массового одномерного анализа (mass-univariate statistical approach), который базируется на предположении, что активность в одном регионе мозга происходит независимо от активности в других регионах. При этом подходе выбранная статистическая модель (обычно общая линейная модель) последовательно тестируется на каждом вокселе объемного изображения активности мозга (в фМРТ-исследованиях это обычно локальный уровень оксигенации крови, или показатели связности), и затем для выявления статистически достоверных эффектов используется выбранный порог достоверности с учетом поправки на множественные сравнения. Дополнительно предполагается, что функционально значимые эффекты должны группироваться в плотно упакованные пространственные кластеры, поэтому вводится дополнительный порог достоверности по размеру кластера (Worsley et al., 1996). Хотя этот подход за многие годы принес много открытий в локализации функций мозга и продолжает оставаться наиболее часто используемым методом анализа нейровизуализационных данных, растет понимание того, что пространственные зависимости между сигналами, поступающими из разных областей мозга, должны быть надлежащим образом смоделированы. Кроме того,

некоторые эффекты могут быть пространственно распределенными по разным областям мозга и не соответствовать предположению об их группировке в плотно упакованные пространственные кластеры. Для выявления таких эффектов обычные инструменты могут оказаться недостаточно чувствительными. Поэтому растет интерес к исследованию этой пространственно распределенной информации с помощью многомерных подходов, которые часто называют многовоксельным анализом паттернов (МВАП, multi-voxel pattern analysis) (Haynes, Rees, 2006; Norman et al., 2006; Pereira et al., 2009). Использование МВАП в нейровизуализации привело к фундаментальному прогрессу в понимании того, как информация представлена в мозге (см. обзор (Князев, 2022)). К настоящему времени этот подход не использовался для анализа фМРТ-данных активности мозга в процессе обработки самореферентной информации в сравнении с обработкой информации, относящейся к другим людям, и задачей данной работы было восполнить этот пробел. Для этого мы использовали фМРТ-данные, полученные при выполнении задания по оцениванию себя и других людей. По этим данным были опубликованы две статьи (Knyazev et al., 2020, 2021), в которых показатели активации и связности анализировались с использованием массового одномерного анализа. В настоящей работе эти данные были переанализированы с использованием МВАП. Нашей целью было выявление паттернов активности мозга при обработке самореферентной информации в сравнении с обработкой информации, относящейся к другим людям разной степени близости. Дополнительно мы ставили себе целью связать индивидуальные различия в этих паттернах с личностными особенностями, потенциально предрасполагающими или защищающими от возникновения аффективных расстройств.

## МЕТОДИКА

### *Описание выборки*

Первоначально в исследовании приняли участие 50 человек. Большинство участников были студентами, аспирантами и сотрудниками Новосибирского государственного университета. Впоследствии трое участников были исключены из анализа из-за чрезмерных фМРТ-артефактов, таким образом, осталось 47 участников (26 женщин, средний воз-

раст 23.5 года, SD 4.9). Критериями исключения были серьезные медицинские заболевания, эпилепсия в анамнезе, злоупотребление психоактивными веществами или зависимость от них, а также все противопоказания к МРТ. Исследование соответствовало Хельсинкской декларации Всемирной медицинской ассоциации и было одобрено этическим комитетом Института нейронаук и медицины. Все участники дали письменное информированное согласие на участие в эксперименте.

### *Стимулы и экспериментальная процедура*

Предварительно был составлен список из 150 прилагательных, большинство из которых были взяты из личностных опросников, другие – из описаний внешности. Тридцать пять экспертов (преподаватели и студенты Гуманитарного факультета Новосибирского государственного университета) оценивали каждое прилагательное по 5-балльной шкале Лайкерта по эмоциональной валентности и эмоциональной активации. На основе средней оценки были отобраны 25 положительных, 25 нейтральных и 25 негативных прилагательных таким образом, чтобы они в среднем не отличались по длине и количеству гласных. Прилагательные предъявлялись визуально с помощью прямоугольного проекционного экрана с зеркалом и были представлены черным цветом в центре серого экрана.

Вначале (вне сканера) участников попросили выбрать среди своих знакомых трех человек, личности которых они не были обязаны раскрывать экспериментатору. Они должны были выбрать лучшего друга или близкого родственника, с которым в настоящее время у них были самые близкие и доверительные отношения (далее “Друг”), нейтрального незнакомца, с которым у них нет личных отношений (далее “Посторонний”), и человека, который им либо неприятен, либо они находятся с ним в плохих отношениях (далее “Враг”). Затем им объяснили предстоящее задание и представили тренировочную сессию, в которой они должны были оценить случайно выбранный объект (т.е. Я, Друг, Посторонний, или Враг). Далее, в сканере, процедура состояла из тех же четырех блоков, которые у разных испытуемых чередовались псевдослучайно. В начале каждого блока на экране появлялась подсказка (например, “Ты”, или “Друг”), которая оставалась на

экране в течение всего блока. Сначала участника просили подумать в течение одной минуты об этом человеке и вспомнить его/ее характеристики. В последующем задании, которое затем использовалось для анализа, испытуемым в случайном порядке предъявлялись все 75 прилагательных и предлагалось решить, свойственна ли оцениваемому человеку черта, соответствующая каждому прилагательному. Таким образом, порядок слов внутри блока был рандомизирован, и ни одно прилагательное не предъявлялось дважды. Задержка между началом предстоящего фМРТ-кадра и началом предъявления прилагательного варьировалась случайным образом между интервалами 100 и 2350 мс. Участники отвечали нажатием левой (Нет) или правой (Да) кнопки с помощью указательных пальцев левой и правой руки, и прилагательное мгновенно исчезало. Следующее предъявление начиналось через 5 с после начала предъявления предшествующего прилагательного. В сумме каждый блок длился 375 с.

#### *Запись и анализ фМРТ-данных*

фМРТ-данные были получены с помощью последовательности EPI на сканере Philips Ingenia 7FN8GDI 3.0 T. Первые пять сканов в начале каждого сеанса отбрасывались для обеспечения эффекта эквипирации сканера, оставляя 225 сканов в каждом из четырех блоков (TR 2.5 с, TE = 35 мс, flip angle = 90°, FOV = 100, матрица 96 × 94, 25 срезов толщиной 5 мм, без зазора). T1-взвешенные структурные сканы высокого разрешения (1 мм) были получены с помощью последовательности 3D MP-GR (TR = 7.8 мс, TE = 3.76 мс, матрица 252 × 227).

Предварительная обработка проводилась с помощью инструментария SPM-12 и включала коррекцию времени получения каждого среза (slice-time correction), выравнивание (realignment) с помощью трансформации по принципу жесткого тела (rigid body transformation), корегистрацию и нормализацию к шаблону Монреальского неврологического института (MNI), повторный сэмплинг (resampling) до разрешения 2 × 2 × 2 мм и сглаживание (full-width half-maximum, 6 мм). Мы проверяли параметры движения, и порог для качества изображений был установлен на уровне 2 мм для трех плоскостей трансляции. Все участники, которые превысили этот порог движения, были исключены из последу-

ющего анализа (всего было исключено три участника из первоначальной выборки 50 испытуемых). Затем на данных каждого испытуемого была использована общая линейная модель (general linear model, GLM) для расчета эффектов предъявления каждого прилагательного. В отличие от обычного GLM первого уровня, в котором сходные события (например, все стимулы определенной категории) рассматриваются как один фактор, в некоторых случаях, например, при анализе эффектов медиации (Atlas et al., 2014), а также при использовании методов машинного обучения (Schrouff et al., 2013) каждое событие рассматривается как уникальное, со своими входными и выходными характеристиками. В данном случае это позволяет проводить машинное обучение, используя вариацию входных и выходных характеристик. Поэтому была построена матрица дизайна GLM с отдельными регрессорами для каждого предъявления стимула. Каждый такой регрессор моделировался прямоугольной функцией (boxcar function) с точками включения и выключения, соответствующими времени предъявления прилагательного и времени, когда испытуемый нажимал кнопку ответа, соответственно. Проводилась конволюция этой функции с функцией канонического гемодинамического ответа (HRF). Мы использовали гибкий базисный набор, включающий не только канонический гемодинамический ответ, но и его производную по времени. Мы решили не использовать производную дисперсии, которая важна для моделирования медленных и длительных ответов, таких как ответы на тепловой стимул (Atlas et al., 2014), но не является столь важной для моделирования быстрых ответов на визуальные стимулы, увеличивая при этом коллинеарность матрицы дизайна. Поскольку некоторые эпохи данных могут быть загрязнены артефактами движения, мы рассчитали для каждой эпохи каждого испытуемого коэффициенты инфляции дисперсии (variance inflation factors, VIF) — меру коллинеарности каждой эпохи с артефактными регрессорами (т.е. оценками движения головы). Все эпохи с VIF больше 2 были исключены из анализа (среднее количество исключенных эпох = 0.13, SD = 0.33). Таким образом, в GLM-матрицу первого уровня каждого испытуемого были включены регрессоры для вызванных каждым стимулом ответов, 12 поперечных регрессоров движения головы (x, y, z, крен, шаг, рысканье и квадраты этих векто-

ров), вектор-индикатор для временных точек, оцененных как выбросы (см. предыдущий раздел), и столбец единиц. В качестве целевого эффекта рассматривалась конволюция BOLD (blood-oxygen-level-dependent) сигнала с HRF, а остальные регрессоры использовались как не представляющие интереса ковариаты (так называемые nuisance covariates), что является стандартным подходом в такого рода анализе, позволяющим выявить главный эффект под контролем всех потенциальных посторонних факторов.

Полученные в результате анализа первого уровня трехмерные образы весовых бета-коэффициентов для каждого события использовались в качестве входных сигналов для МВАП. Такой подход предпочтительнее, чем использование “сырых” фМРТ-данных, так как он позволяет устранить артефакты и берет в расчет форму гемодинамического ответа (Schrouff et al., 2013). GLM первого уровня делался с использованием инструментария SPM-12 и собственных скриптов. В результате входная матрица для МВАП содержала 47 испытуемых, по 300 NIFTI-файлов у каждого. Каждый NIFTI-файл представлял собой трехмерное изображение бета-коэффициентов, моделирующих фМРТ-ответ на каждое из предъявленных прилагательных в процессе оценки себя, друга, постороннего и неприятного человека.

МВАП проводился с использованием инструментария PRoNTo, созданного в среде Matlab специально для анализа нейровизуализационных данных (<http://www.ml-nl.cs.ucl.ac.uk/pronto>) (Schrouff et al., 2013). PRoNTo использует методы (так называемый kernel trick), позволяющие преодолеть проблему высокой размерности нейровизуализационных данных, позволяя, таким образом, использовать в качестве входных данных изображение всего мозга, а не выбирать какую-то область или наиболее активные воксели, как это часто делают при использовании методов МВАП для анализа фМРТ-данных. Мы в качестве маски первого уровня использовали маску серого вещества. Мы использовали МВАП для попарных различий активности мозга, сопровождающей оценивание себя в сравнении с оцениванием каждого из других объектов (друг, посторонний и враг), а также для различения “Я” и “Другие” (то есть любой из других людей). Для учета влияния отрицательной, нейтральной или положительной эмоциональной валентности прилагательных,

а также сделанного испытуемым выбора (“Да” или “Нет”) мы включили в дизайн ковариаты, в которых валентность каждого стимула кодировалась как  $-1$ ,  $0$  или  $1$ , а варианты ответа — как  $1$  или  $0$  соответственно.

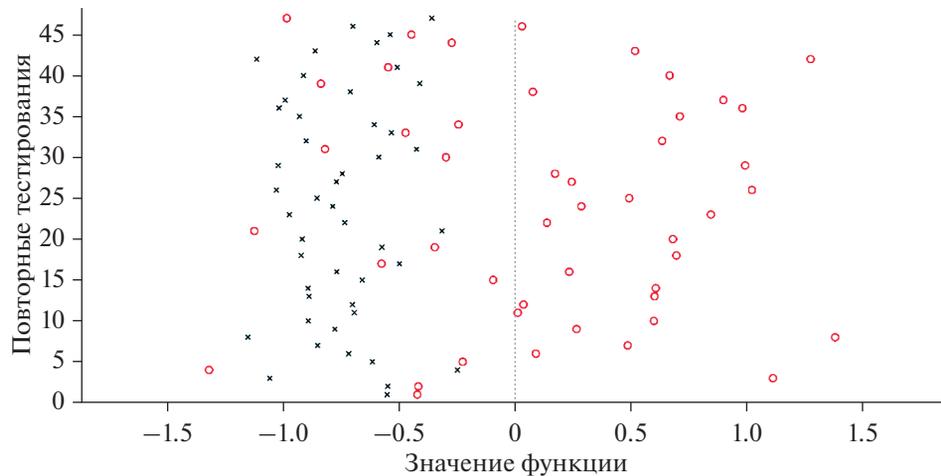
В качестве машины была использована машина опорных векторов (support-vector machine). Для определения эффективности модели использовалась перекрестная валидация по схеме leave-one-subject-out (LOSO), когда обучение машины проводится на данных всех испытуемых, кроме одного, а тестирование на оставшемся, и эта процедура повторяется для всех испытуемых. Для оценки статистической значимости использовался пермутационный тест ( $n = 1000$ ). Порог достоверности устанавливался с учетом поправки на множественные тесты. В нашем случае с учетом поправки Бонферрони порог достоверности был равен  $p = 0.05/4 = 0.0125$ .

Для выявления эффектов личностных переменных были использованы три шкалы из русскоязычных адаптаций международных опросников: шкалы Неполноценности и Подавленности из Личностного профиля по Айзенку (Knyazev et al., 2004a) и шкала чувствительности к награде из опросника Карвера и Уайта (Knyazev et al., 2004b). Первые две шкалы оценивают личностные особенности, предрасполагающие к аффективным расстройствам, а последняя — являющиеся фактором защиты.

## РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ

Мы провели серию бинарных классификаций с использованием машины опорных векторов. При классификации Я-Другие категория “Другие” включает в три раза больше данных, чем категория “Я”, что создает опасность того, что элементы малочисленного класса могут чаще распознаваться неправильно, чем элементы более многочисленного класса (Branco et al., 2016). PRoNTo предоставляет возможность использовать случайные подвыборки пере-представленного класса, чтобы как можно ближе соответствовать недопредставленному классу (Schrouff et al., 2013), и в данном случае мы использовали эту опцию.

Точность предсказаний в бинарных классификациях Я-Друг, Я-Посторонний, Я-Враг и Я-Другие составляла 76, 88, 80 и 78 процентов соответственно. Во всех случаях точность предсказаний была достоверно (на уровне  $p < 0.001$ )



**Рис. 1.** Распределение предсказаний в процессе повторных тестирований классификации категорий Я-другие. Положительные значения функции соответствуют категории “Я”, а отрицательные — категории “Другие”. Кружки и крестики показывают истинное распределение стимулов первой и второй категорий соответственно.

**Fig. 1.** Distribution of predictions in the retests of the self-other categories classification. Positive function values correspond to the self category, and negative values correspond to the other category. The circles and crosses show the true distribution of the first and second category stimuli, respectively.

выше 50% уровня случайных совпадений. Она максимальна для классификации Я-Посторонний и минимальна для классификации Я-Друг. На рис. 1 показано распределение предсказаний по шкале значений функции принятия решений в процессе повторных тестирований классификации Я-Другие. Можно отметить, что, хотя было некоторое количество ошибочных классификаций “Я” как “Другие”, не было ни одного случая, когда “Другие” ошибочно классифицировались как “Я”.

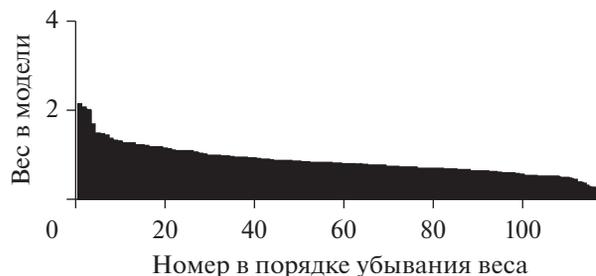
PRoNTo предоставляет возможность восстановления весов модели в исходном пространстве признаков (в данном случае это бета-коэффициенты всех вокселей внутри выбранной маски) и преобразования вектора весов в изображение, или карту. Эта карта содержит для каждого признака соответствующий вес линейной модели, который вместе с другими весами определяет прогностическую функцию и который связан с тем, насколько данный признак способствует решению задачи классификации. Важно отметить, что все воксели с отличным от нуля весом вносят вклад в функцию принятия решения, поскольку анализ является многомерным. Веса могут быть суммированы в пределах регионов, в соответствии с заданным атласом. Мы использовали атлас Automated Anatomical Labeling (AAL) (Tzourio-Mazoyer et al., 2002).

Для каждого региона определялся нормализованный вклад, а затем эти вклады ранжировались в порядке убывания. На гистограмме, представленной на рис. 2, видно, что при классификации Я-Другие наибольший вклад в функцию принятия решения дает небольшое количество областей, которые представлены в табл. 1. Первые три строчки занимают ППК и ЗПК, входящие в КСС (Qin, Northoff, 2011). На рис. 3 показаны все области мозга, дававшие вклад в функцию принятия решения при классификации Я-Другие. При использовании МВАП все воксели с весом, отличным от нуля, вносят вклад в функцию принятия решения, поскольку анализ является многомерным. Поэтому, вопреки общепринятой практике статистического параметрического картирования, которое является массовым одномерным подходом, здесь не рекомендуется делать пороговую карту весов (Schrouff et al., 2013). Видно, что наиболее “горячая” зона для выбора варианта “Я” расположена в области МПК и ППК, а для выбора варианта “Другие” — в теменных и затылочных срединных областях.

Поскольку существуют данные о различиях между мужчинами и женщинами в активности мозга при обработке стимулов, имеющих отношение к себе и другим людям (van Vuuren et al., 2020), мы дополнительно провели классификацию Я-Другие отдельно у муж-

чин и женщин. Точность классификации была 59.5 и 73.1% соответственно, и она достоверно отличалась от уровня случайных совпадений только во втором случае ( $p = 0.009$ ). Интересно отметить, что у мужчин, как и у женщин, не было ни одного случая ошибочной классификации другого как “Я”, но было гораздо больше, чем у женщин, классификаций “Я” как другой. В обеих группах наибольший вклад в функцию принятия решения давали ППК и ЗПК.

Для выявления связи личностных особенностей с паттернами активности мозга в процессе оценивания себя и других людей необходимо было “сжать” полученную при обучении модели информацию до приемлемого количества переменных. Для этого был применен анализ главных компонент, в качестве переменных для которого использовались усреднённые в пределах каждой области мозга из атласа AAL и нормированные на количество вокселей в этой области весовые коэффициенты, рассчитанные в каждой итерации при тестировании модели по методу LOSO. Таким образом, входной матрицей для анализа главных компонент была матрица  $n \times r$ , где  $n$  – количество итераций, равное количеству испытуемых, а  $r$  – количество обла-



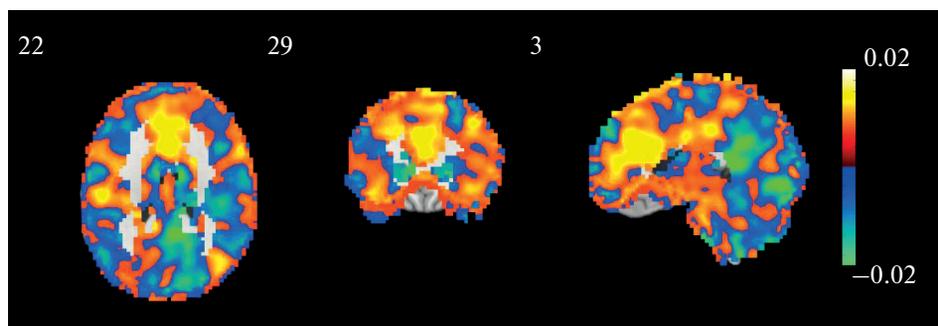
**Рис. 2.** Распределение весов областей мозга в функции принятия решений при классификации Я-Другие.

**Fig. 2.** Distribution of the weights of the brain areas in the decision-making function in the self-other classification.

стей мозга. Мы рассматривали лишь первый компонент, объясняющий наибольшее количество вариации (16%). Положительные нагрузки на этот компонент давали левая и правая ППК, правая миндалина и левая нижняя лобная извилина, а отрицательные – правая нижняя затылочная кора, левый и правый таламус и левая миндалина (рис. 4). Факторные оценки этого компонента коррелировали отрицательно со шкалами Неполноценности ( $r = -0.386$ ,  $p = 0.007$ ) и Подавленности

**Таблица 1.** Области мозга, давшие вклад более 1.2% в функцию принятия решения при классификации Я-Другие, рассортированные в порядке убывания относительного вклада в предсказательную силу модели  
**Table 1.** Brain regions that contributed more than 1.2% to the decision function for self-other classification, sorted in descending order of relative contribution to the predictive power of the model

Название области по атласу	Относительный вес в процентах	Размер (количество вокселей)
Cingulum_Ant_L	2.12	420
Cingulum_Ant_R	2.04	357
Cingulum_Post_R	1.98	77
Pallidum_L	1.67	62
Lingual_R	1.47	664
Parietal_Inf_R	1.46	382
Frontal_Inf_Oper_L	1.42	286
Occipital_Sup_R	1.34	382
Cingulum_Mid_L	1.32	557
Angular_L	1.29	329
Frontal_Sup_Medial_L	1.25	821
Cingulum_Mid_R	1.24	605
Frontal_Inf_Orb_L	1.24	493
Precuneus_R	1.20	893
Parietal_Sup_R	1.20	526



**Рис. 3.** Области мозга, давшие вклад в функцию принятия решения при классификации Я-Другие. Теплые тона показывают области, активность которых увеличивает вероятность того, что объект классификации – “Я”, а холодные – вероятность того, что объект классификации – “Другие”. Интенсивность цвета индексирует расстояние от границы классификации в сторону “Я” или “Другие”.

**Fig. 3.** Brain areas that have contributed to the decision function of self-other classification. Warm tones show areas whose activity increases the probability that the object of classification is “Me”, and cool tones show areas whose activity increases the probability that the object of classification is “Other”. Color intensity indexes the distance from the classification boundary toward self or other.

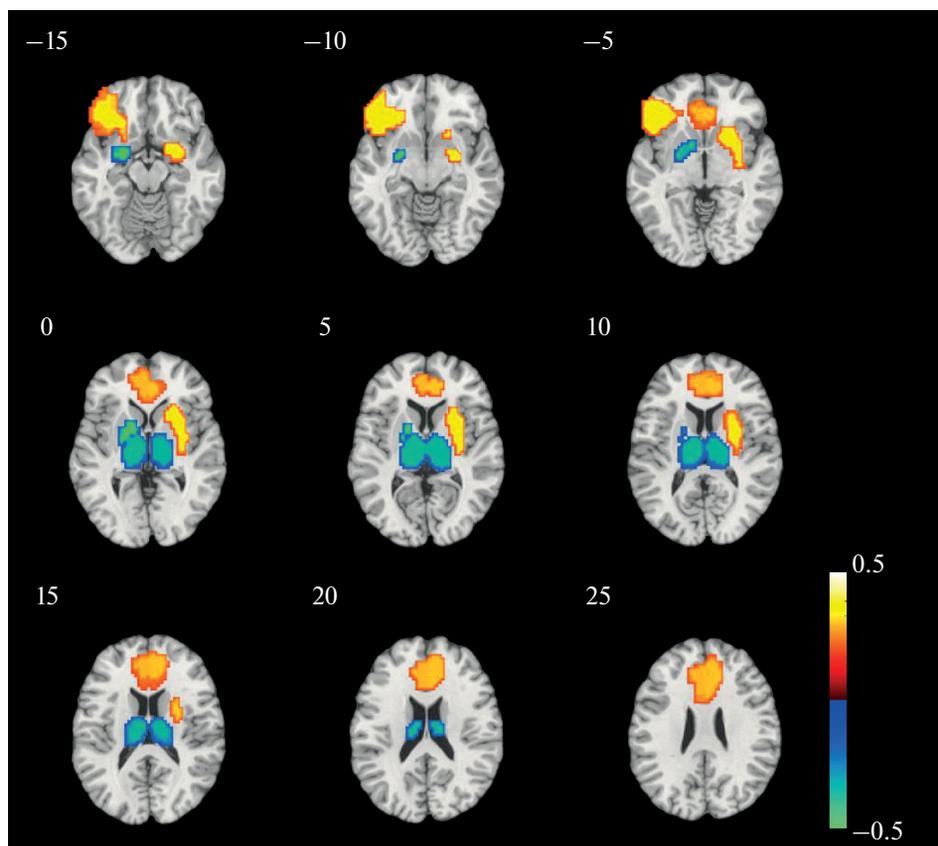
( $r = -0.329$ ,  $p = 0.024$ ) и положительно со шкалой чувствительности к награде ( $r = 0.508$ ,  $p < 0.001$ ).

## ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Целью данной работы было использование МВАП для выявления паттернов активности мозга, сопровождающей оценивание себя и других людей. Поскольку МВАП основан на методах машинного обучения, он включает две стадии – стадию обучения и стадию тестирования, позволяя тем самым оценить предсказательную силу модели. В нашем случае тестирование проводилось по методу LOSO, который позволяет оценить воспроизводимость выявленных паттернов в выборке испытуемых. Во всех протестированных попарных классификациях количество правильных идентификаций было достоверно выше уровня случайных совпадений и составляло 75–88% от общего числа случаев. Это типичный уровень правильных предсказаний в такого рода экспериментах (см., например, обзор (Князев, 2022)). Интересен тот факт, что оценка себя иногда сопровождалась таким же паттерном активности, как и оценка других людей, но обратного не наблюдалось ни разу. Можно думать, что оценивание себя может включать два процесса, один из которых интроспекция, а второй – оценивание себя как бы со стороны. Если оба эти процесса имеют отражение в активности мозга, то машина обучается классифицировать эту активность как оценивание себя, ес-

ли в ней преобладает паттерн, характерный для интроспекции, и как оценивание других, если он отсутствует. Этот паттерн всегда отсутствует при оценивании других, и машина безошибочно классифицирует эти случаи в соответствующую категорию, но если при оценивании себя преобладает второй процесс, то машина может ошибочно классифицировать его как оценивание другого человека. Интересно, что по нашим данным этот вариант гораздо чаще встречается у мужчин, чем у женщин, то есть можно думать, что в среднем женщины при оценке себя более склонны анализировать свои внутренние состояния и чувства, а мужчины более склонны смотреть на себя “со стороны”, что затрудняет работу классификатора в последнем случае. Это, безусловно, пока лишь спекуляция, и этот вопрос требует дальнейшего изучения на больших выборках.

Результаты нашего анализа показывают, что различия при оценивании себя и другого человека минимальны, если это близкий человек, и существенно больше, если это посторонний. Это соответствует полученным нами ранее данным анализа связности, которые тоже показали схожесть паттернов связности при оценивании себя и близкого человека, и их отличие от паттернов, сопровождающих оценивание посторонних людей (Knyazev et al., 2020). Это согласуется с представлением о том, что в сферу “Я” человек включает не только непосредственно себя, но и многое другое, с чем он себя ассоциирует,



**Рис. 4.** Области мозга, давшие наиболее сильные положительные (теплые тона) и отрицательные (холодные тона) нагрузки на первый компонент в анализе главных компонент, проведенном на матрице весовых коэффициентов, полученных при обучении модели в классификации Я-Другие.

**Fig. 4.** Brain regions that showed the strongest positive (warm tones) and negative (cold tones) loadings on the first component in the principal component analysis performed on the weights matrix obtained by training the model in the self-other classification.

прежде всего близких и любимых людей (Gallagher, 2013).

Помимо оценки предсказательной силы модели, наибольший интерес при использовании МВАП представляет выявляемый паттерн активности мозга. В отличие от методов, основанных на массовом одномерном анализе, которые выявляют ограниченный набор структур, участвующих в самореферентных процессах, наши данные показывают, что эти структуры широко распределены по разным областям мозга. Прежде всего это области КСС, включая разные части поясной извилины, участие которых в процессах обработки самореферентной информации — твердо установленный факт (Northoff, Vermpohl, 2004; Qin, Northoff, 2011). Некоторые авторы критикуют теорию Норсоффа о том, что КСС является сидением “Я”, на основе того, что метаанализы Норсоффа с соавторами рассматривали лишь работы, в которых воспри-

ятие “Я” контрастировалось с восприятием других людей и, соответственно, по мнению этих авторов, КСС не является корковым представительство “Я” в широком смысле, а лишь участвует в различии “Я” — не “Я” (Gillihan, Farah, 2005; Legrand, Ruby, 2009; Vogetley, Gallagher, 2011). Результаты нашего анализа укладываются в это представление, но показывают также, что и различие “Я” — не “Я” вовлекает большое количество разных областей мозга, помимо КСС.

Весомый вклад зрительной коры (см. табл. 1 и рис. 3) в успешность классификации паттернов связан, очевидно, с тем, что информация предъявлялась через зрительный канал. Участие сенсорных зон коры в разнообразных когнитивных процессах — обычная находка в исследованиях с использованием МВАП. Показано, что даже кодирование абстрактных понятий может вовлекать сенсорные и моторные области коры (Barsalou et al.,

2003). Наконец, вовлечение латеральной префронтальной коры (см. рис. 4) в процессы различения “Я” – не “Я” также многократно показано ранее (Damasio, 2003; Gallagher, Frith, 2003; Gillihan, Farah, 2005; Morin, Michaud, 2007; Northoff et al., 2006; Vanderwal et al., 2008). Функционально эти области связаны с процессами регуляции внимания, исполнительного контроля и рабочей памяти, и их участие в выполнении практически любого экспериментального задания не удивительно. Однако полученные нами ранее данные указывают на то, что они могут быть более специфично связаны с процессами различения “Я” – не “Я”. Было выявлено, что если оценивание себя и близкого человека в большей степени связано с вовлечением дефолтной сети, непосредственно участвующей в процессах самореференции и интроспекции, то оценивание посторонних людей связано с доминированием сетей регуляции внимания (Knyazev et al., 2020). Можно думать, что посторонние люди относятся как бы к объектам внешнего мира и их восприятие активирует процессы направленного вовне внимания, в то время как восприятие себя и близкого человека активирует самореферентные процессы, связанные с активностью дефолтной сети и КСС. Таким образом, в восприятии себя и других людей важную роль играет баланс активности КСС и латеральной префронтальной коры.

По нашим данным наиболее информативными для выбора варианта “Я” являются вентральные области МПК и ППК, а для выбора варианта “Другие” – теменные и затылочные срединные области (см. рис. 3). Это хорошо согласуется с предположением, сделанным на основе метаанализа фМРТ-исследований, согласно которому вентральная МПК отвечает за обработку информации, относящейся к себе, а задние срединные структуры, такие как ЗПК и прилегающие к нему части клиновидной коры, могут преимущественно обрабатывать информацию, относящуюся к другим (Denny et al., 2012; Qin, Northoff, 2011). Пфайфер с соавторами показали, что относительное участие МПК и задних срединных структур в обработке информации о себе и других меняется в процессе развития (Pfeifer et al., 2007).

Дополнительной целью нашего исследования была попытка связать индивидуальные различия в паттернах активности мозга при обработке самореферентной информации с

личностными особенностями, потенциально предрасполагающими или защищающими от возникновения аффективных расстройств. Постановка этой цели основывалась на том, что обработка самореферентной информации, в сравнении с обработкой информации, относящейся к другим людям, потенциально позволяет выявить мозговые корреляты ФСП, являющегося фактором риска возникновения негативных эмоций и депрессии (Grimm et al., 2009). Анализ главных компонент позволил выявить комбинацию структур мозга в виде компонента, объясняющего наибольшее количество межиндивидуальной вариации паттернов весовых коэффициентов, рассчитанных в процессе обучения машины для различения категорий “Я” и “Другие”. Этот компонент включал положительные нагрузки ППК и правой миндалины и отрицательные нагрузки правой нижней затылочной коры и левой миндалины (см. рис. 4) и коррелировал положительно со шкалой чувствительности к награде и отрицательно со шкалами нейротизма. Нужно отметить, что в использованной для анализа главных компонент  $n \times r$ -матрице весовых коэффициентов каждая  $i$ -тая строка содержит коэффициенты, рассчитанные в данной итерации на всех испытуемых за вычетом  $i$ -го. Это нужно учитывать при интерпретации корреляций факторных оценок с личностными переменными и рассматривать их с противоположным знаком. Таким образом, активность выявленной комбинации структур мозга в процессе оценивания себя и других людей может рассматриваться в качестве фактора риска аффективных расстройств.

Роль вентромедиальной префронтальной коры и миндалины в генерации и регуляции негативных эмоций хорошо известна. Многократно показано, например, что у носителей  $s$ -аллели гена транспортера серотонина, демонстрирующих предрасположенность к аффективным расстройствам, изменена активность этих областей мозга и нарушена связь между ними (Dannlowski et al., 2010; Ma et al., 2014; Rao et al., 2007). Интересно, что нагрузки на этот компонент были противоположны для правой и левой миндалины – положительные в первом случае и отрицательные во втором. Межполушарная асимметрия как структурных, так и функциональных характеристик миндалины – хорошо установленный факт (Ocklenburg et al., 2022). Традиционно, начиная с работ Дэвидсона,

принято связывать правое полушарие с негативными, а левое — с положительными эмоциями (Davidson, 1992). Однако работы Дэвидсона и его последователей имели отношение в основном к фронтальным асимметриям. Что касается миндалины, то она преимущественно вовлечена в генерацию негативных эмоций, связанных с восприятием потенциальной угрозы (Zald, 2003), и метаанализы нейровизуализационных исследований показывают большую реактивность левой, чем правой миндалины, в ответ на стимулы угрозы (Fusar-Poli et al., 2009; Murphy et al., 2003; Wager et al., 2003). Кроме того, согласно гипотезе Марковича, левая миндалина сильнее, чем правая, реагирует на вербальные стимулы угрозы (Markowitsch, 1998), что подтверждается экспериментальными данными (Phelps et al., 2001), а в нашей работе мы использовали именно вербальные стимулы.

В качестве ограничения данной работы нужно отметить, что сбор экспериментальных данных не был изначально рассчитан на применение методов машинного обучения, поэтому в экспериментальную схему не был заложен повторный сбор данных, которые можно было бы использовать для внутрисубъектного тестирования предсказательной силы модели. Нам пришлось ограничиться поэтому межсубъектным тестированием по методу LOSO. Несмотря на это, полученные результаты показывают плодотворность использования методов МВАП в такого рода экспериментах. Выявленные паттерны активности были достоверными и существенно углубляют понимание механизмов мозгового обеспечения процессов оценивания себя и других людей разной степени близости.

## ВЫВОДЫ

1. Во всех протестированных попарных классификациях количество правильных идентификаций было достоверно выше уровня случайных совпадений и составляло 75–88% от общего числа случаев.

2. Активность мозга отличалась при оценивании себя и других людей, и степень отличия зависела от степени близости человека. Различия были минимальны при оценивании близкого человека.

3. Прогностически значимые структуры были широко распределены по разным областям мозга и, помимо КСС, дававших наибольший вклад, включали области зритель-

ной, латеральной префронтальной и многих других корковых зон.

4. При классификации “Я”–“Другие” наиболее информативными для выбора варианта “Я” являлись вентральные области МПК и ППК, а для выбора варианта “Другие” — теменные и затылочные срединные области.

5. Факторные оценки выявленной с помощью анализа главных компонент комбинации структур мозга, включающей ППК, левую и правую миндалины и правую нижнюю затылочную кору, коррелировали положительно со шкалой чувствительности к награде и отрицательно со шкалами нейротизма. Активность этой комбинации структур может рассматриваться в качестве фактора риска аффективных расстройств.

6. Полученные результаты показывают плодотворность использования методов машинного обучения для анализа данных такого рода экспериментов.

## ФИНАНСИРОВАНИЕ

Исследование выполнено при финансовой поддержке Российского научного фонда (проект № 22-15-00142).

## СПИСОК ЛИТЕРАТУРЫ

- Князев Г.Г. Кодирование смысла в активности мозга. Журн. высш. нервн. деят. им. И.П. Павлова. 2022. 72 (6): 800–825.
- Atlas L.Y., Lindquist M.A., Bolger N., Wager T.D. Brain mediators of the effects of noxious heat on pain. *Pain*. 2014. 155: 1632–1648.
- Barsalou L.W., Simmons W.K., Barbey A.K., Wilson C.D. Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*. 2003. 7: 84–91.
- Branco P., Torgo L., Ribeiro R.P. A Survey of Predictive Modeling on Imbalanced Domains. *ACM Computing Surveys*. 2016. 49: 1–50.
- Damasio A. Mental self: the person within. *Nature*. 2003. 423:227.
- Dannlowski U., Konrad C., Kugel H., Zwitterlood P., Domschke K., Schoning S., Ohrmann P., Bauer J., Pyka M., Hohoff C., Zhang W.Q., Baune B.T., Heindel W., Arolt V., Suslow T. Emotion specific modulation of automatic amygdala responses by 5-HTTLPR genotype. *Neuroimage*. 2010. 53: 893–898.

- Davidson R.J.* Anterior cerebral asymmetry and the nature of emotion. *Brain and Cognition*. 1992. 20: 125–151.
- Denny B.T., Kober H., Wager T.D., Ochsner K.N.* A meta-analysis of functional neuroimaging studies of self and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. *J. Cognitive Neuroscience*. 2012. 24: 1742–1752.
- Finlayson-Short L., Davey C.G., Harrison B.J.* Neural correlates of integrated self and social processing. *Social cognitive and affective neuroscience*. 2020. 15: 941–949.
- Finlayson-Short L., Harrison B.J., Davey C.* Self-other referential neural processing in social anxiety disorder and major depressive disorder. *NeuroImage: Clinical*. 2021. 30: 102669.
- Frith U., Frith C.D.* Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*. 2003. 358: 459–473.
- Fossati P., Hevenor S.J., Graham S.J., Grady C., Keightley M.L., Craik F., Mayberg H.* In search of the emotional self: an fMRI study using positive and negative emotional words. *American J. Psychiatry*. 2003. 160: 1938–1945.
- Fusar-Poli P., Placentino A., Carletti F., Allen P., Landi P., Abbamonte M., Barale F., Perez J., McGuire P., Politi P.L.* Laterality effect on emotional faces processing: ALE meta-analysis of evidence. *Neuroscience letters*. 2009. 452 (3): 262–267.
- Gallagher S.* A pattern theory of self. *Frontiers in Human Neuroscience*. 2013. 7: 443.
- Gallagher H.L., Frith C.D.* Functional imaging of ‘theory of mind’. *Trends in Cognitive Sciences*. 2003. 7: 77–83.
- Gillihan S.J., Farah M.J.* Is self special? A critical review of evidence from experimental psychology and cognitive neuroscience. *Psychological Bulletin*. 2005. 131: 76–97.
- Grigg O., Grady C.L.* The default network and processing of personally relevant information: Converging evidence from task-related modulations and functional connectivity. *Neuropsychologia*. 2010. 48: 3815–3823.
- Grimm S., Ernst J., Boesiger P., Schuepbach D., Hell D., Boeker H., Northoff G.* Increased self-focus in major depressive disorder is related to neural abnormalities in subcortical-cortical midline structures. *Human Brain Mapping*. 2009. 30: 2617–2627.
- Hamilton J.P., Furman D.J., Chang C., Thomason M.E., Dennis E., Gotlib I.H.* Default-mode and task-positive network activity in major depressive disorder: implications for adaptive and maladaptive rumination. *Biological Psychiatry*. 2011. 70: 327–333.
- Haynes J.D., Rees G.* Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*. 2006. 7: 523–534.
- Heatherton T.F., Wyland C.L., Macrae C.N., Demos K.E., Denny B.T., Kelley W.M.* Medial prefrontal activity differentiates self from close others. *Social Cognitive and Affective Neuroscience*. 2006. 1: 18–25.
- Hood B.* The self illusion: How the social brain creates identity. Oxford, UK: Oxford University Press. 2012.
- Ingram R.E., Luxton D.D.* Vulnerability-stress models. Development of psychopathology: A vulnerability-stress perspective. 2005. 46.
- Kelley W.M., Macrae C.N., Wyland C.L., Caglar S., Inati S., Heatherton T.F.* Finding the self? An event-related fMRI study. *J. Cognitive Neuroscience*. 2002. 14: 785–794.
- Knyazev G.G., Savostyanov A.N., Bocharov A.V., Rudych P.D.* Intrinsic Connectivity Networks in the Self- and Other-Referential Processing. *Frontiers in Human Neuroscience*. 2020. 14: 579703.
- Knyazev G.G., Savostyanov A.N., Bocharov A.V., Rudych P.D.* How self-appraisal is mediated by the brain. *Frontiers in Human Neuroscience*. 2021. 15: 700046.
- Knyazev G.G., Savostyanov A.N., Bocharov A.V., Tamozhnikov S.S., Saprigyn A.E.* Task-positive and task-negative networks and their relation to depression: EEG beamformer analysis. *Behavioral Brain Research*. 2016. 306: 160–169.
- Knyazev G.G., Savostyanov A.N., Bocharov A.V., Brak I.V., Osipov E.A., Filimonova E.A., Saprigyn A.E., Aftanas L.I.* Task-positive and task-negative networks in major depressive disorder: A combined fMRI and EEG study. *J. Affective Disorders*. 2018. 235: 211–219.
- Knyazev G.G., Belopolsky V.I., Bodunov M.V., Wilson G.D.* The factor structure of the Eysenck Personality Profiler in Russia. *Personality and Individual Differences*. 2004a. 37: 1681–1692.
- Knyazev G.G., Slobodskaya H.R., Wilson G.D.* Comparison of the construct validity of the Gray–Wilson Personality Questionnaire and the BIS/BAS scales. *Personality and Individual Differences*. 2004b. 37: 1565–1582.
- Legrand D., Ruby P.* What is self-specific? Theoretical investigation and critical review of neuroimaging results. *Psychological Reviews*. 2009. 116: 252–282.

- Ma Y.N., Li B.F., Wang C.B., Shi Z.H., Sun Y., Sheng F., Zhang Y.F., Zhang W.X., Rao Y., Han S.H.* 5-HTTLPR polymorphism modulates neural mechanisms of negative self-reflection. *Cerebral Cortex*. 2014. 24: 2421–2429.
- Markowitsch H.J.* Differential contribution of right and left amygdala to affective information processing. *Behavioural Neurology*. 1998. 11 (4): 233–244.
- Morin A., Michaud J.* Self-awareness and the left inferior frontal gyrus: Inner speech use during self-related processing. *Brain Research Bulletin*. 2007. 74: 387–396.
- Murphy F.C., Nimmo-Smith I., Lawrence A.D.* Functional neuroanatomy of emotions: A meta-analysis. *Cognitive, Affective, & Behavioral Neuroscience*. 2003. 3 (3): 207–233.
- Norman K.A., Polyn S.M., Detre G.J., Haxby J.V.* Beyond mind-reading: multi-voxel pattern analysis of fMRT data. *Trends in Cognitive Science (Regul. Ed.)*. 2006. 10: 424–430.
- Northoff G., Bermpohl F.* Cortical midline structures and the self. *Trends in Cognitive Sciences*. 2004. 8: 102–107.
- Northoff G., Heinzel A., de Greck M., Bermpohl F., Dobrowolny H., Panksepp J.* Self-referential processing in our brain – A meta-analysis of imaging studies on the self. *NeuroImage*. 2006. 31: 440–457.
- Ocklenburg S., Peterburs J., Mundorf A.* Hemispheric asymmetries in the amygdala: A comparative primer. *Progress in Neurobiology*. 2022. 214: 102283.
- Pereira F., Mitchell T., Botvinick M.* Machine learning classifiers and fMRT: a tutorial overview. *Neuroimage*. 2009. 45: 199–209.
- Pfeifer J.H., Lieberman M.D., Dapretto M.* “I know who you are but what am I?!”: Neural bases of self- and social knowledge retrieval in children and adults. *J. Cognitive Neuroscience*. 2007. 19: 1323–1337.
- Phelps E.A., O'Connor K.J., Gatenby J.C., Gore J.C., Grillon C., Davis M.* Activation of the left amygdala to a cognitive representation of fear. *Nature Neuroscience*. 2001. 4 (4): 437–441.
- Qin P., Northoff G.* How is our self related to midline regions and the default-mode network? *NeuroImage*. 2011. 57: 1221–1233.
- Rao H.Y., Gillihan S.J., Wang J.J., Korczykowski M., Sankoorikal G.M.V., Kaercher K.A., Brodtkin E.S., Detre J.A., Farah M.J.* Genetic variation in serotonin transporter alters resting brain function in healthy individuals. *Biological Psychiatry*. 2007. 62: 600–606.
- Schrouff J., Rosa M.J., Rondina J.M., Marquand A.F., Chu C., Ashburner J., Phillips C., Richiardi J., Mourao-Miranda J.* PRoNTo: Pattern Recognition for Neuroimaging Toolbox. *Neuroinformatics*. 2013. 11 (3): 319–337.
- Soch J., Deserno L., Assmann A., Barman A., Walter H., Richardson-Klavehn A., Schott B.H.* Inhibition of information flow to the default mode network during self-reference versus reference to others. *Cerebral Cortex*. 2017. 27: 3930–3942.
- Tzourio-Mazoyer N., Landeau B., Papathanassiou D., Crivello F., Etard O., Delcroix N., Mazoyer B., Joliot M.* Automated Anatomical Labeling of activations in SPM using a Macroscopic Anatomical Parcellation of the MNI MRI single-subject brain. *NeuroImage*. 2002. 15 (1): 273–289.
- Turk D.J., Heatherton T.F., Macrae C.N., Kelley W.M., Gazzaniga M.S.* Out of contact, out of mind: the distributed nature of the self. *Annals of the New York Academy of Sciences*. 2003. 1001: 65–78.
- van Buuren M., Walsh R.J., Sijtsma H., Hollarek M., Lee N.C., Bos P.A., Krabbendam L.* Neural correlates of self- and other-referential processing in young adolescents and the effects of testosterone and peer similarity. *Neuroimage*. 2020. 219: 117060.
- Vanderwal T., Hunyadi E., Grupe D.W., Connors C.M., Schultz R.T.* Self, Mother and Abstract Other: an fMRI study of reflective social processing. *Neuroimage* 2008. 41: 1437–1446.
- Vogeley K., Gallagher S.* The self in the brain. In *The Oxford Handbook of the Self* (ed. S. Gallagher). 2011. Oxford: Oxford University Press. 111–136.
- Wager T.D., Phan K.L., Liberzon I., Taylor S.F.* Valence, gender, and lateralization of functional brain anatomy in emotion: a meta-analysis of findings from neuroimaging. *Neuroimage*. 2003. 19 (3): 513–531.
- Watson D., Clark L.A., Harkness A.R.* Structures of personality and their relevance to psychopathology. *J. Abnormal Psychology*. 1994. 103: 18–31.
- Worsley K.J., Marrett S., Neelin P., Vandal A.C., Friston K.J., Evans A.C.* A unified statistical approach or determining significant signals in images of cerebral activation. *Human Brain Mapping*. 1996. 4: 58–73.
- Zald D.H.* The human amygdala and the emotional evaluation of sensory stimuli. *Brain Research Reviews*. 2003. 41 (1): 88–123.

## MULTI-VOXEL PATTERN ANALYSIS OF fMRI DATA DURING SELF-AND OTHER-REFERENTIAL PROCESSING

G. G. Knyazev<sup>a, #</sup>, A. N. Savostyanov<sup>a, b, c</sup>, P. D. Rudych<sup>a</sup>, and A. V. Bocharov<sup>a, b</sup>

<sup>a</sup> Federal State Budgetary Scientific Institution "Scientific Research Institute of Neurosciences and Medicine",  
Novosibirsk, Russia

<sup>b</sup> Novosibirsk State University, Novosibirsk, Russia

<sup>c</sup> Federal Research Center Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia

<sup>#</sup>e-mail: knyazev@physiol.ru

The study of brain activity in the processing of self-referential information, as compared to the processing of information related to other people, is based on the application of mass-univariate analysis, based on the assumption that activity in one region is independent of activity in other regions. Recently, there has been a growing interest in neuroimaging to investigate spatially distributed information using multivariate approaches such as multivoxel pattern analysis (MVPA). In this paper, we used MVPA to analyze fMRI data recorded during self-evaluation and evaluation of other people of varying proximity. In all pairwise classifications tested, the number of correct identifications was significantly higher than the level of random matches. Predictively significant structures were widely distributed over different brain regions and included areas of the visual, lateral prefrontal, and many other cortical areas in addition to the cortical midline structures that contributed the most. In the self-other classification, ventral areas of the medial prefrontal and cingulate cortex were the most informative for the self condition, whereas parietal and occipital medial areas were the most informative for the other condition. The combination of brain structures, which included the anterior cingulate cortex and both amygdalae, revealed by principal component analysis, correlated positively with the psychometric scale of sensitivity to reward, and negatively with neuroticism scales. Overall, the results show the fruitfulness of using machine learning methods to analyze data from such kinds of experiments.

*Keywords:* self, self-evaluation, other-evaluation, fMRI, machine learning, multi-voxel pattern analysis