

УДК 159.91

## ОБЗОР ОПЫТА ИСПОЛЬЗОВАНИЯ ПОТЕНЦИАЛОВ, СВЯЗАННЫХ С СОБЫТИЕМ, В ИССЛЕДОВАНИЯХ ИМПЛИЦИТНОГО КОМПОНЕНТА КОГНИТИВНЫХ ИСКАЖЕНИЙ

© 2024 г. М. В. Яценко<sup>1, 2</sup>, И. В. Брак<sup>3</sup>, Е. Д. Артеменко<sup>1, \*</sup>

<sup>1</sup>Лаборатория социальной и когнитивной информатики,  
Национальный исследовательский университет «Высшая школа экономики», Санкт-Петербург, Россия

<sup>2</sup>Алтайский государственный университет, Барнаул, Россия

<sup>3</sup>Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия  
\*e-mail: edartemenko@hse.ru

Поступила в редакцию 29.02.2024 г.

После доработки 17.04.2024 г.

Принята к публикации 22.07.2024 г.

В настоящем обзоре рассматриваются нейрофизиологические корреляты имплицитного компонента когнитивных искажений в процессах восприятия и принятия решений. В обзоре определен ведущий методический алгоритм для анализа имплицитного компонента когнитивных искажений, обоснован его выбор с точки зрения целей данного обзора. В качестве основного подхода к определению нейрофизиологических показателей имплицитных процессов выбран метод регистрации потенциалов, связанных с событием (ВП). Анализ литературных данных позволил выявить компоненты ВП, воспроизводимые в работах разных авторов в исследованиях различных вариантов неявной предвзятости при восприятии и принятии решений у человека, что может свидетельствовать о наличии общих нейрофизиологических механизмов, связанных с имплицитными процессами предвзятости при когнитивных искажениях. В обзоре также обсуждаются перспективы использования других подходов к анализу данных ЭЭГ для получения новой информации о феномене неявной предвзятости при принятии решений.

**Ключевые слова:** когнитивные искажения, имплицитный компонент, потенциалы, связанные с событиями, ВП, тест имплицитных ассоциаций, ИАТ.

**DOI:** 10.31857/S0044467724040012

### ВВЕДЕНИЕ

Феномен когнитивных искажений сопровождает человека на протяжении всей жизни и чаще всего проявляется в условиях восприятия и анализа текущей и критической информации, формируя характерное поведение. Данный эффект обусловлен прежде всего субъективной картиной реальности, формируемой особенностями функционирования рецепторного аппарата человека, особенностями протекания нервных процессов, типологическими особенностями, жизненным опытом социального взаимодействия, в том числе формированием субъективных убеждений. Когнитивные искажения, с одной стороны, могут приводить к превалированию у человека иррационального мышления, неточностям в интерпретации событий, ошибочным умозаключениям и др. С другой стороны, существование данного феномена может

быть обусловлено с эволюционной точки зрения, поскольку шаблонное, стереотипное поведение при искаженном восприятии информации давало некоторое преимущество, обеспечив жизнеспособность феномена и позволив ему сохраниться до настоящего времени (см., например, Peters, 2022).

Д. Канеман и А. Тверски в 1972 г. ввели понятие «cognitive bias» («когнитивные искажения»). С того времени количество исследований, посвященных данному вопросу, все возрастает. Согласно определению, когнитивные искажения представляют собой систематические ошибки, шаблонные отклонения, которые возникают у человека в процессе мышления на основании убеждений, присутствующих в когнитивных процессах (Канеман, 2016), в частности в процессах принятия решения. Такой важный аспект в жизни человека, как принятие решения, зачастую оказывается обусловлен не рациональными рассуждениями, а спонтанным,

основанным на искаженном восприятии информации процессом. Основные разновидности когнитивных искажений известны: эффект прайминга, ошибка поспешного вывода, профессиональная деформация, прокрастинация, эффект отрицания вероятности, предвзятость подтверждения и др. (Боброва, 2021; Le Bouc, Pessiglione, 2022). Таким образом, актуальность предпринятого обзора обусловлена все возрастающим интересом к данной проблеме со стороны психологов, нейрофизиологов, социологов, маркетологов и специалистов других областей, описывающих данный феномен с применением самого различного инструментария.

В процессе когнитивной обработки информации можно выделить компоненты, которые происходят неосознанно, то есть имплицитно (Barsade et al., 2009). Имплицитные процессы оказывают влияние на конечный результат восприятия информации, и особую роль в этом играют имплицитные когнитивные искажения, или неявные предвзятости. Неявная предвзятость — это бессознательная ассоциация или убеждение по отношению к определенной информации (Bhat et al., 2021). Неявная предвзятость включает в себя подсознательные чувства, отношения, предрассудки и стереотипы, которые развиваются и запечатлеваются у человека на протяжении всей жизни. В результате на принятие решений влияют не только факты и наблюдения, но в том числе подсознательные реакции. Таким образом, имплицитные когнитивные искажения могут лежать в основе формирования потенциально неоптимального, дезадаптивного поведения (Shah, Bohlen, 2024).

Появление понятия имплицитных когнитивных искажений привело к появлению запроса на разработку методов, позволяющих идентифицировать, выделить и подвергнуть анализу данный феномен. В настоящее время известно не менее 20 процедур, которые позволяют решить этот вопрос, однако наиболее востребованным методом является тест имплицитных ассоциаций (далее — IAT) (Greenwald et al., 1998). Так, согласно данным Nosek B.A. и соавт. (Nosek et al., 2011), на применение IAT в исследованиях неявной предвзятости в 2010 г. приходилась почти половина публикаций, а за все время — более 40% работ. Проведенный нами обзор исследований показал, что в подавляющем числе работ именно тест имплицитных ассоциаций используется в качестве инструмента выявления и фиксации имплицитного компонента когнитивного процесса, а процедура IAT используется для поиска нейрофизиологических коррелятов имплицитных процессов, что свидетельствует об универсальном характере предложенной методики.

Несмотря на существование достаточного количества подходов к выделению нейрокоррелятов, при исследованиях имплицитных процессов

преимущественно используется метод регистрации потенциалов, связанных с событием (ВП). Применение такого подхода обусловлено тем, что связанные с событием потенциалы обеспечивают непрерывное измерение между целевым стимулом и реакцией, позволяя выделить эффект экспериментальной манипуляции на определенном этапе обработки (Pesciarelli et al., 2021). Таким образом, связка IAT и ВП нашла широкое применение в изучении имплицитных процессов когнитивных искажений, при этом устойчивость выделяемых компонентов ВП неявной предвзятости продемонстрирована на совершенно различном материале. Вышесказанное обуславливает актуальность оценки и выявления общих, стабильно воспроизводимых нейрофизиологических коррелятов имплицитных процессов.

Таким образом, целью настоящей работы является анализ опыта использования потенциалов, связанных с событием, в исследованиях имплицитного компонента когнитивных искажений, выявляемого при помощи теста имплицитных ассоциаций.

#### *Инструментарий для определения имплицитных процессов*

Тест Имплицитных Ассоциаций (IAT) был разработан Э. Гринвальдом, Д. МакГи и Дж. Шварцем (Greenwald et al., 1998) с целью выявления различий между неосознаваемой оценкой каких-либо событий и отношения к ним, с одной стороны, и реакцией и поведением человека, с другой. Применение IAT позволяет выявить неосознаваемые когнитивные искажения, которые проявляются в виде имплицитных стереотипов, обусловленных убеждениями и предвзятостью человека. Ключевым элементом тестов, выявляющих имплицитные процессы, является измерение времени реакции, которое может быть различным при предъявлении конгруэнтных и неконгруэнтных пар стимулов. В случае стереотипных условий задач оно будет сокращаться, так как ответ в таком случае будет обусловлен автоматическими процессами обработки и реагирования.

Кроме IAT, для решения специфичных задач по выявлению имплицитных ассоциаций был разработан и валидирован тест V-IRAP, который представляет собой модифицированную версию Процедуры Имплицитной Реляционной Оценки, или IRAP. Данный тест, в частности, использовался в ряде исследовательских работ для выявления неосознаваемых причин недоверия к вакцинам, основанных на когнитивных ошибках восприятия, и показал высокую надежность в дифференциации позитивных и негативных представлений участников о вакцинации (Simione et al., 2022). Для изучения аффективных состояний был разработан «Тест имплицитного позитивного

и негативного аффекта» (The Implicit Positive and Negative Affect Test – IPANAT) (Quirin et al., 2009), позволяющий диагностировать психическое состояние человека и степень его аффективности. При этом тест позволяет выделять два независимых друг от друга имплицитных компонента – позитивный и негативный (Осин, 2012).

Таким образом, можно сказать, что для выявления имплицитных компонентов отношения к проблеме или событию существует и применяется инструментарий в виде разработанных модифицируемых тестов имплицитных ассоциаций, процедура применения которых постоянно совершенствуется.

За время своего существования IAT показал высокую эффективность при соблюдении методологических и процедурных рекомендаций по его применению (Greenwald et al., 2009; Meissner et al., 2019; Howell et al., 2022).

Обзор 122 исследований (184 независимые выборки, 14900 испытуемых), проведенный Greenwald и соавт. (Greenwald et al., 2009), показал, что среднее значение для предсказания поведения, суждений и физиологических показателей с помощью показателей теста имплицитных ассоциаций (IAT) составляет  $r = 0.274$ . Параллельные эксплицитные (т.е. самоотчетные) показатели, доступные в 156 из этих выборок (13068 испытуемых), также были эффективны в предсказании (средний  $r = 0.361$ ), но с гораздо большей вариативностью величины эффекта. Как IAT, так и самоотчеты продемонстрировали инкрементную валидность, когда каждая мера предсказывала дисперсию критерия сверх той, что предсказывалась другой.

Валидность IAT была показана и в ряде сравнительных исследований. Например, в исследовании Cai и соавт. (Cai et al., 2021) решалась задача валидировать связь неявной самооценки, измеренной с помощью теста имплицитных ассоциаций (IAT), и автоматической самооценки, измеряемой при помощи модифицированной парадигмы Go/NoGo, предполагающей запись времени реакции (RT); стимулы, используемые в обеих экспериментальных сессиях, были одинаковы. Анализ результатов позволил авторам подтвердить, что IAT действительно демонстрирует результаты автоматического процесса самооценки. В работе Portengen и соавт. (Portengen et al., 2022) показано совпадение в выявляемых нейрофизиологических показателях для результатов, полученных в ходе процедуры IAT и задачи формирования впечатления (IFT). Несмотря на то, что в некоторых исследованиях отмечается возможность сознательной корректировки неявной предвзятости под влиянием перцептивного внимания и мониторинга конфликтов и ошибок (van Nunspeet et al., 2014), результаты, получаемые при помощи IAT, оказываются устойчивы к замаскированному праймингу (Pesciarelli et al., 2021), что,

с одной стороны, позволяет говорить о стабильности и воспроизводимости сформированной неявной оценки и подтверждает валидность применения теста имплицитных ассоциаций при ее оценке, с другой, требует от исследователя соблюдения методологических и процедурных рекомендаций.

Для поиска нейрофизиологических коррелятов неявной предвзятости в большинстве работ используют метод регистрации потенциалов, связанных с событием (ВП). Привязка компонента ВП ко временной шкале от начала появления стимула позволяет исследовать динамику когнитивного процесса и отличительные ВП-характеристики неявной предвзятости, а широкая изученность характеристик и компонентов ВП позволяет с высоким уровнем достоверности атрибутировать имплицитные процессы с описанными нейрофизиологическими механизмами. Кроме того, в ходе анализа литературных источников было выделено два направления исследований. Первое связано с изучением мозговых механизмов обработки информации от имплицитных стимулов IAT для задач, связанных с когнитивными особенностями и физиологическими состояниями. Второе направлено на поиск нейрофизиологических коррелятов имплицитных процессов в аспекте исследования социальных явлений человеческого общества. В связи с этим для реализации целей настоящего обзора представляют интерес, с одной стороны, работы, описывающие нейрофизиологические корреляты имплицитных процессов с точки зрения их автоматического характера, с другой стороны, их преломление в выявлении мозговых механизмов обработки информации, отражающей имплицитное отношение к социально-культурным феноменам.

#### *Нейрофизиологические корреляты имплицитных процессов*

В исследовании G.F. Nealy и соавт. (Nealy et al., 2015) на примере ассоциации между целевой категорией «Я» и двумя категориями оценочных атрибутов («Природа/Природный», «Созданный/Рукотворный») продемонстрировано, как показатели ВП, лежащие в основе как конгруэнтных, так и неконгруэнтных типов реакций, могут свидетельствовать о наличии нейронных механизмов в условиях быстрых ассоциаций, вызывающих эффект IAT. Результаты исследования выявили статистически значимые связи между D-показателем IAT (D рассчитывается как разница среднего значения времени реакции конгруэнтных ответов и среднего значения неконгруэнтных, деленная на стандартное отклонение для всех времен реакции) и ранними временными окнами ВП. LORETA-анализ позволил выявить ряд областей мозга, управляющих этими отношениями ВП-IAT, в частности

лево-височную, островковую, поясную, медиальную лобную и теменную кору в височных областях, соответствующих активности, связанной с компонентами ВП N200 и P300.

В исследовании связи индивидуальных различий в болевой чувствительности с неявным негативным предубеждением в отношении боли (Qian et al., 2020) оценивались поведенческие данные IAT (RT – время реакции, ACC – точность, D – характеристика неявной предвзятости) и ВП-данные (амплитуды N100, N200, P200 и P300). Исследование проводилось в группах участников с низкой и высокой болевой чувствительностью. Авторы предположили, что имплицитное негативное отношение к боли будет выявляться как на основе поведенческих показателей (показатели времени реакции), так и в компонентах ВП. В результате участники группы с высокой чувствительностью продемонстрировали большие показатели D-критерия, что свидетельствует о наличии неявной предвзятости. Значение амплитуды компонента P300 также отличалось в большую сторону по сравнению с участниками группы с меньшей чувствительностью. При этом в ранних компонентах ВП (N100, P200 и N200) различий выявлено не было, что позволило авторам сделать вывод об обусловленности межгрупповых различий в имплицитном негативном предубеждении скорее поздними процессами когнитивного контроля, нежели ранними перцептивными процессами.

В работе J. Hilgard и соавт. (Hilgard et al., 2015) проводилось исследование нейрофизиологических характеристик проактивного и реактивного контроля при конгруэнтных/неконгруэнтных ответах во время выполнения IAT. Авторами была обнаружена более выраженная для неконгруэнтных, чем конгруэнтных испытаний медиально-лобная негативность (MFN) на латентности 250–450 мс. В то же время значения амплитуды снижались для испытаний с переключением в сравнении с испытаниями без переключения, что позволило авторам предположить наличие отдельных процессов контроля в механизмах обработки конгруэнтных и неконгруэнтных стимулов. Кроме того, авторы отмечают влияние степени выраженности предвзятости: более выраженная поведенческая предвзятость характеризуется более высокими значениями амплитуд компонентов ВП, связываемыми как с переключением во время ответов, так и с конгруэнтностью самих ответов.

Вызывает интерес исследование, посвященное оценке имплицитного компонента пищевого поведения и его отражению в показателях компонентов ВП (Lahtinen et al., 2019). Авторы оценивали связь эффекта IAT с амплитудами компонентов ВП, описывающих оценку стимулов, мотивационное внимание и мониторинг конфликтов. Кроме того, исследовалось влияние метаболического состояния

(то есть голодания или сытости) на нейронные ответы о неявных предпочтениях в отношении еды. Чтобы оценить влияние метаболического статуса на результаты IAT, серии экспериментов проводились в двух условиях: натощак и после приема пищи участниками. Полученные результаты продемонстрировали, что на групповом уровне компоненты N400, P300 и поздний позитивный компонент (LPP) вызываются в задаче IAT, связанной с пищевыми продуктами. Кроме того, на индивидуальном уровне потенциалы N400 и LPP коррелировали с D-показателем IAT. Авторы предположили, что эффект IAT отражает нейронные отклонения неявного пищевого поведения. При этом имплицитное предпочтение пищи, измеренное с помощью D-показателя IAT, оказалось более выраженным в состоянии натощак, чем после приема пищи. Таким образом, состояние голода увеличивает неявную гедонистическую мотивацию к еде.

Работа Wang и соавт. (Wang et al., 2017) посвящена изучению неявной регуляции эмоций во время обработки изображений лиц с помощью задачи PI («Задача прайминга-идентификации»), по механизму являющейся аналогом теста имплицитных ассоциаций. Авторы продемонстрировали, что неявная регуляция эмоций связана с более выраженными значениями амплитуды компонента N170 без изменения амплитуд раннего заднего негативного компонента (EPN) и LPP, что может указывать на автоматизм протекания процесса неявной регуляции эмоций на ранней стадии восприятия, а не на поздней стадии избирательного внимания при обработке эмоций.

В исследовании Cui и соавт. (Cui et al., 2019) изучалась роль дозы никотина в мотивационных реакциях участников на сигналы о курении. В результате анализа было показано, что в группе курильщиков изображения, связанные с сигаретой, вызывают более высокие значения амплитуды раннего заднего негативного компонента ВП (на латентности 208–288 мс) (EPN), чем нейтральные изображения. Аналогичным образом условие Give-DENIC (безникотиновая сигарета) вызывало отклонение амплитуды раннего заднего негативного компонента ВП (EPN) в сравнении с условием Give-NIC (никотиновая сигарета), что говорит о повышении мотивационной составляющей неявной предвзятости в условиях никотинового голода.

В схожем исследовании Tschuemperlin и соавт. (Tschuemperlin et al., 2020) изучались нейрофизиологические корреляты неявных ассоциаций с алкоголем у недавно воздержавшихся от употребления алкоголя пациентов с расстройствами употребления алкоголя (AUD) в ходе выполнения теста быстрых имплицитных ассоциаций (BIAT). Хотя никаких поведенческих групповых различий

обнаружено не было, потенциалы, связанные с событием, пациентов и контрольной группы характеризовались различными микросостояниями, начиная с 320 мс. Для пациентов с AUD характер, время и локализация источника эффектов предполагали наличие больших усилий в отношении семантической и саморелевантной интеграции на латентности около 400 мс во время неконгруэнтных испытаний и ослабление эмоциональной обработки в течение позднего периода положительного потенциала.

Таким образом, представленные исследования свидетельствуют об автоматизме протекания имплицитных процессов, в частности неявной предвзятости и его детерминированности, в т.ч. физиологическими состояниями. При этом разница в компонентах ВП, отражающая реализацию имплицитных процессов, связана в основном

с поздними компонентами ВП (см. суммирование эффектов в таблице 1). Выявленные различия в амплитудах ранних компонентов ВП, по-видимому, отражают регуляцию эмоционально-мотивационных составляющих имплицитных процессов.

Подтвердив, что имплицитные процессы могут быть описаны через компоненты ВП, мы переходим к обзору исследований, посвященных выявлению мозговых механизмов обработки информации, отражающей имплицитное отношение к социально-культурным феноменам.

### *Опыт исследования вызванных потенциалов имплицитного компонента социальных явлений*

Большая часть исследований в обозначенной тематике связана с изучением многочисленных социальных феноменов когнитивных искажений.

**Таблица 1.** Компоненты вызванного потенциала (ВП), связанные с имплицитными процессами в представленных работах

**Table 1.** ERP components related to implicit processes in the presented works

| Компонент ВП | Эффект ВП                                          | Цель исследования                                   | Пример исследования           |
|--------------|----------------------------------------------------|-----------------------------------------------------|-------------------------------|
| N100         | Нисходящий контроль                                | Суицидальное поведение                              | Camsari et al., 2023          |
| N170         | Распознавание лиц                                  | Неявная предвзятость при восприятии лиц другой расы | Anzures et al., 2021          |
|              |                                                    | Регуляции эмоций во время обработки изображений лиц | Wang et al., (2017)           |
| N200         | Когнитивный контроль                               | Нейронные механизмы IAT                             | Healy et al., 2015            |
|              |                                                    | Интернет-зависимость                                | Chen et al., 2018             |
| P200         | Избирательное внимание                             | Моральная оценка                                    | Tao et al., 2022              |
|              |                                                    | Суицидальное поведение                              | Camsari et al., 2023          |
|              |                                                    | Восприятие лиц другой расы                          | Anzures et al., 2021          |
| EPN          | Избирательное внимание на эмоц. стимулах           | Мотивационные реакции на стимулы                    | Cui et al., 2019              |
| P300         | Категоризация стимулов                             | Нейронные механизмы IAT                             | Healy et al., 2015            |
|              |                                                    | Предубеждения в отношении боли                      | Qian et al., 2020             |
|              |                                                    | Пищевое поведение                                   | Lahtinen et al., 2019         |
|              |                                                    | Моральная оценка                                    | Tao et al., 2022              |
|              |                                                    | Интернет-зависимость                                | Chen et al., 2018             |
|              |                                                    | Гендерные ожидания                                  | Portengen et al., 2022        |
| N400         | Семантические несоответствия                       | Сексуальные предпочтения                            | Williams, Themanson, 2011     |
|              |                                                    | Пищевое поведение                                   | Lahtinen et al., 2019         |
| LPP          | Эмоциональная значимость и семантическая обработка | Сексуальная ориентация                              | Williams, Themanson, 2011     |
|              |                                                    | Моральная оценка                                    | Tao et al., 2022              |
|              |                                                    | Пищевое поведение                                   | Lahtinen et al., 2019         |
|              |                                                    | Суицидальное поведение                              | Camsari et al., 2023          |
|              |                                                    | Самооценка при дисфории                             | Lou et al., 2021              |
|              |                                                    | Аффективные реакции                                 | Bosshard et al., 2016         |
| MFN          | Проактивный контроль ошибок/конфликта              | Проактивный контроль IAT                            | Hilgard et al., 2015          |
|              |                                                    | Неявная предвзятость к жителям других регионов      | Wagner-Altendorf et al., 2023 |

Авторы фокусируются на выделении неявной предвзятости человека в отношении социальных явлений, например, таких, как расовые и внутригрупповые предубеждения, суицидальное поведение, моральная оценка, гендерные ожидания, отношение к брендам и многое другое. В этом разделе мы ставим перед собой задачу представить разнообразие социальных явлений, на примере которых изучается нейрональный механизм неявной предвзятости.

В исследовании нейрофизиологических процессов, которые отражают групповую неявную предвзятость, J.K. Williams и J.R. Themanson (Williams, Themanson, 2011) на материале отношения участников к сексуальным предпочтениям других людей обнаружили, что ранние компоненты (N100, P200) не отражают различий в зависимости от условий IAT, следовательно, более ранние процессы внимания и восприятия не связаны с поведенческими эффектами теста. При этом более поздние компоненты N400 и LPP модулируются условиями стимула (конгруэнтные/неконгруэнтные).

В исследовании D. Тао и соавт. (Tao et al., 2022) было проведено сравнение временной динамики между явной и неявной моральной оценкой сценариев причинения вреда/заботы, с применением регистрации потенциалов, связанных с событием. Было показано, что как эксплицитная, так и имплицитная моральные оценки включают взаимодействие между эмоциональными процессами и моральным познанием и представляют собой два относительно независимых процесса, что проявилось в различиях амплитуд компонентов ВП. При этом средние и поздние стадии обработки интеграции эмоционального возбуждения и моральной оценки, которые включают тщательную обработку и когнитивный контроль, отразились в следующих компонентах ВП: в лобном P200, теменном P300, теменном LPP и FSW (лобная медленная волна).

T.A. Wagner-Altendorf и соавт. (Wagner-Altendorf et al., 2023) исследовали неявный фаворитизм внутренней группы, к которой принадлежит человек, в сравнении с внешней группой, к которой человек не принадлежит. Для этого оценивалась неявная предвзятость у западных и восточных немцев в отношении городов Восточной и Западной Германии. Полученные результаты показали, что участники исследования демонстрировали сокращение времени реакции при сочетании названий европейских городов и негативных оценочных слов по сравнению с позитивными, в обеих группах. Кроме того, было обнаружено увеличение медиально-центральной негативности (MFN) на латентности 400–800 мс после стимула в неконгруэнтных условиях пары «восточногерманский город/позитивное слово» (у западных немцев) и пары «западногерманский город/позитивное слово» (у восточных немцев). Таким образом, было продемонстрировано существование неявных внутригрупповых

предубеждений у западно- и восточногерманских участников.

D. D. Camsari и соавт. (Camsari et al., 2023) представили работу по поиску ВП-маркеров суицидального поведения подростков, в которой авторы использовали модифицированный IAT для выявления неявных ассоциаций смерти/самоубийства (DS-IAT). Авторы предположили, что подростки с суицидальными мыслями и поведением будут иметь более выраженные ассоциации между понятиями «смерть» и «Я», по сравнению с контрольной группой, при этом тяжесть явных суицидальных мыслей и поведения у пациентов будет коррелировать с силой ассоциаций между «смертью» и «я». В результате отмечается, что как в экспериментальной, так и в контрольной группах у подростков были выявлены большие ассоциации с сочетаниями «жизнь-я», чем со «смерть-я». Однако подростки из экспериментальной группы имели более выраженные имплицитные ассоциации с сочетанием «смерть-я». Корреляционный анализ также выявил наличие значимых корреляций показателя D и компонента N100 в левой теменно-затылочной коре. Кроме того, авторы обнаружили групповые различия по компонентам N100 в области теменной и латеральной затылочной коры, компоненту P200 в левой дорсолатеральной области префронтальной коры и компоненту LPP в лобно-центральной и височно-теменной областях.

Исследование (Lou et al., 2021) было направлено на изучение реакций мозга людей с дисфорией во время теста имплицитных ассоциаций на самооценку. Участникам было предложено пройти два блока заданий в IAT — на самопозитивность и самонегативность. Результаты показали, что различия в реакциях мозга наблюдались между дисфорической и контрольной группами в позднем позитивном компоненте (LPP) в пределах латентности 400–1000 мс. В отличие от контрольной группы, в группе с дисфорией стимулы, отображающие самонегативность, модулировали более высокие значения амплитуды LPP по сравнению со стимулами самопозитивности. Кроме того, полученные результаты показали, что в экспериментальной группе участники в большей степени демонстрировали неявную самонегативную ассоциацию. Таким образом, авторы пришли к выводу, что реакции ВП, а не поведенческие характеристики, больше связаны с групповыми различиями. А высокие амплитуды LPP в ассоциации самонегативности и ее корреляция с более высокими баллами по шкале депрессивных симптомов могут быть интерпретированы как нейронный индекс пониженной имплицитной самооценки.

В исследовании G. Anzures и соавт. (Anzures, 2021) изучались особенности потенциалов, связанных с событием при восприятии лиц, а также влияние неявной предвзятости на восприятие лиц другой расы. Авторы предположили, что умение распознавать лица, неявная расовая предвзятость

или взаимодействие между ними могут объяснить значительную вариабельность ранних реакций ВП при восприятии лиц другой расы (на примере компонентов P100, N170 и P200). Были обнаружены более высокие значения амплитуд данных компонентов ВП и увеличение латентности компонента N170 при распознавании лиц другой расы по сравнению с лицами своей расы.

В исследовании гендерных различий (Portengen et al., 2022) была обнаружена значительная разница в компоненте P3 между стимулами, которые подтверждали или нарушали гендерные ожидания.

В работе L. Chen и соавт. (Chen et al., 2018) изучались характеристики ВП неявной когнитивной предвзятости при интернет-зависимости. Были сформированы две группы: экспериментальная и контрольная, сопоставленные по полу и возрасту. Регистрировались поведенческие данные IAT и данные ВП (компоненты P100, N200, P300 и N400). Исследование выявило значительную разницу во времени реакции между экспериментальной и контрольной группами, что свидетельствует о более выраженных для участников с интернет-зависимостью положительных неявных ассоциациях с сигналами, связанными с Интернетом. Такие сигналы также характеризовались более высокими амплитудами N200 и P300 в затылочных долях у людей, страдающих интернет-зависимостью.

В исследовании Bosshard и соавт. (Bosshard et al., 2016) оценивалась связь аффективных реакций участников по отношению к брендам-производителям с компонентами ВП. Поведенческие показатели участников исследования, полученные с помощью IAT, подтвердили предварительно собранные самоотчетные данные отношений (по шкале Ликерта). При этом понравившиеся бренды вызывали у участников значительно большие значения амплитуд LPP (в сравнении с нелюбимыми брендами) в правых теменных областях коры, начиная примерно через 800 мс после начала стимула (достигая статистической значимости примерно через 1000 мс) и продолжаясь до конца эпохи записи (2000 мс).

Таким образом, использование потенциалов, связанных с событием, для изучения социально обусловленных имплицитных компонентов когнитивных искажений находится в фокусе активного внимания исследователей и применимо к широкому кругу явлений.

## ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Представленные в обзоре исследования демонстрируют высокий потенциал применения регистрации ВП для изучения нейрофизиологических коррелятов имплицитных компонентов когнитивных процессов, в том числе когнитивных искажений.

В приведенных работах можно выделить воспроизводимые ВП-корреляты имплицитных компонентов когнитивной предвзятости. Например, наиболее системно воспроизводится поздний позитивный потенциал (LPP), потенциал P300 и P200, в меньшей степени — N400, медиально-фронтальная негативность (MFN), N200, N170, N100. Таблица 1 суммирует представленные в обзоре компоненты ВП, соотнося их с эффектами и тематическим направлением исследуемых социальных явлений.

Данный факт, безусловно, свидетельствует, что в основе имплицитных компонентов IAT находятся определенные нейрофизиологические механизмы для обработки поступающей информации. Если обобщить представленные данные, можно выделить следующие нейрофизиологические механизмы, принимающие участие в реализации неявных искажений:

P200, амплитуда которого модулируется в случаях возникновения различий в скрытой памяти или несоответствия валентности информации во время перцептивной обработки;

N2, показывающий высокую амплитуду во время неконгруэнтных заданий IAT из-за наличия конфликта в ответах;

P300, отражающий эффект конгруэнтности на оценку и категоризацию стимулов, связанных с принятием решений, которые контролируют выбор правильной двигательной реакции;

N400, амплитуда которого оказывается чувствительна к семантическим несоответствиям при выполнении IAT;

MFN, демонстрирующий высокую амплитуду при выполнении конгруэнтных заданий, что отражает нервные процессы, поддерживающие проактивный контроль;

LPP, амплитуда которого увеличивается при возникновении эмоционального эффекта конгруэнтности стимулов и отношения;

Таким образом, использование регистрации потенциалов, связанных с событием, в исследовании имплицитных процессов находит широкое применение в достаточном количестве исследований, что позволяет выявлять последовательные нейрофизиологические механизмы обработки поступающей информации. Кроме того, анализ представленных работ показал, что, несмотря на разные цели исследования, авторы выделяют компоненты ВП, которые чаще всего связаны с обработкой имплицитной информации. По-видимому, это отражает включение в ее обработку мозговых механизмов, реагирующих на содержащиеся в IAT триггеры, характерные для данных исследований, например, LPP, связанный с эмоциональным компонентом, или P300, отражающий момент категоризации информации и принятие решения.

Однако для расширения представлений об имплицитных и эксплицитных процессах и их

нейрофизиологических основаниях необходимо привлекать новые подходы к анализу нейрофизиологических данных.

Так, представляет интерес подход В. Schiller и соавт. (Schiller et al., 2016), в исследовании которых был применен метод анализа микросостояний потенциалов, связанных с событием. Авторы пытаются ответить на вопрос, чем обусловлен эффект IAT, иными словами, почему происходит увеличение времени реакции при связывании неконгруэнтных стимулов. Для решения данного вопроса был использован метод анализа микросостояний ВП (Schiller et al., 2023), зарегистрированных в процессе выполнения IAT двумя разнородными группами участников. Авторы фиксировали поведенческие показатели теста (D-показатель) и производили усреднение полученных ВП. На основе силуэтного графика проводилась кластеризация ВП и выделение микросостояний, что позволило идентифицировать протекающие психические процессы во временном диапазоне от 0 до 1000 мс. Проведенный sLORETA-анализ позволил локализовать области мозга, лежащие в основе каждого психического процесса. Сравнение микросостояний конгруэнтных и неконгруэнтных ВП в двух группах участников позволило авторам прийти к выводу, что различия во времени реакции и имплицитные предубеждения обусловлены количественными (временем протекания), а не качественными различиями в основных психических процессах. Это позволило исключить предположение о том, что большая длительность времени реакции при имплицитных предубеждениях обусловлена включением в обработку информации еще одного, дополнительного мыслительного процесса.

Кроме того, для изучения имплицитных процессов все чаще применяют анализ ЭЭГ-сигнала различными методами машинного обучения. Например, в исследовании С. Chaïyapan и др. (2021) целью являлось выявление взаимосвязи между имплицитными событиями обучения и характеристиками их ЭЭГ-сигналов путем поиска отличительных признаков, связанных с производительностью участников, и проверки эффективности различных классификаторов для нахождения этих различий. В своей работе авторы извлекали из сигнала ЭЭГ, зарегистрированного во время выполнения когнитивных задач, данные многомасштабной энтропии (MSE) как признака, позволяющего классифицировать функциональные состояния нервной системы, связанные с событиями имплицитного обучения. Кроме того, для оптимизации процесса классификации функциональных состояний был использован метод Искусственной Пчелиной Семьи (ABC), который основан на имитации интеллектуальных действий медоносных пчел во время поиска источника пищи. Результаты показали, что система

может корректно идентифицировать различия между показателями участников, используя данные MSE и метод ABC, с вероятностью 95%.

В ряде нейромаркетинговых исследований с помощью алгоритмов машинного обучения авторы ставили цель предсказывать потребительские предпочтения покупателей. В этом аспекте наиболее эффективными алгоритмами оказались глубокие нейронные сети (DNN), точность прогнозирования потребительских предпочтений которых достигала 94%. Кроме того, точность прогнозирования увеличивалась при использовании данных ЭЭГ в сочетании с другими физиологическими методами, например, айтрекингом и кожно-гальванической реакцией (Aldayel, 2021).

Таким образом, можно сделать вывод, что для получения исчерпывающей информации о нейрофизиологических механизмах обработки стимулов IAT, выявляющих неявную предвзятость, возникает необходимость использования комбинации различных подходов к анализу ЭЭГ-сигнала, что позволит не только выделить нейромаркеры имплицитных процессов, но и расширить представления об их нейрофизиологической основе.

## ЗАКЛЮЧЕНИЕ

Проведенный обзор исследований взаимосвязи потенциалов, связанных с событием, с имплицитными процессами показал, что данная тема характеризуется нарастающим интересом со стороны исследователей в области психологии, нейрофизиологии, социологии и др., особенно в последние годы. Литературные источники, посвященные данной тематике, отражают широкое тематическое разнообразие материала, на примере которого исследуются нейромаркеры неявной предвзятости. Значительную группу составляют исследования социально-психологических и культурных феноменов человеческого общества.

Однако, несмотря на широкое тематическое разнообразие, результаты исследований демонстрируют стабильное воспроизведение отдельных компонентов связанных с событием потенциалов, что, безусловно, свидетельствует о наличии базового нейрофизиологического механизма обработки подобного рода информации и автоматизме его протекания. При этом разница в компонентах вызванных потенциалов, отражающая реализацию имплицитных процессов, связана в основном с поздними компонентами ВП.

Отметим также, что существенно расширить представления о нейрофизиологических основаниях протекания имплицитных процессов позволит применение комбинации различных методов к анализу ЭЭГ-активности, в том числе в сочетании с другими физиологическими методами.

## ВКЛАД АВТОРОВ

М.В. Яценко, Е.Д. Артеменко — концепция и руководство работой; И.В. Брак — обсуждение результатов; М.В. Яценко — написание текста; Е.Д. Артеменко — редактирование текста статьи.

## ФИНАНСИРОВАНИЕ

Исследование выполнено за счет гранта Российского научного фонда № 23-78-10035, <https://rscf.ru/project/23-78-10035/>.

## КОНФЛИКТ ИНТЕРЕСОВ

Авторы заявляют об отсутствии конфликта интересов.

## СОБЛЮДЕНИЕ ЭТИЧЕСКИХ СТАНДАРТОВ

Настоящая статья не содержит описания каких-либо проведенных авторами исследований с участием людей или животных в качестве объектов.

## УКАЗАНИЕ НА ДОСТУПНОСТЬ ПЕРВИЧНЫХ ДАННЫХ

По запросу авторы готовы предоставить аннотированный банк статей, вошедших в настоящий литературный обзор.

## СПИСОК ЛИТЕРАТУРЫ

- Боброва Л.А. Когнитивные искажения. Социальные и гуманитарные науки. Отечественная и зарубежная литература. Серия 3, Философия: Реферативный журнал. 2021. (2): 69–79.
- Канеман Д. Думай медленно... решай быстро. М.: АСТ, 2016. 653 с.
- Осин Е.Н. Измерение позитивных и негативных эмоций: разработка русскоязычного аналога методики PANAS. Психология. Журнал Высшей школы экономики. 2012. 9 (4): 91–110.
- Aldayel M., Ykhlef M., Al-Naffjan A. Recognition of Consumer Preference by Analysis and Classification EEG Signals. *Front. Hum. Neurosci.* 2021. 14: 604639. <https://doi.org/10.3389/fnhum.2020.604639>
- Anzures G., Mildort M. Do perceptual expertise and implicit racial bias predict early face-sensitive ERP responses? *Brain Cogn.* 2021. 147: 105671. <https://doi.org/10.1016/j.bandc.2020.105671>
- Barsade S.G., Ramarajan L., Westen D. Implicit affect in organizations. *Research in Organizational Behavior.* 2009. 29: 135–162. <https://doi.org/10.1016/j.riob.2009.06.008>
- Bhat D., Kollu T., Ricci J.A., Patel A. What's in a Name? Implicit Bias Affects Patient Perception of Surgeon Skill. *Plast. Reconstr. Surg.* 2021. 147 (6): 948–956. <https://doi.org/10.1097/PRS.00000000000008171>
- Bosshard S.S., Bourke J.D., Kunaharan S., Koller M., Walla P., Heinonen J. Established liked versus disliked brands: Brain activity, implicit associations and explicit responses. *Cogent Psychology.* 2016. 3:1. <https://doi.org/10.1080/23311908.2016.1176691>
- Camsari D.D., Lewis C.P., Sonmez A.I., Ozger C., Fatih P., Yuruk D., Shekunov J., Vande Voort J.L., Croarkin P.E. Event-Related Potential Markers of Suicidality in Adolescents. *Int. J. Neuropsychopharmacol.* 2023. 26 (8): 566–575. <https://doi.org/10.1093/ijnp/pyad039>
- Cai H., Wu L. The self-esteem implicit association test is valid: Evidence from brain activity. *Psych J.* 2021 Jun;10(3):465–477. doi: 10.1002/pchj.422. Epub 2021 Jan 28. PMID: 33511787.
- Chaiyanan C., Iramina K., Kaewkamnerdpong B. Investigation on Identifying Implicit Learning Event from EEG Signal Using Multiscale Entropy and Artificial Bee Colony. *Entropy.* 2021. 23 (5): 617. <https://doi.org/10.3390/e23050617>
- Chen L., Zhou H., Gu Y., Wang S., Wang J., Tian L., Zhu H., Zhou Z. The Neural Correlates of Implicit Cognitive Bias Toward Internet-Related Cues in Internet Addiction: An ERP Study. *Front. Psychiatry.* 2018. 9: 421. <https://doi.org/10.3389/fpsy.2018.00421>
- Cui Y., Engelmann J.M., Gilbert D.G., Waters A.J., Cinciripini P.M., Robinson J.D. The impact of nicotine dose and instructed dose on smokers' implicit attitudes to smoking cues: An ERP study. *Psychol Addict Behav.* 2019. 33(8): 710–720. <https://doi.org/10.1037/adb0000523>
- Greenwald A.G., McGhee D.E., Schwartz J.L. Measuring individual differences in implicit cognition: the implicit association test. *J. Pers. Soc. Psychol.* 1998. 74 (6): 1464–80. <https://doi.org/10.1037//0022-3514.74.6.1464>
- Greenwald A.G., Poehlman T.A., Uhlmann E.L., Banaji M.R. Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *J Pers Soc Psychol.* 2009. 97 (1): 17–41. <https://doi.org/10.1037/a0015575>
- Healy G.F., Boran L., Smeaton A.F. Neural Patterns of the Implicit Association Test. *Front. Hum. Neurosci.* 2015. 9: 605. <https://doi.org/10.3389/fnhum.2015.00605>
- Hilgard J., Bartholow B.D., Dickter C.L., Blanton H. Characterizing switching and congruency effects in the Implicit Association Test as reactive and proactive cognitive control. *Social Cognitive and Affective Neuroscience.* 2015. 10 (3): 381–388. <https://doi.org/10.1093/scan/nsu060>

- Howell J.L., Gasser M.L., Kaysen D., Lindgren K.P. Understanding parental vaccine refusal: Implicit and explicit associations about vaccines as potential building blocks of vaccine beliefs and behavior. *Social Science & Medicine*. 2022. 310: 115275.  
<https://doi.org/10.1016/j.socscimed.2022.115275>
- Lahtinen A., Juvonen K., Lapveteläinen A., Kolehmainen M., Lindholm M., Tanila H., Kantanen T., Sinikallio S., Karhunen L., Närväinen J. Metabolic state as a modulator of neural event-related potentials for food stimuli in an implicit association test. *Physiol. Behav.* 2019. 209: 112589.  
<https://doi.org/10.1016/j.physbeh.2019.112589>
- Le Bouc R., Pessiglione M. A neuro-computational account of procrastination behavior. *Nat. Commun.* 2022. 13 (1): 5639.  
<https://doi.org/10.1038/s41467-022-33119-w>
- Lou Y., Lei Y., Astikainen P., Peng W., Otieno S., Leppänen P.H.T. Brain responses of dysphoric and control participants during a self-esteem implicit association test. *Psychophysiology*. 2021. 58 (4): e13768.  
<https://doi.org/10.1111/psyp.13768>
- Meissner F., Grigutsch L.A., Koranyi N., Müller F., Rothermund K. Predicting Behavior With Implicit Measures: Disillusioning Findings, Reasonable Explanations, and Sophisticated Solutions. *Front. Psychol.* 2019. 10: 2483.  
<https://doi.org/10.3389/fpsyg.2019.02483>
- Nosek B.A., Hawkins C.B., Frazier R.S. Implicit social cognition: from measures to mechanisms. *Trends Cogn Sci.* 2011. 15 (4): 152–159.  
<https://doi.org/10.1016/j.tics.2011.01.005>
- Pesciarelli F., Leo I., Serafini L. Electrophysiological correlates of unconscious processes of race. *Sci Rep.* 2021. 11: 11646.  
<https://doi.org/10.1038/s41598-021-91133-2>
- Peters U. What is the function of confirmation bias? *Erkenntnis*, 2022. 87 (3): 1351–1376.  
[doi:https://doi.org/10.1007/s10670-020-00252-1](https://doi.org/10.1007/s10670-020-00252-1)
- Qian Y., Xiaoyun L., Weiwei P. Individual Variation in Pain Sensitivity and Implicit Negative Bias Toward Pain. *Psychosomatic Medicine*. 2020. 82 (8): 796–804.  
<https://doi.org/10.1097/PSY.0000000000000844>
- Portengen C.M., Huffmeijer R., van Baar A.L., Endendijk J.J. Measuring the neural correlates of the violation of social expectations: A comparison of two experimental tasks. *Social Neuroscience*. 2022. 17 (1): 58–72.  
<https://doi.org/10.1080/17470919.2022.2032327>
- Quirin M., Kazén M., Kuhl J. When nonsense sounds happy or helpless: The Implicit Positive and Negative Affect Test (IPANAT). *Journal of Personality and Social Psychology*. 2009. 97: 500–516.  
<https://doi.org/10.1037/a0016063>
- Schiller B., Gianotti L.R., Baumgartner T., Nash K., Koenig T., Knoch D. Clocking the social mind by identifying mental processes in the IAT with electrical neuroimaging. *Proc. Natl. Acad. Sci. USA*. 2016. 113 (10): 2786–2791.  
<https://doi.org/10.1073/pnas.1515828113>
- Schiller B., Sperl M.F.J., Kleinert T., Nash K., Gianotti L.R.R. EEG Microstates in Social and Affective Neuroscience. *Brain Topogr.* 2023.  
<https://doi.org/10.1007/s10548-023-00987-4>
- Shah H.S., Bohlen J. Implicit Bias. 2023 Mar 4. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing, 2024.
- Simione L., Vagni M., Maiorano T., Giostra V., Pajardi D. How Implicit Attitudes toward Vaccination Affect Vaccine Hesitancy and Behaviour: Developing and Validating the V-IRAP. *Int. J. Environ. Res. Public Health*. 2022. 19: 4205.  
<https://doi.org/10.3390/ijerph19074205>
- Tao D., Leng Y., Peng S., Xu J., Ge S., Deng H. Temporal dynamics of explicit and implicit moral evaluations. *International Journal of Psychophysiology*. 2022. 172: 1–9.  
<https://doi.org/10.1016/j.ijpsycho.2021.12.006>
- Tschuempferlin R.M., Batschelet H.M., Moggi F., Koenig T., Roesner S., Keller A., Pfeifer P., Soravia L.M., Stein M. The Neurophysiology of Implicit Alcohol Associations in Recently Abstinent Patients With Alcohol Use Disorder: An Event-Related Potential Study Considering Gender Effects. *Alcohol Clin Exp Res*. 2020. 44(10): 2031–2044.  
<https://doi.org/10.1111/acer.14444>
- van Nunspeet F., Ellemers N., Derks B., Nieuwenhuis S. Moral concerns increase attention and response monitoring during IAT performance: ERP evidence. *Social Cognitive and Affective Neuroscience*. 2014. 9 (2): 141–149.
- Wagner-Altendorf T.A., van der Lugt A.H., Kroeber A., Cirkel A., Heldmann M., Münte T.F. Differences in Implicit Attitudes in West and East Germans as Measured by the Go/NoGo Association Task and Event-related EEG Potentials. *Cognitive and Behavioral Neurology*. 2023. 36 (3): 145–158.  
<https://doi.org/10.1097/WNN.0000000000000338>
- Wang Y., Li X. Temporal course of implicit emotion regulation during a Priming-Identify task: an ERP study. *Sci Rep.* 2017. 2(7): 41941.  
<https://doi.org/10.1038/srep41941>
- Williams J.K., Themanson J.R. Neural correlates of the implicit association test: evidence for semantic and emotional processing. *Social Cognitive and Affective Neuroscience*. 2011. 6 (4): 468–476.  
<https://doi.org/10.1093/scan/nsq065>

## EVENT-RELATED POTENTIALS IN STUDIES OF THE IMPLICIT COMPONENT OF COGNITIVE BIASES

M. V. Yatsenko<sup>a, b</sup>, I. V. Brak<sup>c</sup>, E. D. Artemenko<sup>a, #</sup>

<sup>a</sup>*Social & Cognitive Informatics lab, HSE University, Saint Petersburg, Russia*

<sup>b</sup>*Altai State University, Barnaul, Russia*

<sup>c</sup>*Novosibirsk State University, Novosibirsk, Russia*

<sup>#</sup>*e-mail: edartemenko@hse.ru*

The review analyzes neurophysiological correlates of the implicit component of cognitive bias in the processes of perception and decision-making. The review identifies the leading methodological pipeline for analysis of the implicit component of cognitive bias, and justifies its choice in terms of the objectives of this review. The method of recording event-related potentials (ERPs) was chosen as the main approach to determining neurophysiological indicators of implicit processes. The analysis of literature allowed us to identify ERP components reproduced in the works of different authors using different variants of experimental designs for studying implicit bias, which may indicate the presence of common neurophysiological mechanisms associated with implicit processes in cognitive bias. The possibility of using other approaches to the analysis of EEG data to obtain new information about the mechanisms of implicit components in cognitive bias is also discussed.

**Keywords:** cognitive bias, implicit component, event-related potentials, ERP, Implicit Association Test, IAT.