



# НЕЙРОБИОЛОГИЧЕСКИЕ МЕХАНИЗМЫ СОЦИАЛЬНОГО ВЛИЯНИЯ

**КЛЮЧАРЕВ В.А.\***, Национальный исследовательский университет «Высшая Школа Экономики», Москва, Россия; Санкт-Петербургский государственный университет, Санкт-Петербург, Россия; Университет Базеля, Швейцария,  
e-mail: vasily.klucharev@unibas.ch

**ЗУБАРЕВ И.П.\*\***, Санкт-Петербургский государственный университет, Санкт-Петербург, Россия,  
e-mail: ipzubarev@gmail.com

**ШЕСТАКОВА А.Н.\*\*\***, Санкт-Петербургский государственный университет, Санкт-Петербург, Россия,  
e-mail: anna.shestakova@helsinki.fi

В последние годы развитие неинвазивных методов исследования активности головного мозга человека позволило пролить свет на нейробиологические механизмы социального влияния на принятие решений. В данном обзоре представлены современные нейробиологические исследования социального влияния и обсуждаются ключевые теоретические модели возможных нейробиологических механизмов, лежащих в основе социального влияния. Ряд исследований демонстрирует тесную связь социального влияния с нейрофизиологической системой обучения с подкреплением и контроля деятельности. Вероятно, отклонение от поведения (мнения) группы интерпретируется нервной системой как поведенческая ошибка или «ошибка предсказания вознаграждения», что запускает процесс изменения поведения на основе дофаминергического механизма обучения с подкреплением.

**Ключевые слова:** социальное влияние, конформизм, нейроэкономика, нейробиология, ростральная цингулярная зона, дофаминергическая система, ошибка предсказания вознаграждения.

## Введение

Идея о том, что индивидуальные решения человека в значительной степени обусловлены поведением окружающих людей, является краеугольным камнем социальной психологии (Ajzen, Fishbein, 1980; Asch, 1955; Cialdini, Goldstein, 2004; Milgram, 1974). Однако до настоящего времени большинство нейробиологических исследований рассматривали механизмы принятия решений отдельно от социального окружения. Достижения нейробиологии последних лет не только позволили рассмотреть биологические основы принятия решений, но также способствовали развитию новой междисциплинарной области знания – нейробиологии принятия решений, или нейроэкономики (Glimcher, Rustichini, 2004),

### Для цитаты:

Ключарев В.А., Зубарев И.П., Шестакова А.Н. Нейробиологические механизмы социального влияния // Экспериментальная психология. 2014. Т. 7. № 4. С. 20–36.

\* Ключарев В.А. Кандидат биологических наук, профессор, руководитель Департамента психологии, Национальный Исследовательский университет «Высшая школа экономики», Москва, Россия; Санкт-Петербургский государственный университет, Санкт-Петербург, Россия; Университет Базеля, Швейцария. E-mail: vasily.klucharev@unibas.ch

\*\* Зубарев И.П. Санкт-Петербургский государственный университет, Санкт-Петербург, Россия. E-mail: ipzubarev@gmail.com

\*\*\* Шестакова А.Н. PhD, старший научный сотрудник, биологический факультет, Санкт-Петербургский государственный университет, Санкт-Петербург, Россия. E-mail: anna.shestakova@helsinki.fi



объединившей нейробиологию, экономические науки, социальную психологию и ряд других дисциплин, знания которых легли в основу новых представлений о механизмах принятия решений (Ключарев, Шмидс, Шестакова, 2011). Данный научный обзор представляет собой анализ основных проведенных в последние годы нейробиологических исследований, результаты которых позволили предположить существование связи социального влияния с фундаментальными нейробиологическими (дофаминергическими) механизмами процессов обучения.

### Конформное поведение

Первые нейробиологические исследования социального влияния были сфокусированы на изучении нейробиологических механизмов конформного поведения. Под конформным поведением понимается особая форма социального влияния, при которой индивид изменяет свои убеждения, оценки и поведение в соответствии с нормами определенной социальной (референтной) группы даже в отсутствие прямого на то указания или требования. Согласно представлениям Роберта Чалдини и Джефа Голдштейна (Cialdini, Goldstein, 2004), конформное поведение возникает под воздействием двух типов социальных норм:

- *прескриптивные (предписывающие) нормы* содержат в себе указания на ожидаемую, требуемую модель поведения, и, как правило, содержат моральную оценку такого поведения;
- *дескриптивные (описательные) нормы* описывают типичное поведение большинства людей в данной группе, в данной ситуации вне зависимости от моральной окраски такого поведения.

В то время как прескриптивные нормы (например, «Нельзя садиться за руль в нетрезвом виде») выполняют регулятивную, нормативную функцию, регламентируя человеческое поведение, а их несоблюдение чревато теми или иными санкциями; дескриптивные нормы, напротив, лишь описывают доминирующее поведение большинства людей и, таким образом, носят скорее информативный характер: примером таких норм может служить склонность людей придерживаться модных тенденций в одежде, или мода на курение (или употребление наркотиков) в подростковых субкультурах. Дескриптивные нормы (и ассоциированное с ними социальное одобрение – валидация) также весьма эффективны. Именно дескриптивные нормы в значительной мере определяют, будут ли граждане уклоняться от налогов или, например, будут ли домохозяйства придерживаться экологических норм (Kenrick, 2003). Интересно, что хотя в повседневной жизни влияние описательных норм достаточно велико, большинство людей не осознают их воздействия (Bryan, Test, 1967).

Конформное поведение может определяться различными видами мотивации (Cialdini, Goldstein, 2004):

- стремлением к наиболее точной интерпретации реальности и выбору наиболее эффективной поведенческой стратегии;
- стремлением получить одобрение окружающих;
- стремлением поддерживать положительную самооценку, избегать когнитивного диссонанса.

В работах в области социальной психологии особо выделяется роль общественного одобрения как фактора, подкрепляющего такое конформное поведение (Cialdini, Goldstein, 2004), в то время как исследователи в области поведенческой экономики при анализе причин конформного поведения акцентируют внимание на санкциях, которые, как правило, вводятся за нарушение общественных норм (Fehr, Fischbacher, 2004). Однако, примени-



тельно к обоим указанным подходам, соблюдение социальных норм может быть описано в терминах теории обучения с подкреплением. В данном случае *социальная валидация* подкрепляет сходные с нормой поведенческие паттерны и наказывает отклонения от нормы. Аналогичным образом можно предположить, что результат любого действия может оцениваться и нервной системой с точки зрения не только достижения индивидуальных целей, но также и соответствия результата социальным нормам. С этой точки зрения рассогласования результата поведения с индивидуальными целями и социальными нормами трактуются как ошибки, требующие коррекции поведения. Можно ли экспериментально проверить, действительно ли социальное влияние модифицирует наше поведение посредством фундаментального механизма контроля результатов деятельности?

В ряде нейроэкономических исследований (Klucharev et al., 2009, 2011; Shestakova et al., 2012) были продемонстрированы свидетельства в пользу того, что механизмы социального влияния являются частью фундаментального нейронального механизма контроля деятельности, описанного в рамках модели обучения с подкреплением (Sutton, Barto, 1998). В работе Ключарева с соавт. (Klucharev et al., 2009) было показано, что отклонение от нормативного мнения группы генерирует нейрональный ответ, схожий с так называемой *ошибкой предсказания вознаграждения* (ОП) – ключевым параметром в теории обучения с подкреплением. Предполагается, что ОП-сигнал является *сигналом обучения* и кодирует необходимость модифицировать поведение с тем, чтобы избежать повторения поведенческих ошибок в будущем. Таким образом, была сформулирована гипотеза о том, что в ситуации рассогласования мнения индивида с мнением референтной группы в его нервной системе генерируется ОП-сигнал, обозначающий необходимость изменить мнение, чтобы соответствовать социальной норме. Рассмотрим роль данного сигнала в процессе обучения подробнее.

### Дофаминергическая система обучения с подкреплением

В основе адаптивного поведения лежит способность субъекта обучаться, основываясь на информации о результатах своих действий (Thorndike, 1911). Теория обучения с подкреплением формализует механизм этого процесса (Sutton, Barto, 1998). При этом разность *ожидаемого* и *полученного* результатов соответствует ОП, являющейся сигналом к обучению. Получая информацию о результате своего действия, субъект всякий раз генерирует ОП и новое предсказание о ценности (V) данной ситуации/стимула (Sutton, Barto, 1998):

$$V_{\text{new}} = V_{\text{old}} + \eta (R - V_{\text{old}}),$$

где  $V_{\text{new}}$  – субъективное предсказание ожидаемой ценности вознаграждения, следующего за данным стимулом (например, за звуковым тоном, предшествующим появлению еды), которая является функцией предыдущего предсказания  $V_{\text{old}}$  и ОП, являющейся разницей между ценностью реально полученного в данной пробе вознаграждения (R) и ожидаемым вознаграждением ( $V_{\text{old}}$ ). Параметр  $\eta$  – индивидуальная скорость обучения данного индивида. ОП ( $R - V_{\text{old}}$ ) рассчитывается как разность между ценностью результата действия и ожиданием организма, а ОП-сигнал генерируется в момент наступления результата действия.

Сходная концепция обучения с подкреплением, получившая название *модели временных различий* (temporal difference model), также постулирует наличие ОП-сигнала, но логически отделяет его от момента наступления результата и моделирует его как некий континуум или перманентное изменение предсказания о будущих значениях результата действия, происходящее вплоть до наступления последнего (Niv, Schoenbaum, 2008). Ряд исследований связывает генерацию ОП с работой дофаминергической системы мозга – совокупности взаимосвязанных нейронных ансамблей, использующих в качестве нейромедиа-



тора дофамин (Schultz, 2007). Динамика импульсации дофаминергических нейронов среднего мозга приматов соответствует теоретической модели ОП-сигнала, возникающего, как свидетельствуют данные исследований, в ситуациях с различными комбинациями ожидаемых и получаемых вознаграждений. В частности, если результат действия оказался выше ожидаемого (позитивная ОП), происходит фазическое повышение частоты импульсации дофаминергических нейронов, тогда как если результат действия оказывается ниже ожидаемого (негативная ОП), то происходит фазическое снижение активности дофаминергических нейронов. Нейроимиджинговые исследования человека подтвердили факт генерации ОП-сигнала при рассогласовании ожидаемого результата с полученным в компактной части черной субстанции среднего мозга (substantia nigra – SNc), в вентральной области покрышки (ventral tegmental area – VTA), прилежащем ядре (nucleus accumbens – NAc), в медиальной префронтальной коре (posterior medial frontal cortex – pmPFC) и в других областях, богатых дофамином (Gehring et al., 1993; McClure et al., 2004; O'Doherty et al., 2004). В целом, модель обучения с подкреплением хорошо описывает некоторые формы обучения, основанные на предсказании немедленного и отсроченного вознаграждения, которое позволяет выбирать и корректировать наиболее выгодные поведенческие стратегии.

Особый отдел лобной коры (рис. 1А) – роstralная цингулярная зона (rostral cingulate zone – RCZ) – подробно изучен в контексте генерации ОП-сигнала. RCZ определяется как область коры на границах 6, 8, 24 и 32 полей Бродмана (Picard, Strick, 1996). Когнитивные исследования указывают на то, что RCZ активируется в случае необходимости изменения поведения как вследствие ошибки (т.е. недостижения цели поведения), так и тогда, когда вероятность такой ошибки велика (Cohen, Ranganath, 2007; Ridderinkhof et al., 2004). В экспериментальных задачах, предусматривающих выбор между несколькими альтернативами, показатели уровня активации RCZ позволяют эффективно предсказывать последующие поведенческие изменения (Cohen, Ranganath, 2007; Kerns et al., 2004). Более того, согласно одной из моделей обучения с подкреплением, в независимости от причин отклонения реального результата от ожидания в положительную или отрицательную сторону (Holroyd, Coles, 2002; Matsumoto et al., 2007), активность RCZ модулируется дофаминергическим ОП-сигналом, кодирующим соотношение ожидаемого и реального результата действия.

Результаты многочисленных исследований указывают на наличие распределенной функциональной сети участвующих в данном процессе нервных центров, включающей, помимо прочих, вентральную часть стриатума – NAc. Так, было продемонстрировано, что NAc вовлечена в предсказание вознаграждения и активируется на сигналы, предшествующие получению подкрепления (Knutson, Wimmer, 2007). В экспериментах на животных было показано, что разрушение связей RCZ с NAc ведет к нарушению процесса принятия решений и обучения в целом (Hauber, Sommer, 2009; Parkinson et al., 2000). На основании изложенных фактов можно предположить существование в головном мозге специализированной системы, вовлеченной в контроль результатов деятельности и генерацию ОП-сигнала обучения и включающей RCZ и NAc.

### **Гипотеза механизма социального влияния**

Исходя из предположения, что всякий результат действия оценивается организмом с точки зрения как его субъективной «физиологической» ценности, так и «социальной» ценности соответствия собственного поведения поведению группы, отклонение от групповой нормы должно интегрироваться в общий сигнал ОП. Можно предположить, что рас-

согласование мнения индивида с мнением референтной группы должно генерировать ОП-сигнал, аналогичный тому, который регистрируется при совершении ошибок в их классическом смысле (например, при денежном проигрыше или при неверном ответе на вопрос). Данный сигнал также должен отражать активацию структур RCZ и NAc, вовлеченных в обработку результатов действия. Такой сигнал должен отражать как величину отклонения от мнения социального окружения, так и степень последующей коррекции поведения. Иными словами, в соответствии с данной гипотезой социального влияния, по силе активации RCZ и NAc можно с достаточной степенью надежности судить о степени подверженности социальному влиянию конкретного индивида. Важно подчеркнуть, что конфликт с социальной нормой является не ошибкой поведения в строгом смысле этого слова, а, по сути, – любым отклонением от поведения большинства (Montague, Lohrenz, 2007).

### **Исследования социального влияния методами неинвазивной нейровизуализации**

Первой работой по исследованию социального влияния методами неинвазивного нейроимиджинга принято считать работу Грегори Бернса и соавторов (Berns et al., 2005), имитирующую парадигму Соломона Эша и посвященную исследованию нейробиологических основ конформного поведения в задаче на мысленное вращение фигур. Однако данное направление исследований не получило дальнейшего развития, и последующие работы по изучению проблемы социального влияния были связаны с разработкой новой исследовательской парадигмы, основанной на применении методов функционального картирования мозга.

Ключарев с соавторами (Klucharev et al., 2009) разработали новую экспериментальную парадигму, основанием которой является сопоставление индивидуального мнения с мнением референтной группы на примере оценки привлекательности лиц. Такая организация эксперимента позволяет моделировать ситуацию конфликта мнения индивида с мнением группы. При этом в ситуации конфликта мнения испытуемого с мнением группы регистрировалась активация RCZ (отражающая, как указывалось выше, ОП-сигнал), а также деактивация NAc, на основании чего авторы сделали вывод о том, что рассогласование мнений испытуемого и мнения большинства автоматически генерирует нейронную активность, напоминающую ОП-сигнал. Более того, амплитуда деактивации NAc показала корреляцию с последующим изменением мнения под действием социального влияния. В целом, в данном и других ранних исследованиях социального влияния было показано, что отличие мнения индивида от мнения окружающих модулирует активность RCZ (Behrens et al., 2008; Berns et al., 2010; Campbell-Meiklejohn et al., 2010; Falk et al., 2010; Klucharev et al., 2009) и NAc (Campbell-Meiklejohn et al., 2010; Klucharev et al., 2009) – областей мозга, вовлеченных в систему контроля результатов деятельности и генерацию ОП-сигнала.

Гипотеза об участии RCZ в процессе развития конформизма была проверена посредством временного подавления активности данной области методом транскраниальной магнитной стимуляции – ТМС (Klucharev et al., 2011). Результаты эксперимента показали, что ТМС-подавление RCZ вызвало снижение степени и вероятности конформных изменений мнения (рис. 1Б); более того, модуляция активности RCZ посредством ТМС приводит к значительному снижению степени подверженности социальному влиянию. Таким образом, полученные данные свидетельствуют о том, что восприимчивость к социальному влиянию может объясняться модуляцией работы *системы контроля результатов деятельности* и, соответственно, приводить к последующим изменениям поведения (Klucharev et al., 2011).

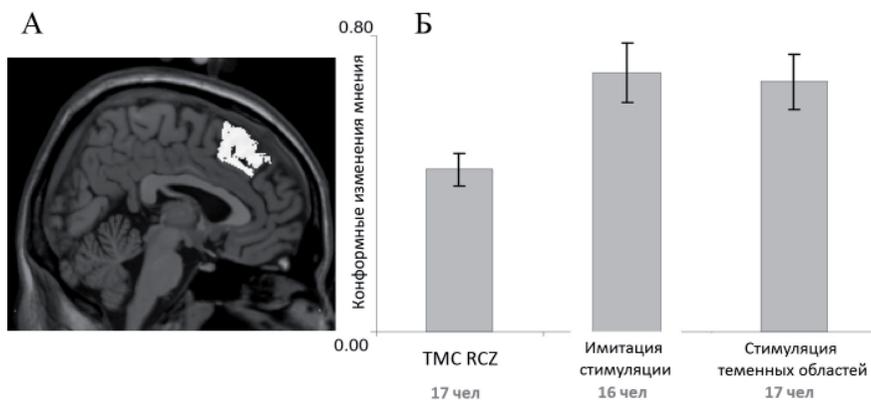


Рис. 1. А. Локализация ростральной цингулярной зоны (RCZ), включая парацингулярную кору, и область RCZ в экспериментах с целевой и контрольной ТМС-стимуляцией (отмечены стрелками). Б. Средняя величина изменения оценки привлекательности лиц в сторону групповой оценки: в группе испытуемых, прошедших ТМС-стимуляцию RCZ (левый столбец); в группе с имитацией стимуляции (средний столбец); в группе испытуемых, прошедших контрольную стимуляцию медиальной теменной области коры (правый столбец). Адаптировано из (Klucharev et al., 2011)

Результаты многочисленных зарубежных исследований также свидетельствуют в пользу высказанного предположения о вовлечении RCZ в процесс изменения мнения индивида под воздействием социального окружения. Так, активация RCZ и NAc была продемонстрирована в исследованиях Кэмпбэлл-Мейклджен и соавторов в ситуации рассогласования мнений испытуемого с мнением двух экспертов в задаче на оценку музыкальных композиций (Campbell-Meiklejohn et al., 2010); при этом амплитуда сигнала была значительно выше у испытуемых, более склонных изменять свои оценки под воздействием социального влияния. Аналогичным образом Бернс и соавторы (Berns et al., 2010) анализировали влияние мнения группы сверстников на музыкальные предпочтения подростков и показали связь активации RCZ с предрасположенностью к изменению мнения в сторону группового (Berns et al., 2010). В исследовании Беренса и коллег (Behrens, Hunt, Rushworth, 2009) была экспериментально проверена гипотеза о том, что социальная информация обрабатывается структурами, ассоциируемыми с системой обучения с подкреплением. Испытуемые выполняли задачу, в которой было необходимо сочетать информацию, получаемую методом проб и ошибок, с подсказками других людей; при этом такие подсказки были верны лишь в части случаев. Полученные в исследовании данные указывают на активацию двух областей медиальной префронтальной коры (включая RCZ), кодирующих «социальную» информацию и информацию, полученную в результате собственных действий испытуемых; более того, оба потока информации интегрировались в вентромедиальных областях префронтальной коры (там же). В одном из новейших исследований конформности, проведенном группой Кейса Изумы (Izuma, Adolphs, 2013), было показано, что направление изменений мнения и активация RCZ в ситуации социального влияния зависят от положительной или отрицательной самоидентификации испытуемого с референтной группой: и в ситуации рассогласования индивидуального мнения с мнением «положительно» самоидентифицируемой группы, и в ситуации совпадения индивидуального мнения с мнением «отрицательно» самоидентифицируемой группы генерируется сигнал, аналогичный ОП-ответу. При этом, как и в перечисленных выше работах, сила активации медиальных областей префронтальной коры (вклю-

чая RCZ) достоверно коррелировала со степенью изменения мнения, в том числе зарегистрированного четыре месяца спустя. Эти и другие фМРТ-исследования указывают на то, что дофаминергическая система, включающая NAc и RCZ, вовлечена в процесс изменения индивидуального мнения под влиянием социального окружения. Примечательно, что, в полном соответствии с выдвинутым предположением, увеличение концентрации дофамина в мозге с помощью *метилфенидата*, как показывают полученные Кэмпбэлл-Мейклджен с сотрудниками (Campbell-Meiklejohn et al., 2012) данные, способствует повышению склонности индивида к изменению собственного мнения в сторону мнения окружающих.

Однако возникает вопрос о том, действительно ли изменения индивидуального мнения, фиксируемые в экспериментальных условиях с применением фМРТ-сканера, отражают реальную его внутреннюю переоценку (*private acceptance*), а не стремление лишь внешне продемонстрировать изменение мнения (*public compliance*)? Известно, что уровень активации NAc кодирует различные степени предпочтения тех или иных стимулов (Elliott et al., 2004) даже в условиях, когда задача не требует непосредственной оценки стимула (O'Doherty et al., 2004; Tricomi et al., 2004). Исследования Малии Мейсон с соавторами и Джамии Заки с соавторами продемонстрировали, что изменения оценок абстрактных символов (Mason et al., 2009) и привлекательности женских лиц (Zaki et al., 2011), вызванные конфликтом с групповыми оценками, также сопровождались соответствующими долгосрочными изменениями в активации NAc. Так, Заки с соавт. (Zaki et al., 2011) показали, что при повторном предъявлении изображений лиц, индивидуальное мнение о которых ранее расходилось с мнением социальной группы, активность NAc отражает предшествующее социальное влияние: она повышена в случае, когда группа ранее оценивала лицо на предъявленном изображении более привлекательным, и понижена в том случае, когда группа находила его менее привлекательным, что, в общем, свидетельствует о интернализации мнения группы испытуемыми. Таким образом, можно выдвинуть предположение о том, что в ситуации социального влияния имеет место реальная переоценка индивидом субъективной ценности стимулов (*private acceptance*).

### **Негативность результата действия (FRN) как возможный электроэнцефалографический маркер социального влияния**

Гипотеза о том, что социальное влияние может быть опосредовано *системой контроля результатов деятельности*, вовлеченной в обучение с подкреплением, находит дополнительное подтверждение в ряде работ, выполненных методом электроэнцефалографии (Chen et al., 2012; Kim et al., 2012; Shestakova et al., 2012). Принципы теории обучения с подкреплением наилучшим образом объясняют динамику возникновения компонента вызванных потенциалов (ВП), называемого «Негативность Результата Действия», *feedback-related negativity* – FRN (Holroyd, Coles, 2002; Nieuwenhuis et al., 2004; Walsh, Anderson, 2012). Наиболее вероятным источником, генерирующим FRN, считается поясная извилина (включая RCZ) (Cohen, Ranganath, 2007; Gehring, Willoughby, 2002; Hewig et al., 2007; Miltner et al., 1997; Nieuwenhuis et al., 2005; Tucker et al., 2003). Считается, что FRN отражает ОП-сигнал – фазическое снижение дофаминергической импульсации, растормаживающее RCZ, что выражается в более негативном отклонении ВП (Holroyd, Coles, 2002).

Используя метод ВП, Шестакова и соавторы (Shestakova et al., 2013) показали, что конфликт с мнением группы вызывает нейрональную активность во фронтотентальной



области коры (негативное отклонение ВП около 200 мс), сходное по характеристикам с FRN (рис. 2). Аналогичные результаты были получены в исследовании Д. Чен и соавторов (Chen et al., 2012) в задаче, где испытуемому предлагалось, аналогично задаче Эша, выбрать, какая линия совпадает по длине с целевой линией.

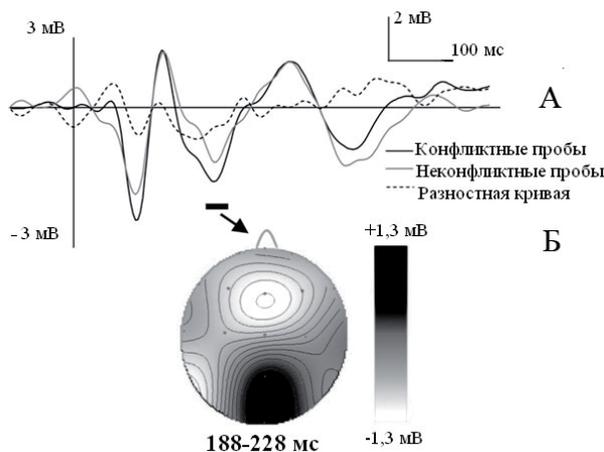


Рис. 2. А. Различия между компонентом FRN в отведении Fz в экспериментальной ситуации рассогласования мнения испытуемого с мнением группы (черный цвет – конфликтные пробы) и аналогичным компонентом в ситуации, когда мнение испытуемого совпадало с мнением группы (серый цвет – неконфликтные пробы). Разностная кривая изображена пунктирной линией. Б. Изопотенциальное распределение вызванного ответа в интервале 188–228 мс (адаптировано из Shestakova et al., 2013)

Результаты ряда исследований с использованием различных фармакологических манипуляций свидетельствуют об усилении амплитуды FRN при введении агонистов (амфетамин) и, наоборот, об уменьшении FRN при введении антагонистов дофамина (галоперидол, прампексол) (de Bruijn et al., 2006; Zirnheld et al., 2004), указывая, таким образом, на участие дофаминергических структур в генерации FRN, хотя и не исключая возможности участия других медиаторных систем (Jocham, Ullsperger, 2009).

В целом, весь объем существующих ЭЭГ- и фМРТ-исследований соответствует механизму социального влияния, формализованному в рамках моделей «обучения с подкреплением» – фундаментальной системы контроля результатов деятельности.

### Нейробиологические корреляты убеждения

В отличие от конформного поведения, когда изменение индивидуального мнения происходит без явного целенаправленного влияния, изменение мнения в результате *убеждения* зачастую является результатом активного направленного процесса. При этом убеждение не является однородным процессом, поскольку индивид может изменить свое мнение вследствие различных причин. Ричард Петти и Джон Качоппо (Petty, Cacioppo, 1986; см. также Petty, Wegener, 1999) подчеркивают два принципиально различных типа убеждения: (а) *прямой*, основанный на логической аргументации, и (б) *косвенный*, основанный на запуске «механизма автоматического согласия». В частности, авторы выделяют ряд факторов, способствующих эффективности процесса убеждения, таких как: 1) авторитет коммуникатора; 2) содержание сообщения; 3) способ подачи сообщения; 4) способность аудитории воспринимать тот или иной тип аргументов.

Немногочисленные исследования нейробиологических механизмов убеждения позволили, тем не менее, выделить ряд объективных маркеров эффективности убеждения. Так, например, в экспериментах, выполненных методом фМРТ, было показано, что показатели уровня и интенсивности активации передней части медиальной префронтальной коры (включая RCZ) во время просмотра испытуемым видеороликов о вреде курения обладают более значительной прогностической способностью относительно того, бросит ли этот человек курить, чем его собственные отчеты об убедительности этих роликов (Chua et al., 2009; Falk et al., 2010; Falk et al., 2011). Авторы заключают, что активация данной области коры может служить более надежным маркером интернализации убеждающего сообщения, нежели данные опросников.

### Когнитивный диссонанс и социальное влияние

Идея согласованности когнитивных структур, знаний, элементов восприятия и ожиданий, связанная с концепцией *когнитивного диссонанса* (Festinger, 1959) и теорией структурного баланса (Heider, 1958), является краеугольным положением теории социального влияния. Данная идея основывается на утверждении о том, что предпочтительной для индивида является ситуация, когда новая информация согласуется с уже имеющимися оценками, мнениями и суждениями; в случае, когда поступающая информация вызывает несогласование с уже имеющейся, индивид стремится минимизировать это несогласование.

Под когнитивным диссонансом понимается состояние психического дискомфорта, вызванное столкновением конфликтующих представлений: идей, верований, ценностей или эмоциональных реакций. Предполагается, что психологический дискомфорт, возникающий в ситуации, когда поведение индивида несогласуется с его психологическими установками, вынуждает большинство людей изменять свои установки. Экспериментально было показано, что после вынужденного совершения неприятных действий, как, например, поедание необычной пищи (Zimbardo et al., 1965) или написания (студентом) эссе в пользу увеличения стоимости обучения в вузе (Steele et al., 1981), большинство испытуемых склонно относиться к ситуациям или темам, ранее вызывавшим психологический дискомфорт, более положительно, уменьшая, таким образом, возникающий когнитивный диссонанс. Подобный диссонанс может возникать, например, если отношение индивида к некоторому объекту «А» отличается от отношения группы людей, которые ему нравятся, т.е. возникает диссонанс: «А» нравится мне, но не нравится моим друзьям. Аналогичным образом, когнитивный диссонанс возникает и в ситуации, когда мнение индивида об объекте «А» совпадает с мнением членов референтной для индивида группы. В обоих случаях индивид стремится устранить несогласование посредством изменения своего отношения либо к объекту, либо к окружающим. Тут уместно вновь обратиться к исследованию К. Изумы (Izuma, Adolphs, 2013), который полагает, что в основе нейробиологического механизма, обеспечивающего изменение мнения индивида под действием социального влияния, может лежать именно стремление сократить когнитивный диссонанс. Действительно, активация RCZ наблюдается не только в ситуациях влияния социального окружения, но и в ситуациях развития психического дискомфорта, вызванного столкновением в сознании противоречащих друг другу представлений или эмоций (Izuma et al., 2010; van Veen et al., 2009). Результаты проведенного исследования, свидетельствовавшие о вовлеченности RCZ в процесс переработки когнитивной информации и преодоления психического дискомфорта, позволили К. Изуме сделать вывод о том, что активность RCZ отражает некий *индекс вну-*



*тренней согласованности мнения или поведения индивида*, т.е. внутренний диссонанс в ситуации социального влияния (Izuma, Adolphs, 2013; Izuma, 2013). Однако, по нашему мнению, модель когнитивного диссонанса может являться частным случаем обучения с подкреплением. Если принять мнение социальной группы в качестве одного из параметров, сигнализирующих о вероятности получения вознаграждения, то *модель временных различий* предсказывает генерацию ОП-сигнала как в ситуации, когда мнение испытуемого расходится с мнением «предпочитаемой» референтной группы, так и тогда, когда оно совпадает с оценкой «не предпочитаемой» группы. Иными словами, в обеих ситуациях имеет место рассогласование результата действия и ожидаемого результата, что и было показано в исследовании Изумы (Izuma et al., 2013). В целом, по нашему мнению, когнитивный диссонанс может отражать конфликт результата текущей деятельности и ожиданий, связанных с долгосрочными целями индивида – процесса, функционально сходного с механизмом ОП-сигнала.

### Заключение

Результаты исследований специфики социального влияния, выполненные методами неинвазивной нейровизуализации, подтверждают гипотезу о том, что механизм, посредством которого осуществляется социальное влияние, тесно связан с *системой контроля результата деятельности*. Согласно данной гипотезе, нервная система постоянно отслеживает не только базовые результаты поведения (получение вознаграждения или наказания), но и соответствие индивидуального поведения нормативному поведению референтной социальной группы. С этой точки зрения как отсутствие ожидаемого вознаграждения, так и отклонение поведения (мнения) от поведения группы являются поведенческой ошибкой, отражающейся в ОП-сигнале, генерируемом при участии дофаминергической системы мозга. Именно этот сигнал и запускает изменение нашего поведения с целью привести его в соответствие с социальной нормой.

Однако почему человеку свойственно такое автоматическое конформное поведение? Вероятно, тому существуют серьезные эволюционные предпосылки. Одним из объяснений может служить концепция *эволюционно стабильного* поведения большинства (Dawkins, 1976), в соответствии с которой, в неизменной среде большинство особей в популяции будет придерживаться одной и той же стратегии только в том случае, если эта стратегия (поведение) имеет преимущества перед всеми альтернативами, а отклонения от такой оптимальной стратегии большинства наказываются в ходе естественного отбора; таким образом, как с точки зрения *эволюционной теории*, так и с точки зрения *теории игр*, рационально следовать за большинством.

Исследования нейробиологических механизмов, лежащих в основе регулирования социального поведения в ситуациях социального воздействия, представляют безусловный интерес не только для фундаментальной науки, поскольку не только существенно расширяют имеющуюся методологию исследований социального поведения, предоставляя возможность строить более точные модели социального поведения и осуществлять их экспериментальную проверку, но также находят широкое применение в таких сферах, как социальная реклама, пропаганда здорового образа жизни и т.п. Нейробиологические данные воссоздают перед нами картину работы динамической системы мозга, постоянно отслеживающей результаты нашего поведения, сравнивающей их с нашими ожиданиями и поведением нашей социальной группы и автоматически меняющей наше поведение для достижения более оптимального результата. Несмотря на то, что начало нейробиологическим ис-

следованиям было положено лишь несколько лет назад, их результаты уже успешно продемонстрировали плодотворность синтеза нейробиологических подходов и методов социальных наук в целях формулирования фундаментального нейробиологического механизма социального влияния – процесса, изо дня в день модулирующего наше мнение, повседневное поведение и даже научную деятельность.

#### Финансирование

Данная работа выполнена при поддержке Федеральной Целевой Программы «Кадры» – Мероприятие 1.5 (Соглашение № 8488), а также Российского Фонда Фундаментальных Исследований (грант 11-06-00449-а) и Гранта СПбГУ 0.37.522.2013.

#### Литература

1. Ключарев В.А., Шмидс А., Шестакова А.Н. Нейроэкономика: нейробиология принятия решений // Экспериментальная психология. 2011. Т. 4. № 2. С. 14–35.
2. Ajzen I., Fishbein M. Understanding attitudes and predicting social behavior. Englewood Cliffs, NJ: Prentice-Hall, 1980. 278 p.
3. Asch S.E. Opinions and social pressure // Scientific American. 1955. Vol. 193. № 5. P. 31–35.
4. Behrens T.E., Hunt L.T., Rushworth M.F. The computation of social behavior // Science. 2009. Vol. 324. (5931). P. 1160–1164. doi: 10.1126/science.1169694
5. Behrens T.E., Hunt L.T., Woolrich M.W., Rushworth M.F. Associative learning of social value // Nature. 2008. Vol. 456. P. 245–249.
6. Berns G.S., Capra C.M., Moore S., Noussair C. Neural mechanisms of the influence of popularity on adolescent ratings of music // Neuroimage. 2010. Vol. 49. P. 2687–2696.
7. Berns G.S., Chappelow J., Zink C.F., Pagnoni G., Martin-Skurski M.E., Richards J. Neurobiological correlates of social conformity and independence during mental rotation // Biological Psychiatry. 2005. Vol. 58. P. 245–253.
8. Bryan J.H., Test M.A. Models and helping: naturalistic studies in aiding behavior // Journal of Personality and Social Psychology. 1967. Vol. 6. P. 400–407.
9. Campbell-Meiklejohn D.K., Bach D.R., Roepstorff A., Dolan R.J., Frith C.D. How the opinion of others affects our valuation of objects // Current Biology. 2010. Vol. 20. P. 1165–1170.
10. Campbell-Meiklejohn D.K., Simonsen A., Jensen M., Wohlert V., Gjerloff T., Scheel-Kruger J., Moller A., Frith C.D., Roepstorff A. Modulation of social influence by methylphenidate // Neuropsychopharmacology. 2012. Vol. 37. № 6. P. 1517–1525.
11. Chen J., Wu Y., Tong G. Y., Guan X.M., Zhou X.L. ERP correlates of social conformity in a line judgment task // BMC Neuroscience. 2012. Vol. 13. P. 13–43. doi: 10.1186/1471-2202-13-43.
12. Chua H.F., Liberzon I., Welsh R. C., Strecher V.J. Neural correlates of message tailoring and self-relatedness in smoking cessation programming // Biological Psychiatry. 2009. Vol. 65. P. 165–168.
13. Cialdini R.B., Goldstein N.J. Social influence: compliance and conformity // Annual Reviews in Psychology. 2004. Vol. 55. P. 591–621.
14. Cohen M.X., Ranganath C. Reinforcement learning signals predict future decisions // Journal of Neuroscience. 2007. Vol. 27. P. 371–378.
15. Dawkins R. The Selfish Gene. New York: Oxford University Press, 1976. 512 p.
16. de Bruijn E.R.A., Sabbe B.G.C., Hulstijn W., Ruigt G.S.F., Verkes R.J. Effects of antipsychotic and antidepressant drugs on action monitoring in healthy volunteers // Brain Research. 2006. Vol. 1105. P. 122–129.
17. Elliott R., Newman J.L., Longe O.A., Deakin J.F.W. Instrumental responding for rewards is associated with enhanced neuronal response in subcortical reward systems // NeuroImage. 2004. Vol. 21. P. 984–990.
18. Falk E.B., Berkman E.T., Mann T., Harrison B., Lieberman M.D. Predicting persuasion-induced behavior change from the brain // Journal of Neuroscience. 2010. Vol. 30. P. 8421–8424.
19. Falk E.B., Berkman E.T., Whalen D., Lieberman M.D. Neural activity during health messaging predicts reductions in smoking above and beyond self-report // Health Psychology. 2011. Vol. 30. P. 177–185.



20. *Fehr E., Fischbacher U.* Social norms and human cooperation // *Trends in Cognitive Sciences*. 2004. Vol. 8. P. 185–190.
21. *Festinger L., Carlsmith J.M.* Cognitive consequences of forced compliance // *Journal of Abnormal Psychology*. 1959. Vol. 58. № 39. P. 203–210.
22. *Gehring W.J., Goss B., Coles M.G.H., Meyer D.E., Donchin E.* A neural system for error detection and compensation // *Psychological Science*. 1993. Vol. 4. P. 385–390.
23. *Gehring W.J., Willoughby A.R.* The medial frontal cortex and the rapid processing of monetary gains and losses // *Science*. 2002. Vol. 295. P. 2279–2282.
24. *Glimcher P.W., Rustichini A.* Neuroeconomics: the consilience of brain and decision // *Science*. 2002. Vol. 306. P. 447–452.
25. *Hauber W., Sommer S.* Prefrontostriatal circuitry regulates effort-related decision making // *Cerebral Cortex*. 2009. Vol. 19. P. 2240–2247.
26. *Heider F.* *The Psychology of Interpersonal Relations*. New York: Wiley, 1958. 322 p.
27. *Hewig J., Trippe R., Hecht H., Coles M.G.H., Holroyd C.B., Miltner W.H.R.* Decision-making in blackjack: an electrophysiological analysis // *Cerebral Cortex*. 2007. Vol. 17. P. 865–877.
28. *Holroyd C.B., Coles M.G.H.* The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity // *Psychology Reviews*. 2002. Vol. 109. P. 679–709.
29. *Izuma K.* The neural basis of social influence and attitude change [Electronic edition] // *Current Opinion in Neurobiology*. 2013. Vol. 23. URL: <http://dx.doi.org/10.1016/j.conb.2013.03.009> (Accessed at: 15. 11.2014).
30. *Izuma K., Adolphs R.* Social manipulation of preference in the human brain [Electronic edition] // *Neuron*. 2013. Vol. 78. P. 563–573. URL: <http://dx.doi.org/10.1016/j.neuron.2013.03.023> (Accessed at: 15. 11.2014).
31. *Izuma K., Matsumoto M., Murayama K., Samejima K., Sadato N., Matsumoto K.* Neural correlates of Cognitive dissonance and choice-induced preference change // *Proceedings of National Academy of Sciences USA*. 2010. Vol. 107. P. 22014–22019.
32. *Jocham G., Ullsperger M.* Neuropharmacology of performance monitoring // *Neuroscience and Biobehavioral Reviews*. 2009. Vol. 33. P. 48–60.
33. *Kenrick D.T., Li N.L., Butner J.* Dynamical evolutionary psychology: Individual decision rules and emergent social norms // *Psychological Review*. 2003. Vol. 110. P. 3–28.
34. *Kerns J.G., Cohen J.D., MacDonald A.W., Cho R.Y., Stenger V.A., Carter C.S.* Anterior cingulate conflict monitoring and adjustments in control // *Science*. 2004. Vol. 303. P. 1023–1026.
35. *Kim B.R., Liss A., Rao M., Singer Z., Compton R.J.* Social deviance activates the brain's error-monitoring system // *Cognitive Affective Behavioral Neuroscience*. 2012. Vol. 12. P. 65–73.
36. *Klyucharev V., Hytonen K., Rijpkema M., Smidts A., Fernandez G.* Reinforcement learning signal predicts social conformity // *Neuron*. 2009. Vol. 61. P. 140–151.
37. *Klyucharev V., Munneke M.A.M., Smidts A., Fernandez G.* Downregulation of the posterior medial frontal cortex prevents social conformity // *Journal of Neuroscience*. 2011. Vol. 31. P. 11934–11940.
38. *Knutson B., Wimmer G.E.* Splitting the difference: how does the brain code reward episodes? // *Ann NY Acad Sci*. 2007. Vol. 1104. P. 54–69.
39. *Mason M.F., Dyer R., Norton M.I.* Neural mechanisms of social influence // *Organization Behavioral Human Development*. 2009. Vol. 110. P. 152–159
40. *Matsumoto M., Matsumoto K., Abe H., Tanaka K.* Medial prefrontal cell activity signaling prediction errors of action values // *Nature Neuroscience*. 2007. Vol. 10. P. 647–656.
41. *McClure S.M., York M.K., Montague P.R.* The neural substrates of reward processing in humans: the modern role of fMRI // *Neuroscientist*. 2004. Vol. 10. P. 260–268.
42. *Milgram S.* *Obedience to Authority: An Experimental View*. New York: Harper and Row, 1974. 256 p.
43. *Miltner W.H.R., Braun C.H., Coles M.G.H.* Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a “generic” neural system for error detection // *Journal of Cognitive Neuroscience*. 1997. Vol. 9. P. 788–798.
44. *Montague P.R., Lohrenz T.* To detect and correct: norm violations and their enforcement // *Neuron*. 2007. Vol. 56. P. 14–18.
45. *Nieuwenhuis S., Holroyd C.B., Mol N., Coles M.G.H.* Reinforcement-related brain potentials from medial



- frontal cortex: origins and functional significance // *Neuroscience and Biobehavioral Reviews*. 2004. Vol. 28. P. 441–448.
46. *Nieuwenhuis S., Slagter H.A., von Geusau N.J.A., Heslenfeld D.J., Holroyd C.B.* Knowing good from bad: differential activation of human cortical areas by positive and negative outcomes // *European Journal of Neuroscience*. 2005. Vol. 21. P. 3161–3168.
47. *Niv Y., Schoenbaum G.* Dialogues on prediction errors // *Trends Cognitive Sci.* 2008. Vol. 12. P. 265–272.
48. *O'Doherty J.P., Dayan P., Schultz J., Deichmann R., Friston K., Dolan R.J.* Dissociable roles of ventral and dorsal striatum in instrumental conditioning // *Science*. 2004. Vol. 304. P. 452–454.
49. *Parkinson J.A., Willoughby P.J., Robbins T.W., Everitt B.J.* Disconnection of the anterior cingulate cortex and nucleus accumbens core impairs Pavlovian approach behavior: further evidence for limbic cortical-ventral striatopallidal systems // *Behavioral Neuroscience*. 2000. Vol. 114. № 1. P. 42–63.
50. *Petty R., Cacioppo J.* Central and Peripheral Routes to Persuasion: An Individual Difference Perspective // *Journal of Personality and Social Psychology*. 1986. Vol. 51. № 5. P. 1032–1043.
51. *Petty R.E., Wegener D.T.* The Elaboration Likelihood Model: Current status and controversies. // *Dual process theories in social psychology* / Eds. S. Chaiken, Y. Trope. New York: Guilford Press, 1999. P. 41–72.
52. *Picard N., Strick P.L.* Motor areas of the medial wall: a review of their location and functional activation // *Cerebral Cortex*. 1996. Vol. 6. P. 342–353.
53. *Ridderinkhof K.R., Ullsperger M., Crone E.A., Nieuwenhuis S.* The role of the medial frontal cortex in Cognitive control // *Science*. 2004. Vol. 306. P. 443–447.
54. *Schultz W.* Behavioral theories and the neurophysiology of reward // *Annual Review of Psychology*. 2007. Vol. 57. P. 87–115.
55. *Shestakova A., Rieskamp J., Tugin S., Ossadtchi A., Krutitskaya J., Klucharev V.* Electrophysiological precursors of social conformity // *Journal of Social Cognitive and Affective Neuroscience*. 2013. Vol. 8. № 7. P. 756–763. doi: 10.1093/scan/nss064.
56. *Steele C.M., Southwick L.L., Critchlow B.* Dissonance and alcohol – drinking your troubles away // *Journal of Personality and Social Psychology*. 1981. Vol. 41. P. 831–846.
57. *Sutton R., Barto A.* Reinforcement learning: an introduction. Cambridge, MA: MIT Press, 1998. 551 p.
58. *Thorndike E.L.* Animal Intelligence: Experimental Studies. New York: Macmillan, 1911. 297 p.
59. *Tricomi E.M., Delgado M.R., Fiez J.A.* Modulation of caudate activity by action contingency // *Neuron*. 2004. Vol. 41. P. 281–292.
60. *Tucker D.M., Luu P., Frishkoff G., Quiring J., Poulsen C.* Frontolimbic response to negative feedback in clinical depression // *Journal of Abnormal Psychology*. 2003. Vol. 112. P. 667–678.
61. *Veen V. van, Krug M.K., Schooler J.W., Carter C.S.* Neural activity predicts attitude change in Cognitive dissonance // *Nature Neuroscience*. 2009. Vol. 12. P. 1469–1474.
62. *Walsh M., Anderson R.* Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioral choice // *Neuroscience and Biobehavioral Reviews*. 2012. Vol. 36. P. 1870–1884.
63. *Zaki J., Schirmer J., Mitchell J.P.* Social influence modulates the neural computation of value // *Psychological Science*. 2011. Vol. 22. P. 894–900.
64. *Zimbardo P.G., Weisenberg M., Firestone I.* Communicator effectiveness in producing public conformity and private attitude change // *Journal of Personality*. 1965. Vol. 33. P. 233–255.
65. *Zirnheld P.J., Carroll C.A., Kieffaber P.D., O'Donnell B.F., Shekhar A., Hetrick W.P.* Haloperidol impairs learning and error-related negativity in humans // *Journal Cognitive Neuroscience*. 2004. Vol. 16. № 6. P. 1098–1112.



# NEUROBIOLOGICAL MECHANISMS OF SOCIAL INFLUENCE

**KLYUCHAREV V.A.**\*, National Research University Higher School of Economics, Moscow, Russia; St. Petersburg State University, St. Petersburg, Russia; University of Basel, Switzerland,  
e-mail: vasily.klyucharev@unibas.ch

**ZUBAREV I.P.**\*\* , St. Petersburg State University, St. Petersburg, Russia,  
e-mail: ipzubarev@gmail.com

**SHESTAKOVA A.N.**\*\*\* , St. Petersburg State University, St. Petersburg, Russia,  
e-mail: anna.shestakova@helsinki.fi

Recently, the development of non-invasive methods of research of activity of the human brain shed new light on the neurobiological mechanisms of social influence on decision-making. This review presents modern neurobiological studies of social influence and discusses the key theoretical models of possible neurobiological mechanisms underlying social influence. Recent studies have demonstrated changes in the work of neurophysiological system of reinforcement learning and control of activities due to the influence of the social environment. Probably, the deviation of individual opinion from the group behavior (opinions) is interpreted by the nervous system as behavioral error or “reward prediction error”, which starts the process of behavior change, based on the dopaminergic mechanism of reinforcement learning.

**Keywords:** social influence, conformity, neuroeconomics, neuroscience, rostral cingulate zone, dopaminergic system, reward prediction error.

## Funding

This work was supported by the Federal Targeted Program “Kadry”, Activity 1.5 (Agreement № 8488), the Russian Foundation for Basic Research (grant 11-06-00449-a) and St. Petersburg State University Grant 0.37.522.2013.

## References

1. Ajzen I., Fishbein M. *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ, Prentice-Hall, 1980. 278 p.
2. Asch S.E. Opinions and social pressure. *Scientific American*, 1955, vol. 193, no. 5, pp. 31–35.

## For citation:

Klyucharev V.A., Zubarev I.P., Shestakova A.N. Neurobiological mechanisms of social influence. *Ekspierimental'naya psikhologiya = Experimental Psychology (Russia)*, 2014, vol. 7, no. 4, pp. 20–36 (In Russ., abstr. in Engl.).

\* *Klyucharev V.A.*, Professor, Cand. Sci. in Biology, National Research University Higher School of Economics, Moscow, Russia; St. Petersburg State University, St. Petersburg, Russia; University of Basel, Switzerland. E-mail: vasily.klyucharev@unibas.ch

\*\* *Zubarev I.P.*, St. Petersburg State University, St. Petersburg, Russia. E-mail: ipzubarev@gmail.com

\*\*\* *Shestakova A.N.* Ph.D., Senior Research Fellow, Chair of Higher Nervous Activity and Psychophysiology, Biological Faculty, St. Petersburg State University, St. Petersburg, Russia. E-mail: anna.shestakova@helsinki.fi



3. Behrens T.E., Hunt L.T., Rushworth M.F. The computation of social behavior. *Science*, 2009, vol. 324 (5931), pp. 1160–1164. doi: 10.1126/science.1169694
4. Behrens T.E., Hunt L.T., Woolrich M.W., Rushworth M.F. Associative learning of social value. *Nature*, 2008, vol. 456, pp. 245–249.
5. Berns G.S., Capra C.M., Moore S., Noussair C. Neural mechanisms of the influence of popularity on adolescent ratings of music. *Neuroimage*, 2010, vol. 49, pp. 2687–2696.
6. Berns G.S., Chappelow J., Zink C.F., Pagnoni G., Martin-Skurski M.E., Richards J. Neurobiological correlates of social conformity and independence during mental rotation. *Biological Psychiatry*, 2005, vol. 58, pp. 245–253.
7. Bryan J.H., Test M.A. Models and helping: naturalistic studies in aiding behavior. *Journal of Personality and Social Psychology*, 1967, vol. 6, pp. 400–407.
8. Campbell-Meiklejohn D.K., Bach D.R., Roepstorff A., Dolan R.J., Frith C.D. How the opinion of others affects our valuation of objects. *Current Biology*, 2010, vol. 20, pp. 1165–1170.
9. Campbell-Meiklejohn D.K., Simonsen A., Jensen M., Wohlert V., Gjerloff T., Scheel-Kruger J., Moller A., Frith C.D., Roepstorff A. Modulation of social influence by methylphenidate. *Neuropsychopharmacology*, 2012, vol. 37, no. 6, pp. 1517–1525.
10. Chen J., Wu Y., Tong G.Y., Guan X.M., Zhou X.L. ERP correlates of social conformity in a line judgment task. *BMC Neurosci*, 2012, vol. 13, pp. 13–43. doi: 10.1186/1471-2202-13-43
11. Chua H.F., Liberzon I., Welsh R.C., Strecher V.J. Neural correlates of message tailoring and self-relatedness in smoking cessation programming. *Biological Psychiatry*, 2009, vol. 65, pp. 165–168.
12. Cialdini R.B., Goldstein N.J. Social influence: compliance and conformity. *Annual Reviews in Psychology*, 2004, vol. 55, pp. 591–621.
13. Cohen M.X., Ranganath C. Reinforcement learning signals predict future decisions. *Journal of Neuroscience*, 2007, vol. 27, pp. 371–378.
14. Dawkins R. *The Selfish Gene*. New York, Oxford University Press, 1976. 512 p.
15. de Bruijn E.R.A., Sabbe B.G.C., Hulstijn W., Ruigt G.S.F., Verkes R.J. Effects of antipsychotic and antidepressant drugs on action monitoring in healthy volunteers. *Brain Research*, 2006, vol. 1105, pp. 122–129.
16. Elliott R., Newman J.L., Longe O.A., Deakin J.F.W. Instrumental responding for rewards is associated with enhanced neuronal response in subcortical reward systems. *NeuroImage*, 2004, vol. 21, pp. 984–990.
17. Falk E.B., Berkman E.T., Mann T., Harrison B., Lieberman M.D. Predicting persuasion-induced behavior change from the brain. *Journal of Neuroscience*, 2010, vol. 30, pp. 8421–8424.
18. Falk E.B., Berkman E.T., Whalen D., Lieberman M.D. Neural activity during health messaging predicts reductions in smoking above and beyond self-report. *Health Psychology*, 2011, vol. 30, pp. 177–185.
19. Fehr E., Fischbacher U. Social norms and human cooperation. *Trends Cognitive Sci*, 2004, vol. 8, pp. 185–190.
20. Festinger L., Carlsmith J.M. Cognitive consequences of forced compliance. *Journal Abnorm Psychology*, 1959, vol. 58 (39), pp. 203–210.
21. Gehring W.J., Goss B., Coles M.G.H., Meyer D.E., Donchin E. A neural system for error detection and compensation. *Psychological Science*, 1993, vol. 4, pp. 385–390.
22. Gehring W.J., Willoughby A.R. The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, 2002, vol. 295, pp. 2279–2282.
23. Glimcher P.W., Rustichini A. Neuroeconomics: the consilience of brain and decision. *Science*, 2002, vol. 306, pp. 447–452.
24. Hauber W., Sommer S. Prefrontostriatal circuitry regulates effort-related decision making, *Cerebral Cortex*, 2009, vol. 19, pp. 2240–2247.
25. Heider F. *The Psychology of Interpersonal Relations*. New York, Wiley, 1958, 322 p.
26. Hewig J., Trippe R., Hecht H., Coles M.G.H., Holroyd C.B., Miltner W.H.R. Decision-making in blackjack: an electrophysiological analysis. *Cerebral Cortex*, 2007, vol. 17, pp. 865–877.
27. Holroyd C.B., Coles M.G.H. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychology Reviews*, 2002, vol. 109, pp. 679–709.
28. Izuma K. The neural basis of social influence and attitude change. *Current Opinion in Neurobiology*, 2013, vol. 23. Available at: <http://dx.doi.org/10.1016/j.conb.2013.03.009>



29. Izuma K., Adolphs R. Social manipulation of preference in the human brain. *Neuron*, 2013, vol. 78, pp. 563–573. <http://dx.doi.org/10.1016/j.neuron.2013.03.023>
30. Izuma K., Matsumoto M., Murayama K., Samejima K., Sadato N., Matsumoto K. Neural correlates of Cognitive dissonance and choice-induced preference change. *Proceedings of National Academy of Sciences USA*, 2010, vol. 107, pp. 22014–22019.
31. Jocham G., Ullsperger M., Neuropharmacology of performance monitoring. *Neuroscience and Biobehavioral Reviews*, 2009, vol. 33, pp. 48–60.
32. Kenrick D.T., Li N.L., Butner J. Dynamical evolutionary psychology: Individual decision rules and emergent social norms. *Psychological Review*, 2003, vol. 110, pp. 3–28.
33. Kerns J.G., Cohen J.D., MacDonald A.W., Cho R.Y., Stenger V.A., Carter C.S. Anterior cingulate conflict monitoring and adjustments in control. *Science*, 2004, vol. 303, pp. 1023–1026.
34. Kim B.R., Liss A., Rao M., Singer Z., Compton R.J. Social deviance activates the brain's error-monitoring system. *Cognitive Affective Behavioral Neuroscience*, 2012, vol. 12, pp. 65–73.
35. Klucharev V., Hytonen K., Rijpkema M., Smidts A., Fernandez G. Reinforcement learning signal predicts social conformity. *Neuron*, 2009, vol. 61, pp. 140–151.
36. Klucharev V., Munneke M.A.M., Smidts A., Fernandez G. Downregulation of the posterior medial frontal cortex prevents social conformity. *Journal of Neuroscience*, 2011, vol. 31, pp. 11934–11940.
37. Klyucharev V.A., Shmids A., Shestakova A.N. Neuroekonomika: neurobiologiya prinyatiya reshenii [Neuroeconomics: neurobiology of decision making], *Eksperimental'naya psikhologiya [Experimental Psychology (Russia)]*, 2011, vol. 4, no. 2, pp. 14–35 (In Russ.; abstr. in Engl.).
38. Knutson B., Wimmer G.E. Splitting the difference: how does the brain code reward episodes? *Ann NY Acad Sci*, 2007, vol. 1104, pp. 54–69.
39. Mason M.F., Dyer R., Norton M.I. Neural mechanisms of social influence. *Organization Behavioral Human Devevelopment*, 2009, vol. 110, pp. 152–159
40. Matsumoto M., Matsumoto K., Abe H., Tanaka K. Medial prefrontal cell activity signaling prediction errors of action values. *Nature Neuroscience*, 2007, vol. 10, pp. 647–656.
41. McClure S.M., York M.K., Montague P.R. The neural substrates of reward processing in humans: the modern role of fMRI. *Neuroscientist*, 2004, vol. 10, pp. 260–268.
42. Milgram S., *Obedience to Authority: An Experimental View*. New York, Harper and Row, 1974. 256 p.
43. Miltner W.H.R., Braun C.H., Coles M.G.H. Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a “generic” neural system for error detection. *Journal of Cognitive Neuroscience*, 1997, vol. 9, pp. 788–798.
44. Montague P.R., Lohrenz T. To detect and correct: norm violations and their enforcement. *Neuron*, 2007, vol. 56, pp. 14–18.
45. Nieuwenhuis S., Holroyd C.B., Mol N., Coles M.G.H. Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance. *Neuroscience and Biobehavioral Reviews*, 2004, vol. 28, pp. 441–448.
46. Nieuwenhuis S., Slagter H.A., von Geusau N.J.A., Heslenfeld D.J., Holroyd C.B. Knowing good from bad: differential activation of human cortical areas by positive and negative outcomes. *European Journal of Neuroscience*, 2005, vol. 21, pp. 3161–3168.
47. Niv Y., Schoenbaum G. Dialogues on prediction errors. *Trends in Cognitive Sciences*, 2008, vol. 12, pp. 265–272.
48. O'Doherty J.P., Dayan P., Schultz J., Deichmann R., Friston K., Dolan R.J. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, 2004, vol. 304, pp. 452–454.
49. Parkinson J.A., Willoughby P.J., Robbins T.W., Everitt B.J. Disconnection of the anterior cingulate cortex and nucleus accumbens core impairs Pavlovian approach behavior: further evidence for limbic cortical-ventral striatopallidal systems. *Behavioral Neuroscience*, 2000, vol. 114, no. 1, pp. 42–63.
50. Petty R.E., Wegener D.T. The Elaboration Likelihood Model: Current status and controversies. In S. Chaiken, Y. Trope (eds.), *Dual process theories in social psychology*. New York: Guilford Press, 1999, pp. 41–72.
51. Petty R., Cacioppo J. Central and Peripheral Routes to Persuasion: An Individual Difference Perspective. *Journal of Personality and Social Psychology*, 1986, vol. 51, no. 5, pp. 1032–1043.



52. Picard N., Strick P.L. Motor areas of the medial wall: a review of their location and functional activation. *Cerebral Cortex*, 1996, vol. 6, pp. 342–353.
53. Ridderinkhof K.R., Ullsperger M., Crone E.A., Nieuwenhuis S. The role of the medial frontal cortex in Cognitive control. *Science*, 2004, vol. 306, pp. 443–447.
54. Schultz W. Behavioral theories and the neurophysiology of reward. *Annual Review of Psychology*, 2007, vol. 57, pp. 87–115.
55. Shestakova A., Rieskamp J., Tugin S., Ossadtschi A., Krutitskaya J., Klucharev V. Electrophysiological precursors of social conformity. *Journal of Social Cognitive and Affective Neuroscience*, 2013, vol. 8, no. 7, pp. 756–763. doi: 10.1093/scan/nss064.
56. Steele C.M., Southwick L.L., Critchlow B. Dissonance and alcohol – drinking your troubles away. *Journal Pers Soc Psychology*, 1981, vol. 41, pp. 831–846.
57. Sutton R., Barto A. *Reinforcement learning: an introduction*. Cambridge, MA, MIT Press, 1998. 551 p.
58. Thorndike E.L. *Animal Intelligence: Experimental Studies*. New York, Macmillan, 1911, 297 p.
59. Tricomi E.M., Delgado M.R., Fiez J.A. Modulation of caudate activity by action contingency. *Neuron*, 2004, vol. 41, pp. 281–292.
60. Tucker D.M., Luu P., Frishkoff G., Quiring J., Poulsen C. Frontolimbic response to negative feedback in clinical depression. *Journal of Abnormal Psychology*, 2003, vol. 112, pp. 667–678.
61. Veen V. van, Krug M.K., Schooler J.W., Carter C.S. Neural activity predicts attitude change in Cognitive dissonance. *Nature Neuroscience*, 2009, vol. 12, pp. 1469–1474.
62. Walsh M., Anderson R., Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience and Biobehavioral Reviews*, 2012, vol. 36, pp. 1870–1884.
63. Zaki J., Schirmer J., Mitchell J.P. Social influence modulates the neural computation of value. *Psychological Science*, 2011, vol. 22, pp. 894–900.
64. Zimbardo P.G., Weisenberg M., Firestone I. Communicator effectiveness in producing public conformity and private attitude change. *Journal of Personality*, 1965, vol. 33, pp. 233–255.
65. Zirnheld P.J., Carroll C.A., Kieffaber P.D., O'Donnell B.F., Shekhar A., Hetrick W.P. Haloperidol impairs learning and error-related negativity in humans. *Journal of Cognitive Neuroscience*, 2004, vol. 16, no. 6, pp. 1098–1112.