

ФИЛОСОФИЯ И НАУЧНОЕ ПОЗНАНИЕ

К.Ю. Федоров

ОБЪЕДИНЕНИЕ ТЕОРИЙ ПРОГНОСТИЧЕСКОЙ ОБРАБОТКИ И КОГНИТИВНОГО ДИССОНАНСА В КОНТЕКСТЕ ПРОБЛЕМЫ РАЦИОНАЛЬНОГО ВЫБОРА

Федоров Кирилл Юрьевич – аспирант. Московский государственный университет имени М.В. Ломоносова. Российская Федерация, 119991, г. Москва, Ленинские горы, д. 1; e-mail: kirill.fedorov.1@mail.ru

Существенной проблемой при построении теории рационального выбора (и рациональности вообще) на данный момент является вопрос объяснения иррациональности, точнее, того, что для внешнего наблюдателя агент может действовать иррационально, а модели, применяемые для него, могут оказаться нерабочими. Однако сам агент «внутри», если оказаться в его «логике», на поверку может действовать отчетливо рационально. Просто его «рациональность» и «рациональность» внешнего наблюдателя оказываются различны, поэтому не совпадают, что и приводит к несопадению языка описания ими друг друга. Для разрешения этой проблемы и нахождения теоретической основы для связи между физиологическими состояниями организмов и проверяемыми теоретическими построениями, а также получения таким образом возможности, не меняя оптики, работать со всеми агентами вне зависимости от распространенного у них типа рациональности, все чаще привлекаются психологические подходы. Для последних базовой является проблема «укорененности» в «материальном субстрате». Все это порождает вопрос, можно ли посредством инструментария существующих психологических теорий получить возможность построения единого языка описания разных типов рациональности? Данная работа направлена к тому, чтобы показать, какой теоретический механизм может позволить эту проблему решить. Представляется, что объединение теории прогностической обработки и теории когнитивного диссонанса оказывается крайне перспективным для данной задачи. Это обуславливается тем, что обе они занимают с уменьшением неожиданности, непоследовательности (ошибки прогнозирования и диссонанса), при этом первая предоставляет второй информативный контекст в рамках более широкой структуры восприятия–прогнозирования–действия (известной как активный вывод), а также помогает отделить диссонирующие когнитивные элементы от диссонирующих психологических состояний, рассматривая их при этом как часть одного и того же процесса восприятия и познания. Все это позволяет задать базовую теоретическую основу для связи между физиологическими состояниями организмов и проверяемыми теоретическими построениями о степени рациональности того или иного агента, и, таким образом, получить доступ ко внутренним информационным процессам в мозге для последующего анализа.

Ключевые слова: рационализация иррационального, теория прогностической обработки, когнитивный диссонанс, активный вывод, действия

Для цитирования: Федоров К.Ю. Объединение теорий прогностической обработки и когнитивного диссонанса в контексте проблемы рационального выбора // Философский журнал / Philosophy Journal. 2024. Т. 17. № 1. С. 180–196.

Следует отметить, что существует множество подходов к разрешению проблемы слабой рациональности. Представителем одного из наиболее известных является Даниэль Канеман. По его мнению, человек предстает нерациональным (не максимизирующим полезность) и имеющим две структурно и эволюционно различные системы для принятия решений: 1) бессознательная (быстрая, автоматическая, эмоциональная); 2) произвольная (медленная, целенаправленная, объективно рациональная)¹. По большому счету обе системы рациональны, но первая объективно (действующая по логике максимизации полезности, неосознанно для самого субъекта), а вторая субъективно (в смысле осознанная).

Одна из наиболее известных альтернатив представлена Г. Гигеренцером. Он утверждает, что человеческое мышление – это не нечто проинзанное иррациональными (случайными) когнитивными предубеждениями, а скорее рациональное, некий адаптивный инструмент, не идентичный правилам формальной логики или вероятностного исчисления. Им же предлагается понимать многие когнитивные ошибки как адаптивные реакции на мир неопределенности².

Теория прогностической обработки

В своих рассуждениях мы по большей части опираемся на теорию прогностической обработки (Predictive processing theory), считая ее наиболее полновесной и перспективной в деле изучения и описания особенностей работы человеческого мозга.

Согласно этой теории, мозг во многом походит на многоуровневую машину предсказаний, в которой нисходящий поток предсказаний (что мы ожидаем от мира) непрерывно сравнивается и корректируется относительно восходящего потока сенсорных данных (что воспринимают наши органы чувств). Этот процесс постоянного уточнения вероятностей можно описать, используя байесовскую статистику. Ее суть – в определении вероятности наступления события по предшествующим событиям. В случае, если нисходящие (top-down signals) предсказания не соответствуют восходящим (bottom-up signals) сенсорным данным, возникает так называемая ошибка предсказания (sensory prediction error), и модель либо обновляет свои предварительные установки (priors), либо сохраняет их и игнорирует входящие данные как шум. Это известно как многоуровневая иерархическая генеративная модель. Для теории прогностической обработки (далее – ПО) восприятие, познание и действие есть взаимосвязанные результаты процесса вывода. Этот

¹ Kahneman D. A perspective on judgment and choice: mapping bounded rationality // American Psychologist. 2003. Vol. 58. P. 697–720.

² Gigerenzer G. How to make cognitive illusions disappear: Beyond «heuristics and biases» // European Review of Social Psychology. 1991. Vol. 2 (1). P. 83–115.

процесс («активный вывод» по Фристон³) направлен на постепенное уменьшение количества сигналов об ошибках, получаемых организмом по ходу реализации своих собственных предсказаний. То есть, как отмечает Кларк⁴, эта теория говорит о том, что мозг не просто пассивный, управляемый стимулами инструмент, а наоборот, что «прогностический» мозг предсказывает, какие должны поступить сенсорные сигналы, основываясь на имеющихся в его распоряжении наилучших предыдущих моделях.

Вычислительная основа восприятия, познания и действия включает в себя всего три «базовых элемента»⁵:

1) предсказания (вытекающие из долгосрочной многоуровневой «генеративной модели»);

2) сигналы об ошибках предсказания (вычисляемые относительно активных предсказаний);

3) и оценочную, зависящую от контекста «точность» этих сигналов об ошибках предсказаний.

Ошибки с высокой точностью обладают большим постсинаптическим усилением и (следовательно) повышенным влиянием. И наоборот, даже сигнал с большой ошибкой предсказания, если ему присвоена крайне низкая точность, может оказаться системно бессильным, неспособным стимулировать обучение или дальнейшую обработку.

Иными словами, критически важным параметром для сигналов обоих потоков является уровень достоверности. То есть нас интересуют не только данные, но еще и их точность (precision), или вероятностный «вес».

Каждый вышестоящий уровень пытается предсказать поступающий сенсорный сигнал таким, каким он генерируется на расположенном ниже уровне, и делает это на основе закономерности ответов этого нижестоящего уровня. Хорошие примеры этого процесса приводит Кларк⁶. Так, в задаче обработки языка более высокие уровни могут специализироваться на предсказаниях, включающих знания на уровне слов, более низкие уровни – на предсказаниях, включающих знания на уровне букв, еще ниже – на предсказаниях, включающих знания на уровне штрихов, и т.д. В дальнейшем предположения с более высокого уровня используются для контекстуализации и приведения в соответствие с уже имеющейся информацией ответов с более низкого уровня, в то же время эти ответы формируют предположения более высокого уровня. Все это дает возможность высокоуровневым предположениям о том, «какое это слово», устранить двусмысленность на уровне букв, например, между плохо написанными «ev» и «w». Как показывает Кибель⁷, это может нам продемонстрировать, что высшие уровни начинают специализироваться на прогнозировании событий и отношений, «построенных» на основе характеристик и свойств, предсказанных более низкими уровнями.

³ Friston K., Daunizeau J., Kilner J., Kiebel S.J. Action and behavior: A free-energy formulation // *Biological Cybernetics*. 2010. Vol. 102 (3). P. 227–260.

⁴ Clark A. Radical predictive processing // *Southern Journal of Philosophy*. 2015. Vol. 53. Suppl. 1. P. 5.

⁵ Friston K., Daunizeau J., Kilner J., Kiebel S.J. Op. cit.

⁶ Clark A. A nice surprise? Predictive processing and the active pursuit of novelty // *Phenomenology and the Cognitive Sciences*. 2018. Vol. 17 (3). P. 521–534.

⁷ Kiebel S., Daunizeau J., Friston K. Perception and hierarchical dynamics // *Frontiers in Neuroinformatics*. 2009. Vol. 3 (20). P. 1–9.

За счет динамической оценки точности сигналов об ошибках в различных обстоятельствах возможно проводить изменения различных балансов как между обработкой в разных областях мозга (так как ПО является моделью, расположенной во множестве нейронных областей, взаимодействующих сложным, изменяющимся в контексте образом), так и между нисходящим прогнозом и поступающими сенсорными данными. Так, например, читая озадачивающий отрывок из инструкции по эксплуатации во время прослушивания музыки, мне может потребоваться посвятить свое внимание визуально представленным словам, временно игнорируя или снижая воздействие информации из слухового потока. В терминах ПО это означает увеличение точности оценок (а значит, и влияния) ошибок предсказаний, связанных с задачей чтения⁸. В попытке расслышать слова в шумной комнате, чтобы заполнить пробелы в поступающей информации, я вынужден в значительной степени полагаться на свои нисходящие прогнозы. А, например, рассматривание редкой монеты при ярком освещении подразумевает очень серьезное отношение к небольшим вариациям сенсорного сигнала, чтобы лучше распознать хорошо сделанную подделку⁹. В подобных случаях переменная оценка точности во взвешивании ошибки предсказания дает возможность корректировать баланс сил в зависимости от необходимости. А гибкая реализация множеством способов этого алгоритма взвешивания точности делает данную архитектуру удивительно гибкой и контекстно-зависимой¹⁰.

Все эти процессы практически мгновенны. Низшие уровни все время отправляют данные на высшие уровни, а те, в свою очередь, на основе этих данных перестраивают свои гипотезы и спускают к ним предсказания, основанные на них. Если происходит ошибка в этих предсказаниях, то производится либо замена гипотезы на высших уровнях, либо запрос переходит на следующий по высоте уровень.

Сама модель функционирования устанавливается в результате комбинации «эволюционного наследия» и непрерывного обучения на протяжении всей жизни. При этом изменчивость системы при обучении встроена в саму системную архитектуру (которая постоянно занимается прогнозированием), и потому улучшает ее. Мозг использует задачи по прогнозированию для прокладывания пути (корректируемого через обучение) к структурированному знанию о мире, используемому для создания более качественных и совершенных прогнозов.

В соответствии с логикой теории прогностической обработки, действия производятся при помощи тех же ресурсов. Идея¹¹ состоит в том, что у мозга есть два способа соотнести свои предсказания с окружающим миром: либо найти предсказание, наилучшим образом учитывающее текущий сенсорный сигнал (восприятие), либо изменить сенсорный сигнал в соответствии с предсказанием (действие). Кларк¹² приводит такой пример: «Если я предсказываю, что увижу свою кошку, и возникает ошибка, я могу использовать другое предсказание (например, «Я вижу экран компьютера»). Или я мог бы

⁸ Clark A. A nice surprise? P. 521–534.

⁹ Ibid.

¹⁰ Clark A. *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford, 2016.

¹¹ Friston K., Daunizeau J., Kilner J., Kiebel S.J. Op. cit.

¹² Clark A. A nice surprise? P. 523.

пошевелить головой и глазами так, чтобы увидеть кошку (которая, как оказалось, находится прямо здесь, рядом со мной, на столе)».

Действие здесь является результатом своего рода самоисполняющегося пророчества, в котором предсказание избранного подмножества сенсорных последствий служит для его осуществления. Обратим внимание, однако, что действие само по себе также предоставляет мощное средство проверки текущей гипотезы, например, устранения неоднозначности между равновероятными гипотезами¹³. В результате получается картина, в которой восприятие и действие являются взаимодополняющими проявлениями единого адаптивного режима, основным принципом работы которого является уменьшение характерной для организма ошибки прогнозирования.

Все это создает динамическую картину разума, в которой постоянно формируются генеративные, перспективные гипотезы, проверяемые сенсорной информацией, подтверждающей или обновляющей (в зависимости от веса, приданного тем или иным ошибкам прогнозирования) генеративную прогностическую модель, которые есть не что иное, как локально несовершенные, но в высшей степени обобщенные модели мира¹⁴.

Теория прогностической обработки отмечает, что снижение ошибки прогнозирования имеет важное значение для выживания по той причине, что, минимизировав эти ошибки, в том числе через попытки направить поведение, организмы оказываются лучше подготовленными к тому, чтобы делать выводы и узнавать о причинных структурах внешнего мира, избегать неожиданных состояний окружающей среды и столкновений, которые ставят под угрозу их выживание. Кларк¹⁵ пишет: «Таким образом, прогностический мозг – это не столько изолированный механизм вывода «в голове», сколько механизм взаимодействия, ориентированный на действие, обеспечивающий постоянное использование возможностей, связанных с выполнением задач». Так как люди живут в сложном социокультурном и материальном мире, механизмы вывода, подобные описываемым выше, должны быть способными адаптироваться к широкому разнообразию контекстов.

Неожиданность или энтропия – это функция ощущений агента и самого агента (модели). То есть неожиданность может появиться в отношении ожиданий, основанных на модели. Она обусловлена наблюдением и является атрибутом модели. Если модель дает наименьшую среднюю неожиданность для данной конкретной среды среди всех доступных моделей, она считается оптимальной.

Интересным для нас следствием является то, что согласно принципу свободной энергии¹⁶ мы собираем сенсорные сигналы, которые можем

¹³ Seth A. The Cybernetic Bayesian Brain: From Interoceptive Inference to Sensorimotor Contingencies // *Open MIND*. Fr. am M., 2015. P. 1–24; Friston K., Adams R.A., Perrinet L., Breakpear M. Perceptions as hypotheses: Saccades as experiments // *Frontiers in Psychology*. 2012. Vol. 3 (151). P. 1–20.

¹⁴ Clark A. Surfing Uncertainty.

¹⁵ Ibid. P. 295.

¹⁶ Этот термин зачастую используется как полный синоним ПО, и по сути, он является математической формулировкой того, как адаптивные системы (то есть биологические агенты, такие как животные или мозг) сопротивляются естественной тенденции к беспорядку (disorder) (то есть свободной энергии) (см.: Friston K. The free-energy principle: a unified brain theory? // *Nature reviews. Neuroscience*. 2010. No. 11. P. 127–138).

предсказать (см. теория эмуляции¹⁷), что гарантирует, что мы продолжим придерживаться этого в дальнейшем.

Неожиданность может быть уменьшена путем изменения сенсорного ввода (действия), предсказаний этого ввода (восприятия) или модели как таковой (т.е. путем эволюции, направленной на минимизацию свободной энергии или же «максимизацию» физической формы)¹⁸.

Эволюционная оптимизация модели, или иначе говоря, оптимизация, развивающая нейронную систему, отличается от восприятия и влечет за собой изменение формы и архитектуры агента. То есть среда способна обеспечить низкий уровень неожиданности для агента только в том случае, если этот агент был оптимизирован в соответствии с ней. Если агент был адаптирован, например, под насыщенную стимулами среду, то попав в малонасыщенную, он получит дозу неприятного удивления и постарается ее покинуть. В этом смысле каждый организм (от вирусов до людей) можно рассматривать как модель своей экологической ниши, оптимизированной для прогнозирования и отбора образцов из этой ниши¹⁹, а также как воплощенное решение проблемы минимизации свободной энергии, поддерживаемой окружающим его миром.

Если учесть, что «среднее значение свободной энергии за время» – это то, что в физике называется «действие»²⁰, то принцип свободной энергии оказывается уже хорошо знакомым принципом наименьшего действия, перенесенным на область теории информации.

Теория прогностической обработки и действия

Прогностическая обработка предполагает, что прогностическое восприятие должно быть внутренне согласованным, чтобы предоставлять четкие инструкции к действию. Кларк²¹ подчеркивает данный момент, используя пример с бинокулярным соперничеством. При демонстрации правому и левому глазу разных изображений (например, лицо и дом) субъекты обычно испытывают перцептивные переходы от одного восприятия (лицо) к другому (дом). Это чередование восприятий (вместо восприятия смешанного изображения «дом/лицо» (“house/face mash-up”) или, говоря иначе, «чередование системы между двумя полустабильными состояниями»²², как и предполагает ПО, происходит потому, что ни одна гипотеза не учитывает все данные, и само это чередование получает смысл, если целью является совершение действия. Если возможности (opportunities) для действий или доступность (affordances) «домов-лиц» (“face-houses”) неясны и ранее не учитывались (т.е. априорная вероятность «домов-лиц» низка), то очень мало практического смысла их воспринимать. Допустимость же определяется

¹⁷ Grush R. The emulation theory of representation: motor control, imagery, and perception // Behavioral and Brain Sciences. 2004. Vol. 27. P. 377–396.

¹⁸ Sella G., Hirsh A. The application of statistical physics to evolutionary biology // Proceedings of the National Academy of Sciences. 2005. Vol. 102. No. 27. P. 9541–9546.

¹⁹ Friston K., Thornton C., Clark A. Free-energy minimization and the dark-room problem // Frontiers in Psychology. 2012. Vol. 3 (130). P. 1–7.

²⁰ Ibid.

²¹ Clark A. Surfing Uncertainty. P. 33–37.

²² Ibid. P. 36.

нами как отношение между способностями воспринимать и действовать и особенностями окружающей среды, или, другими словами, допустимость действия обусловлена окружающей средой для соответствующего организма²³.

Прогностические гипотезы в первую очередь направляют действие, и ненужный диссонанс или ошибка между противоречивыми действиями в общем случае должны быть устранены, чтобы быть способными направлять поведение плавно и непрерывно²⁴.

В качестве подкрепления этой позиции можно привлечь «гипотезу конкуренции доступности» (“the affordance competition hypothesis”) Сайсека²⁵. В ней он утверждает, что, чтобы «выжить во враждебной среде, нужно быть готовым действовать в кратчайшие сроки, выполняя действия, которые, по крайней мере, частично подготовлены», что в свою очередь требует зависимости обработки «сенсорной информации от действий», чтобы построить представления о допустимых в данных условиях окружающей среды действиях. На основе наличествующей информации эти доступные варианты (или возможные, если мы говорим о действиях) сопоставляются друг с другом, пока не будет выбрано одно-единственное. Выбор же двух противоречивых действий был бы и неэффективным, и непрактичным. Таким образом, любые диссонирующие отношения между гипотезами о действии или в состояниях, подготавливающих к действию, должны быть устранены как можно быстрее.

Когнитивный диссонанс

Возникает вопрос, как может происходить «перекалибровка» описанных выше механизмов, и как туда могут проникать варианты действий, способствующие эффективному разрешению различных трудностей, и вместе с этим, как может проходить отбраковка уже неработающих вариантов действий? Нам представляется, что именно теория когнитивного диссонанса способна подсказать нам ответ на этот вопрос.

Л. Фестингер²⁶ в своей основополагающей работе «Теория когнитивного диссонанса» определяет когнитивный диссонанс как психологический дискомфорт, возникающий в случае, когда человек испытывает два взаимно несовместимых «элемента знания». Для Фестингера²⁷ когнитивные элементы – это то, «что индивид знает относительно самого себя, относительно своего поведения и относительно своего окружения». Эти когнитивные элементы не только «относятся к знанию самого себя» (чувств, желаний, установок и т.д.), но также и к состояниям окружающей среды, и к причинно-следственным связям в мире, в котором человек живет. Несмотря на то, что определение Фестингером когнитивных элементов является расплывчатым (если учесть, что свои изыскания он проводил на заре когнитивной

²³ Kaaronen R. A Theory of Predictive Dissonance: Predictive Processing Presents a New Take on Cognitive Dissonance // *Frontiers in Psychology*. 2018. Vol. 9 (2218). P. 1–15.

²⁴ Ibid.

²⁵ Cisek P. Cortical mechanisms of action selection: the affordance competition hypothesis // *Philosophical Transactions of the Royal Society. B: Biological Sciences*. 2007. No. 362. P. 1568.

²⁶ Фестингер Л. Теория когнитивного диссонанса. СПб., 1999.

²⁷ Там же. С. 25.

революции в психологии, то это не вызывает удивления), из него ясно, что отношение консонанса (гармонии) и диссонанса может сохраняться между разными парами этих элементов²⁸.

По мнению Фестингера²⁹, эти элементы познания образуют «индивидуальную карту действительности», фактически «зеркало», или «карту» реальности. Сохранение этой карты или зеркала реагирующими на реальность посредством активного уменьшения диссонанса помогает нам сохранять контроль над миром и выживать в нем. Таким образом, он³⁰ называет основным смыслом теории диссонанса то, что «реальность, которая воздействует на индивида, будет оказывать давление в направлении приведения когнитивных элементов в соответствие с этой реальностью». Фестингер³¹ подчеркивает, что когнитивный диссонанс «сам по себе является мотивирующим фактором», т.е., находясь в психологически некомфортно состоянии, агент будет стремиться уменьшить диссонанс и достигнуть консонанса (гармоничного состояния)³². Чтобы сделать это (по оригинальной концепции), есть четыре стратегии:

1. Изменение поведенческого когнитивного элемента.

Изменение поведения³³.

2. Изменение когнитивного элемента окружающей среды.

Изменения окружающей среды или перемещение в материальном или социальном мире³⁴.

3. Добавление новых когнитивных элементов.

Фестингер³⁵ пишет, что если диссонанс невозможно устранить, то его можно уменьшить, «добавив новые когнитивные элементы». Для этого применимы три взаимосвязанные подстратегии:

(1) Сосредоточение внимания на поддерживающих убеждениях.

Диссонанс можно уменьшить, сосредоточившись на поддерживающих убеждениях или уменьшив важность конфликтующих убеждений.

(2) Добавление примиряющих когнитивных элементов.

Диссонанс может быть уменьшен путем добавления когнитивных элементов или гипотез, которые «примиряют» два диссоциирующих когнитивных элемента.

(3) Изменение конфликтных когнитивных элементов.

Диссонанс может быть уменьшен даже произвольными изменениями в убеждениях, установках или ценностях.

4. Предотвращение диссонанса.

Помимо уменьшения диссонанса, Фестингер³⁶ подчеркивает, что существуют также «явно выраженные тенденции избегать увеличения диссонанса (или полностью предотвращать его появление)». Другими словами, иногда люди активно избегают ситуаций и информации, которые, вероятно, усилили бы диссонанс.

²⁸ Kaaronen R. Op. cit.

²⁹ Фестингер Л. Указ. соч. С. 26.

³⁰ Там же. С. 27.

³¹ Там же. С. 18.

³² Kaaronen R. Op. cit.

³³ Фестингер Л. Указ. соч. С. 36–37.

³⁴ Там же. С. 36–37.

³⁵ Там же. С. 40.

³⁶ Там же. С. 49.

Сходства и взаимодействие двух теорий

В этом разделе мы постараемся показать, что теория когнитивного диссонанса (далее КД) во многом является «производным» предполагаемых теорией прогностической обработки (ПО) особенностей устройства мозга. А так как диссонанс приводит к изменению поведения, он, по нашему представлению, оказывается механизмом воздействия на мозговые структуры (потоки). И признаком, проявлением этого является крайняя схожесть того, как уменьшается диссонанс в КД и как уменьшается ошибка прогнозирования в ПО.

Сходства (по Кааронену³⁷)

1) Изменение когнитивного элемента окружающей среды.

Как и в КД, ошибка прогнозирования в ПО также может быть уменьшена двумя различными, но взаимосвязанными способами, представляющими собой физическую модуляцию окружения и выборку сенсорных сред³⁸. Во-первых, как пишет Кларк³⁹: «Мы, люди... строим и неоднократно перестраиваем социальные, лингвистические и технологические миры, закономерности которых затем отражаются в генеративных моделях, делающих прогнозы». То есть у нас есть тенденция к созданию сред, которые уменьшают ошибку прогнозирования.

Во-вторых, как Фестингер⁴⁰ подчеркивает активную роль перемещения по социальным средам для уменьшения диссонанса, так и Кларк⁴¹ подчеркивает, что для уменьшения ошибки прогнозирования мы активно перемещаемся, чтобы «выборочно выбирать» воспринимаемый мир. То есть у нас прослеживается тенденция к избирательному сбору сенсорной информации, которая согласуется с нашими прогнозами (что является одной из причин такого явления, как предвзятость подтверждения⁴²).

2) Добавление новых когнитивных элементов.

Подобно предположениям из теории диссонанса, генеративные модели ПО активны в «поиске прогнозов, которые наилучшим образом соответствуют текущим сенсорным данным»⁴³.

3) Сосредоточение внимания на поддерживающих убеждениях.

В иерархической структуре ПО набор высокоуровневых прогнозов может объяснить ошибки прогнозирования и эффективно приказать блокам ошибок, поступающим снизу вверх, «заткнуться» (“shut up”)⁴⁴. Следовательно, входные данные более низкого уровня, которые хорошо предсказываются некоторой прогностической моделью более высокого уровня, «подавляются» (quashed) и «объясняются» (“explained away”)⁴⁵.

³⁷ Kaaronen R. Op. cit.

³⁸ Clark A. Surfing Uncertainty.

³⁹ Ibid. P. 270.

⁴⁰ Фестингер Л. Указ. соч. С. 38.

⁴¹ Clark A. Surfing Uncertainty. P. 70-71.

⁴² Ibid. P. 71.

⁴³ Ibid. P. 121.

⁴⁴ Friston K. A theory of cortical responses // Philosophical Transactions of the Royal Society B: Biological Sciences. 2005. Vol. 360. P. 829. Также: Clark A. Surfing Uncertainty. P. 38.

⁴⁵ Clark A. Surfing Uncertainty. P. 60.

То есть происходит добавление примиряющих когнитивных элементов и изменение конфликтных когнитивных элементов. Основной принцип ПО заключается в том, что ошибка прогнозирования может быть уменьшена путем подбора «новых и лучших» прогнозов или путем модуляции текущих прогнозов таким образом, чтобы они наилучшим образом соответствовали текущим сигналам ошибки^{46, 47}. Это в некотором смысле сопоставимо с представлением Фестингера о том, что когнитивные элементы могут быть добавлены или изменены для согласования диссонирующих элементов. На практике в прогностической иерархии это также может означать, что гипотезы «более высокого уровня» сохраняются путем корректировки или создания новых прогнозов на более низких уровнях. То есть это может включать, например, произвольные изменения отношений, часто связанные с определенными реакциями на КД⁴⁸.

4) Предотвращение диссонанса.

В теориях ПО и «активного вывода»⁴⁹ утверждается, что минимизация ошибки прогнозирования является движущей силой для организмов. Это имеет сходство с фестингеровской⁵⁰ идеей о том, что КД и его активное сокращение являются «самостоятельным мотивирующим фактором», в соответствии с которым следует активно избегать «ситуаций и информации», которые, вероятно, усилили бы диссонанс, чтобы наши когнитивные модели оставались реагирующими на реальность. Аналогичным образом Кларк⁵¹ пишет: «...весь аппарат (обработки на основе прогнозирования) существует только для того, чтобы помочь животным достичь своих целей, избегая фатально неожиданных встреч с миром». Эта идея хорошо совмещается с моделями КД, основанными на действиях.

Подытоживая все вышесказанное: основное обоснование переноса КД в рамки ПО вытекает из идеи Фестингера⁵²: «...реальность, которая воздействует на индивида, будет оказывать давление в направлении приведения когнитивных элементов в соответствие с этой реальностью». Вспомним критику Кларка⁵³ стандартной пассивной когнитивной модели, управляемой стимулами, которая принимает «энергетические сигналы от органов чувств» и «превращает их» в когерентное или согласованное восприятие. В таком случае КД, с точки зрения ПО, становится не несоответствием между двумя пассивными элементами познания, «затронутыми» реальностью (как предлагает Фестингер), а скорее несоответствием между активным нисходящим (ориентированным на действие) предсказанием и восходящими потоками перцептивной информации. Следовательно, уменьшение диссонанса может быть переформулировано как активное и ориентированное на перспективу уменьшение конфликтов между последующими прогнозами и исходными ошибками.

⁴⁶ Friston K. The free-energy principle. P. 127–138.

⁴⁷ Clark A. Surfing Uncertainty. P. 1.

⁴⁸ Kaaronen R. Op. cit.

⁴⁹ Например, см.: Friston K. The free-energy principle. P. 127–138; Clark A. Surfing Uncertainty.

⁵⁰ Фестингер Л. Указ. соч. С. 18.

⁵¹ Clark A. Surfing Uncertainty. P. 250, 293.

⁵² Фестингер Л. Указ. соч. С. 27.

⁵³ Clark A. Radical predictive processing. P. 5.

Взаимодействие

Теория прогностической обработки способна решить ряд трудностей, с которыми в свое время столкнулась теория когнитивного диссонанса. Так, Бовуа и Джоуль⁵⁴ отмечают, что первоначальная теория диссонанса Фестингера⁵⁵ вызывала впечатление, что «диссонирующие отношения» между познаниями всегда порождают «диссонирующее психологическое состояние», что, очевидно, не всегда соответствует действительности. Как нам представляется, иерархическая формулировка теории ПО может эту проблему эффективно решить. Напомним, что в прогностической иерархии все уровни от высоко абстрактных когнитивных представлений до сенсорных данных самого низкого порядка «подотчетны другим, обеспечивая внутренне согласованное представление сенсорных причин на нескольких уровнях описания»⁵⁶. Это говорит о том, что, хотя диссонансные отношения могут возникать на разных уровнях прогностической иерархии, начиная от непосредственных уровней восприятия («Идет дождь, хотя солнечно») и до более абстрактных уровней, вызываемый психологический дискомфорт из теории диссонанса будет являться тем, что находится на более высоком уровне этой иерархической модели.

Также, если КД понимается как одна из форм ошибки прогнозирования, то появляется цель ее уменьшить, для того чтобы обеспечить успешное поведение в нашей экологической нише, в которой присутствует множество конкурирующих возможностей. Создатель теории когнитивного диссонанса Фестингер, похоже, действительно имел это в виду, формулируя первоначальное определение КД (где уменьшение диссонанса заставляет нас поведенчески реагировать на реальность), но эта позиция, по-видимому, почти исчезла из более поздних исследований, которые все больше фокусировались на поведенческом диссонансе и меньше на действии-восприятии в реальном мире⁵⁷. Таким образом, теория прогностического диссонанса, похоже, согласуется с положением Хармон-Джонса и др.⁵⁸, что «процессы диссонанса в первую очередь функционируют для облегчения проведения эффективных действий».

Также стоит отметить, что хотя исследования диссонанса в основном и были сосредоточены на диссонансе между установками или ценностями⁵⁹, такая позиция есть следствие эмпирической методологии (так как выявить диссонанс с помощью интервью или опросов – это самый простой способ получить исследовательские данные⁶⁰) и не встречается в оригинальной теории Фестингера. И сам он, похоже, предполагал возможность более широкого применения своей теории.

Для многих исследователей когнитивного диссонанса представляется иррациональным присущее разрешению диссонанса отбрасывание признанной «лишней» информации. Во всех трех наиболее заметных парадигмах

⁵⁴ Beauvois J., Joule R. A Radical Dissonance Theory. London, 1996. P. xvii.

⁵⁵ Фестингер Л. Указ. соч.

⁵⁶ Friston K. The free-energy principle. P. 129.

⁵⁷ См.: Cooper J. Cognitive Dissonance: 50 Years of a Classic Theory. Los Angeles (CA), 2007.

⁵⁸ Harmon-Jones E., Harmon-Jones C., Levy N. An action-based model of cognitive-dissonance processes // Current Directions in Psychological Science. 2015. Vol. 24. P. 184–189.

⁵⁹ Cooper J. Op. cit.

⁶⁰ Подробнее см.: Ibid.

его исследования (индуцированное соответствие Фестингера и Карлсмита⁶¹, трудное решение Брема⁶² и оправдание усилий Тавриса и Аронсона⁶³) основное внимание уделяется таким, казалось бы, иррациональным моментам, как кажущиеся произвольными изменения отношений и систематическое избегание противоречивой информации. Однако в экологической нише, в которой мы обитаем, ограниченно рациональное, избирательно невежественное, но при этом быстрое и экономное, ориентированное на действие познание может быть наиболее эффективным способом передвижения и принятия решений⁶⁴. Хотя восприятие-действие такого рода не всегда локально оптимально, оно оказывается в высшей степени обобщаемо и в целом «удовлетворительно»⁶⁵. Это связано с тем, что чрезмерное обучение может привести к статистическому переобучению, т.е. обновлению байесовских априорных значений прогностической генеративной модели, так, что она настолько точно соответствовала бы одному набору локальных условий, что ее больше нельзя было бы обобщить на другие контексты.

К вопросу об объединении двух теорий

Таким образом, можно утверждать, что когнитивный диссонанс является одним из проявлений ПО. А предложенная Каароненом⁶⁶ теория прогностического диссонанса (predictive dissonance) является одним из вариантов такого объединения, и на наш взгляд, достаточно удачного. Помимо причин, указанных выше, положительную роль в этом играет и видение им когнитивного диссонанса как одной из форм «ошибки прогнозирования» (в рамках прогностической обработки). Тогда то, что аналогичные процессы обнаружены у других (также задействующих теорию прогностической обработки) организмов, совершенно не удивительно.

Дело в том, что несмотря на широкое распространение и известность КД, ее эволюционные истоки все еще плохо изучены⁶⁷. Однако в самом уменьшении диссонанса, как представляется, нет ничего контрэволюционного: «Напротив, механизмы уменьшения ошибок прогнозирования абсолютно необходимы для поддержания готовности к действию в среде с высокой условностью и неопределенностью, где консервативные реакции на диссонанс или ошибку прогнозирования также предотвращают перенастройку наших прогностических генеративных моделей к местным преходящим (ephemeral) условиям»⁶⁸.

⁶¹ Festinger L., Carlsmith J. Cognitive consequences of forced compliance // *Journal of Abnormal Psychology*. 1959. Vol. 58. P. 203–210.

⁶² Brehm J. Postdecision changes in the desirability of alternatives // *Journal of Abnormal Psychology*. 1956. Vol. 52 (3). P. 384–389.

⁶³ Tavris C., Aronson E. *Mistakes Were Made (but not by me): Why we Justify Foolish Beliefs, Bad Decisions, and Hurtful Acts*. Boston (MA), 2008.

⁶⁴ Clark A. *Surfing Uncertainty*. P. 1.

⁶⁵ Simon H. Rational choice and the structure of the environment // *Psychological Review*. 1956. Vol. 63 (2). P. 129–138; *Idem*. Theories of bounded rationality // *Decision and Organization*. 1972. Vol. 1. P. 161–176.

⁶⁶ Kaaronen R. *Op. cit.*

⁶⁷ Egan L., Santos L., Bloom P. The origins of cognitive dissonance: evidence from children and monkeys // *Psychological Science*. 2007. Vol. 18. P. 978.

⁶⁸ Kaaronen R. *Op. cit.*

Стоит также отметить, что отсутствие четкого определения когнитивных элементов и исходно гибкие формулировки в прописанной Фестингером теории открывают широкое пространство для интерпретаций, главным условием которых становится в таком случае то, чтобы они (интерпретации) вписывались в указанные и прописанные в теории КД рамки и подчинялись указанным закономерностям (либо как минимум не имели с ними существенных противоречий).

Мы предполагаем, что то, что понимается под «когнитивным элементом» в теории КД, на языке ПО есть результат взаимодействия «модели» (с высших уровней) и «данных» (с низших уровней), передающийся дальше по цепочке. То есть то, что в теории КД понимается под одним элементом, в теории ПО оказывается результатом взаимодействия двух составляющих, который уже сам на следующем уровне оказывается составляющей нового цикла сопоставления (следствием этого является и то, что диссонансы могут возникать множество раз, на разных уровнях, причем и из одного изначального набора моделей и данных, что согласуется и с выводом Кааронена). Мы можем воспользоваться данной интерпретацией потому, что КД не вступает в противоречие с понимаемыми таким образом элементами, теория не разрушается и для своего применения с понимаемыми так когнитивными элементами не требует существенной переработки, разве лишь стилистической. Это наше предположение подкрепляет и то, что и «данные», и «модели» из ПО подпадают под фестингеровское определение (указанное нами ранее) «элементов знания» (когнитивных элементов), противоречие между которыми и приводит к самому когнитивному диссонансу⁶⁹.

Все это делает возможным слияние, а точнее, фактически встраивание (поглощение) теории КД в ПО, приводя, таким образом, к обогащению и расширению инструментария последней. Гибкость же самой теории КД только способствует подобному ее применению.

Заключение

В данной работе мы постарались показать, как два теоретических подхода (ПО и КД), каждый из которых имеет свою развитую фактологическую базу (которую мы конкретно в данной статье не рассматриваем), взаимодействуют друг с другом. А также предположили, какими теоретическими следствиями этого взаимодействия мы можем воспользоваться в тех или иных задачах.

Подводя итог, мы можем заявить, что ПО и КД могут быть информативными и взаимодополняющими (вплоть до включения теории диссонанса в теорию прогностической обработки) друг для друга. Фактически при этом КД инкорпорируется, поглощается ПО, модернизируя ее. Это обуславливается тем, что обе эти теории занимают уменьшением неожиданности или последовательности (ошибки прогнозирования и диссонанса), при этом ПО способна включить КД в информативный контекст в рамках более широкой структуры восприятия – прогнозирования – действия (известной как активный вывод), а также помогает нам отделить диссонирующие когнитивные элементы от диссонирующих психологических состояний, рассматривая

⁶⁹ Фестингер Л. Указ. соч. С. 25.

их при этом как часть одного и того же процесса восприятия и познания⁷⁰. Таким образом, можно утверждать, что когнитивный диссонанс является одним из проявлений ПО. А предложенная Каароненом⁷¹ теория прогностического диссонанса является одним из вариантов такого объединения. Синтез этих двух теорий позволяет нам не только объяснить трудности, с которыми до этого сталкивалась теория КД, но и через нее (теорию КД) получить возможность приступить к разработке инструментов прямого и контролируемого взаимодействия с потоками данных в мозге, описываемых ПО. Причем учитывая, что диссонанс может происходить на любом уровне иерархии, мы можем получить способ взаимодействия соответственно с любым уровнем (в пределе), что в свою очередь открывает новые перспективы в вопросах о регуляции, рационализации и автономности поведения. По этой причине мы и можем утверждать, что механизм когнитивного диссонанса, рассматриваемого нами как один из основных (а на данный момент и самый доступный для изучения) механизмов «проникновения» и закрепления «вариантов эффективных действий» у отдельных индивидов, является частью теории прогностической обработки.

Все это позволяет задать базовую теоретическую основу для связи между физиологическими состояниями организмов и проверяемыми теоретическими построениями о степени рациональности того или иного агента, а также, в перспективе, проследить (посредством инструментов, основанных на КД) происхождение убеждений, влияющих на поведение (по которому мы только и можем судить о рациональности), и в пределе воздействовать на него посредством этого же механизма.

Список литературы

- Festinger L.* Теория когнитивного диссонанса / Пер. с англ. А. Анистратенко, И. Знаешева. СПб.: Ювента, 1999.
- Beauvois J., Joule R.* A Radical Dissonance Theory. London: Taylor & Francis, 1996.
- Brehm J.* Postdecision changes in the desirability of alternatives // *Journal of Abnormal Psychology*. 1956. Vol. 52 (3). P. 384–389.
- Cisek P.* Cortical mechanisms of action selection: the affordance competition hypothesis // *Philosophical Transactions of the Royal Society. B: Biological Sciences*. 2007. No. 362. P. 1585–1599.
- Clark A.* A nice surprise? Predictive processing and the active pursuit of novelty // *Phenomenology and the Cognitive Sciences*. 2018. Vol. 17 (3). P. 521–534.
- Clark A.* Radical predictive processing // *Southern Journal of Philosophy*. 2015. Vol. 53. Suppl. 1. P. 3–27.
- Clark A.* *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford: Oxford University Press, 2016.
- Cooper J.* *Cognitive Dissonance: 50 Years of a Classic Theory*. Los Angeles (CA): Sage, 2007.
- Egan L., Santos L., Bloom P.* The origins of cognitive dissonance: evidence from children and monkeys // *Psychological Science*. 2007. Vol. 18. P. 978–983.
- Festinger L., Carlsmith J.* Cognitive consequences of forced compliance // *Journal of Abnormal Psychology*. 1959. Vol. 58. P. 203–210.
- Friston K.* A theory of cortical responses // *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2005. Vol. 360. P. 815–836.

⁷⁰ Kaaronen R. Op. cit.

⁷¹ Ibid.

- Friston K.* The free-energy principle: a unified brain theory? // *Nature reviews. Neuroscience.* 2010. No. 11. P. 127–138.
- Friston K., Adams R.A., Perrinet L., Breakspear M.* Perceptions as hypotheses: Saccades as experiments // *Frontiers in Psychology.* 2012. Vol. 3 (151). P. 1–20. DOI: 10.3389/fpsyg.2012.00151.
- Friston K., Daunizeau J., Kilner J., Kiebel S.J.* Action and behavior: A free-energy formulation // *Biological Cybernetics.* 2010. Vol. 102 (3). P. 227–260.
- Friston K., Thornton C., Clark A.* Free-energy minimization and the dark-room problem // *Frontiers in Psychology.* 2012. Vol. 3 (130). P. 1–7. DOI: 10.3389/fpsyg.2012.00130.
- Gigerenzer G.* How to make cognitive illusions disappear: Beyond «heuristics and biases» // *European Review of Social Psychology.* 1991. Vol. 2 (1). P. 83–115.
- Grush R.* The emulation theory of representation: motor control, imagery, and perception // *Behavioral and Brain Sciences.* 2004. Vol. 27. P. 377–396.
- Harmon-Jones E., Harmon-Jones C., Levy N.* An action-based model of cognitive-dissonance processes // *Current Directions in Psychological Science.* 2015. Vol. 24. P. 184–189.
- Hohwy J., Roepstorff A., Friston K.* Predictive coding explains binocular rivalry: an epistemological review // *Cognition.* 2008. Vol. 108. P. 687–701.
- Kaaronen R.* A Theory of Predictive Dissonance: Predictive Processing Presents a New Take on Cognitive Dissonance // *Frontiers in Psychology.* 2018. Vol. 9 (2218). P. 1–15. DOI: 10.3389/fpsyg.2018.02218.
- Kahneman D.* A perspective on judgment and choice: mapping bounded rationality // *American Psychologist.* 2003. Vol. 58. P. 697–720.
- Kiebel S., Daunizeau J., Friston K.* Perception and hierarchical dynamics // *Frontiers in Neuroinformatics.* 2009. Vol. 3 (20). P. 1–9. DOI: 10.3389/neuro.11.020.2009.
- Sella G., Hirsh A.* The application of statistical physics to evolutionary biology // *Proceedings of the National Academy of Sciences.* 2005. Vol. 102. No. 27. P. 9541–9546.
- Seth A.* The Cybernetic Bayesian Brain: From Interoceptive Inference to Sensorimotor Contingencies // *Open MIND* / Ed. by T. Metzinger and J.M. Windt. Fr. am M.: MIND Group, 2015. P. 1–24. DOI: 10.15502/9783958570108.
- Simon H.* Rational choice and the structure of the environment // *Psychological Review.* 1956. Vol. 63 (2). P. 129–138.
- Simon H.* Theories of bounded rationality // *Decision and Organization.* 1972. Vol. 1. P. 161–176.
- Tavris C., Aronson E.* *Mistakes Were Made (but not by me): Why we Justify Foolish Beliefs, Bad Decisions, and Hurtful Acts.* Boston (MA): Houghton Mifflin Harcourt, 2008.

Integrating the predictive processing and the cognitive dissonance theories in the context of rational choice problem

Kirill Yu. Fedorov

Lomonosov Moscow State University. 1 Leninskie Gory, Moscow, 119991, Russian Federation; e-mail: kirill.fedorov.1@mail.ru

A significant problem in constructing the theory of rational choice (and rationality in general) at the moment is the question of explaining irrationality, or rather the fact that an agent can act irrationally for an external observer, and the models used for it may not work. However, the agent himself “inside”, if you find yourself in his “logic”, acts clearly rationally upon verification. It’s just that his “rationality” and the “rationality” of an external observer turn out to be different, hence they do not coincide, and hence this leads to a discrepancy in the language of their description of each other. To resolve this problem and find a theoretical basis for the connection between the physiological states of organisms and testable the theoretical constructs, and thus obtaining the opportunity, without changing the optics, to work with all agents, regardless of the types of rationality they have, psychological approaches have increasingly become involved. But for them, the basic

problem is always “rootedness” in the “material substrate”. This work largely suggests and shows which mechanism can allow this problem to be solved. The paper draws attention to the fact that the combination of the theory of predictive processing and the theory of cognitive dissonance turns out to be extremely promising for this task. This is due to the fact that both of these theories are engaged in reducing surprise or inconsistency (prediction errors and dissonance), with the former providing the latter with valuable and informative context within the broader perception – prediction – action framework (known as active inference) and also helps us to separate dissonant cognitive elements from dissonant psychological states, while considering them as part of the same process of perception and cognition. All this allows us to set a fundamental theoretical basis for the connection between the physiological states of organisms and testable theoretical constructs about the degree of rationality of a particular agent, and, thus, gain access to the internal information processes in the brain for study.

Keywords: rationalization of irrational, predictive processing theory, cognitive dissonance, active inference, actions

For citation: Fedorov, K.Yu. “Ob”edinenie teorii prognosticheskoi obrabotki i kognitivnogo dissonansa v kontekste problemy ratsional’nogo vybora” [Integrating the predictive processing and the cognitive dissonance theories in the context of rational choice problem], *Filosofskii zhurnal / Philosophy Journal*, 2024, Vol. 17, No. 1, pp. 180–196. (In Russian)

References

- Beauvois, J. & Joule, R. *A Radical Dissonance Theory*. London: Taylor & Francis, 1996.
- Brehm, J. “Postdecision changes in the desirability of alternatives”, *Journal of Abnormal Psychology*, 1956, Vol. 52 (3), pp. 384–389.
- Cisek, P. “Cortical mechanisms of action selection: the affordance competition hypothesis”, *Philosophical Transactions of the Royal Society. B: Biological Sciences*, 2007, No. 362, pp. 1585–1599.
- Clark, A. “A nice surprise? Predictive processing and the active pursuit of novelty”, *Phenomenology and the Cognitive Sciences*, 2018, Vol. 17 (3), pp. 521–534.
- Clark, A. “Radical predictive processing”, *Southern Journal of Philosophy*, 2015, Vol. 53, Suppl. 1, pp. 3–27.
- Clark, A. *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford: Oxford University Press, 2016.
- Cooper, J. *Cognitive Dissonance: 50 Years of a Classic Theory*. Los Angeles, CA: Sage, 2007.
- Egan, L., Santos, L. & Bloom, P. “The origins of cognitive dissonance: evidence from children and monkeys”, *Psychological Science*, 2007, Vol. 18, pp. 978–983.
- Festinger, L. *Teoriya kognitivnogo dissonansa* [A theory of cognitive dissonance], trans. by A. Anistratenko, I. Znaesheva. St. Petersburg: Yuventa Publ., 1999. (In Russian)
- Festinger, L. & Carlsmith, J. “Cognitive consequences of forced compliance”, *Journal of Abnormal Psychology*, 1959, Vol. 58, pp. 203–210.
- Friston, K. “A theory of cortical responses”, *Philosophical Transactions of the Royal Society B: Biological Sciences*, 2005, Vol. 360, pp. 815–836.
- Friston, K. “The free-energy principle: a unified brain theory?”, *Nature reviews. Neuroscience*, 2010, No. 11, pp. 127–138.
- Friston, K., Adams, R.A., Perrinet, L. & Breakspear, M. “Perceptions as hypotheses: Saccades as experiments”, *Frontiers in Psychology*, 2012, Vol. 3 (151), pp. 1–20, DOI: 10.3389/fpsyg.2012.00151.
- Friston, K., Daunizeau, J., Kilner, J. & Kiebel, S.J. “Action and behavior: A free-energy formulation”, *Biological Cybernetics*, 2010, Vol. 102 (3), pp. 227–260.
- Friston, K., Thornton, C. & Clark, A. “Free-energy minimization and the dark-room problem”, *Frontiers in Psychology*, 2012, Vol. 3 (130), pp. 1–7, DOI: 10.3389/fpsyg.2012.00130.

- Gigerenzer, G. "How to make cognitive illusions disappear: Beyond 'heuristics and biases'", *European Review of Social Psychology*, 1991, Vol. 2 (1), pp. 83–115.
- Grush, R. "The emulation theory of representation: motor control, imagery, and perception", *Behavioral and Brain Sciences*, 2004, Vol. 27, pp. 377–396.
- Harmon-Jones, E., Harmon-Jones, C. & Levy, N. "An action-based model of cognitive-dissonance processes", *Current Directions in Psychological Science*, 2015, Vol. 24, pp. 184–189.
- Hohwy, J., Roepstorff, A. & Friston, K. "Predictive coding explains binocular rivalry: an epistemological review", *Cognition*, 2008, Vol. 108, pp. 687–701.
- Kaaronen, R. "A Theory of Predictive Dissonance: Predictive Processing Presents a New Take on Cognitive Dissonance", *Frontiers in Psychology*, 2018, Vol. 9 (2218), pp. 1–15, DOI: 10.3389/fpsyg.2018.02218.
- Kahneman, D. "A perspective on judgment and choice: mapping bounded rationality", *American Psychologist*, 2003, Vol. 58, pp. 697–720.
- Kiebel, S., Daunizeau, J. & Friston, K. "Perception and hierarchical dynamics", *Frontiers in Neuroinformatics*, 2009, Vol. 3 (20), pp. 1–9, DOI: 10.3389/neuro.11.020.2009.
- Sella, G. & Hirsh, A. "The application of statistical physics to evolutionary biology", *Proceedings of the National Academy of Sciences*, 2005, Vol. 102, No. 27, pp. 9541–9546.
- Seth, A. "The Cybernetic Bayesian Brain: From Interoceptive Inference to Sensorimotor Contingencies", *Open MIND*, ed. by T. Metzinger and J.M. Windt. Fr. am M.: MIND Group, 2015, pp. 1–24, DOI: 10.15502/9783958570108.
- Simon, H. "Rational choice and the structure of the environment", *Psychological Review*, 1956, Vol. 63 (2), pp. 129–138.
- Simon, H. "Theories of bounded rationality", *Decision and Organization*, 1972, Vol. 1, pp. 161–176.
- Tavris, C. & Aronson, E. *Mistakes Were Made (but not by me): Why we Justify Foolish Beliefs, Bad Decisions, and Hurtful Acts*. Boston, MA: Houghton Mifflin Harcourt, 2008.