

М.А. Сущин

Предсказывающая обработка как теоретический комплекс в когнитивных науках

Сущин Михаил Александрович – кандидат философских наук, старший научный сотрудник. Институт научной информации по общественным наукам РАН. Российская Федерация, 117997, г. Москва, Нахимовский проспект, д. 51/21; e-mail: sushchin@bk.ru

Статья продолжает серию работ о концепции теоретических комплексов в когнитивных науках. Основные положения концепции теоретических комплексов используются для оценки теорий и моделей предсказывающей обработки в когнитивных науках. Предсказывающую обработку предлагается рассматривать как теоретический комплекс в когнитивных науках. Рассматриваются вопросы о структуре и функциях комплекса предсказывающей обработки. Выявляются основные положения этого направления и указывается, что предсказывающая обработка имеет однородную структуру. Показывается, что ключевые положения комплекса действительно обеспечивали дальнейшую пролиферацию предсказывающих теорий и моделей когнитивных и/или ментальных феноменов, как это и предполагается основными положениями концепции теоретических комплексов. Вместе с тем делается вывод, что к настоящему моменту в рамках комплекса предсказывающей обработки значимого теоретического прогресса достигнуто не было. Показывается также, что теоретический прогресс в когнитивных науках не стоит связывать преимущественно с ценностью унификации и что предпочтительная для когнитивных наук объяснительная стратегия (монистическая или плюралистическая) должна определиться в ходе эмпирических исследований.

Ключевые слова: теоретические комплексы, когнитивные науки, предсказывающая обработка, теоретический прогресс, достоинства научных теорий

Об идее теоретических комплексов в когнитивных науках

Данная статья продолжает серию публикаций, в которых была предложена концепция теоретических комплексов с целью надлежащего описания крупных теоретических направлений, или групп теорий, в когнитивных науках (таких как классический вычислительный когнитивизм, коннекционизм, «познание 4 Е» / «воплощенное познание» и др.) [Суцин 2021; Суцин, 2022б]. Проблема понимания крупных когнитивистских теоретических направлений представляет собой одну из важнейших проблем для философии современных когнитивных наук как подраздела философии науки. Крупные теоретические направления были в центре внимания когнитивистов-теоретиков на протяжении десятков лет, о них были написаны сотни статей и книг. Что именно они собой представляют?

Традиционно в философии науки доминировала установка анализа отдельных научных теорий. Безоговорочное доминирование традиционной установки было оспорено Т. Куном [Kuhn, 1996], взявшим в качестве единицы анализа не отдельную теорию, а нечто большее – то, что было названо им парадигмой. В схожем ключе развили свои концепции и такие авторы, как И. Лакатос [Lakatos, 1989] и Л. Лаудан [Laudan, 1978]. Именно к концепциям Куна, Лакатоса и Лаудана чаще всего и обращаются, когда в самом деле пытаются охарактеризовать направления, подобные коннекционизму или «воплощенному познанию» (следует отметить, такие попытки предпринимают весьма редко).

Между тем, как я утверждал в более ранних работах, указанные концепции сталкиваются с рядом трудностей, не позволяющих использовать ни одну из них как таковую в качестве основы для понимания когнитивистских комплексов теорий. Так, пожалуй, наиболее неудовлетворительным следствием концепции Куна, критиковавшимся многими авторами (включая К. Поппера, Д. Шэпира [Shapere, 1984] и др.), оказывается релятивизм, следующий из его тезиса о несоизмеримости парадигм. Проблематичность концепции Лакатоса связана с выдвигаемым им универсалистским критерием оценки научно-исследовательских программ – их способностью предсказывать новые факты. Так, даже в физике предсказание новых фактов отнюдь не всегда оказывается важнее объяснения уже имеющихся фактов (что, по-видимому, побудило Лакатоса изменить подход к оценке научно-исследовательских программ, добавив к более раннему предиктивистскому критерию способность давать непреднамеренные объяснения уже имеющимся фактам [Lakatos, Zahar, 1989]). В свою очередь, трудности концепции Лаудана связаны с поспешным отказом от идеи истины при оценке научного прогресса (на фоне краха попперовской концепции приближения к истине) и выдвиганием в качестве альтернативы критерия решения проблем, а также неясным онтологическим статусом предлагаемых им исследовательских традиций (ни одно утверждение из книги Лаудана «Прогресс и его проблемы» не исключает объективно идеалистические интерпретации исследовательских традиций).

Кроме того, в начале 1990-х гг. философ когнитивных наук Б. фон Эккардт выступила с концепцией «исследовательских концептуальных рамок» (research frameworks), являющейся дальнейшим развитием идей Куна и Лаудана

и специально предназначенной для описания ранних когнитивистских теоретических направлений – классического когнитивизма и коннекционизма [Von Eckardt, 1995]. И хотя отсутствие универсалистских амбиций, свойственных концепциям Куна, Лакатоса и Лаудана, позволяет концепции Фон Эккардт обойти ряд проблем, имеющих у ее предшественников, фатальным для ее представлений оказывается то обстоятельство, что в концепции не было предусмотрено возникновение теоретических направлений, не опирающихся на понятия репрезентации и обработки информации. По мнению Фон Эккардт, идеи репрезентации и обработки информации суть необходимые компоненты исследовательской концептуальной рамки когнитивной науки. Между тем, как сейчас известно, в когнитивных науках появились направления, никоим образом не обращающиеся к этим понятиям, – например, движение радикального воплощенного познания. Без коренного пересмотра концепция Фон Эккардт не в состоянии дать какое-либо описание этого направления.

Итак, я утверждаю, что ни одна из упомянутых концепций не может считаться достаточной для задачи надлежащего описания крупных комплексов теорий в когнитивных науках. Вместе с тем концепции Куна, Лакатоса, Лаудана и Фон Эккардт, безусловно, обладают рядом характеристик, которые хорошо стыкуются с теми или иными аспектами крупных совокупностей теорий в когнитивных науках. Таким образом, свою задачу я видел в том, чтобы разработать концепцию, которая учитывала бы позитивные наработки указанных концепций, но и была бы свободна от связанных с ними проблем. На эту роль я предлагаю концепцию теоретических комплексов в когнитивных науках, основное содержание которой будет сейчас кратко изложено.

Для начала введем два предварительных допущения. Первое допущение: подобно концепции Фон Эккардт идея теоретических комплексов не имеет универсалистских амбиций. Ее цель – дать адекватное описание крупных теоретических направлений в когнитивных науках. Вопрос о возможности использования концепции теоретических комплексов за пределами философии когнитивных наук следует адресовать специалистам в соответствующих разделах философии науки. Второе допущение призвано прояснить онтологический статус когнитивистских комплексов теорий и исключить возможность их объективно идеалистических интерпретаций. Его можно сформулировать так: комплексы теорий необходимо воплощены в отдельных научных теориях и моделях. Последние являются основными (но не единственными) носителями комплексов¹. Назовем данное допущение «номиналистическим допущением».

Два ключевых вопроса, на которые необходимо ответить, чтобы охарактеризовать крупные теоретические направления в когнитивных науках, есть вопросы об их структуре и функциях. Для ответа на эти два вопроса проведем различие между абстрактными и конкретными компонентами составляющих комплексы отдельных научных теорий и моделей. Абстрактные компоненты теорий, зачастую подразумеваемые имплицитно, определяют общее («грубое» и «эскизное», а не детальное) онтологическое видение предмета исследования: «познание есть X » (где X можно заменить на: «вычисление

¹ Для существования комплексов теорий требуются все три мира по Попперу.

на основе ментальных репрезентаций», «параллельную распределенную обработку информации», «воплощенное действие» и т.п.), «когнитивная система организована следующим образом...», «части когнитивных систем взаимодействуют между собой так...». Кроме того, они определяют общее видение методов исследования: «исследование когнитивных систем необходимо проводить в реальном мире, а не в стенах лаборатории» и т.п. Тогда как конкретные компоненты представляют собой любые утверждения, основывающиеся на базовых допущениях и направленные на их уточнение и развитие (в концептуальном и эмпирическом измерениях). Так, ими могут потенциально проверяемые утверждения формы «при условии X во временном интервале Y будет наблюдаться эффект Z » и т.д. и т.п.

При этом именно абстрактные компоненты, которые могут быть общими для ряда теорий, позволяют идентифицировать ту или иную группу теорий в качестве комплекса. Они формируют ядро теоретических комплексов. На этой основе можно выделить два типа структурной организации комплексов теорий в когнитивных науках (т.е. два типа отношения отдельных когнитивистских теорий в рамках комплекса друг к другу). Первый тип – это однородная структура. В этом случае все составляющие комплекс отдельные научные теории и модели разделяют ряд общих абстрактных онтологических и методологических положений. Такой тип структуры комплексов является наиболее распространенным (к нему следует отнести комплексы классического когнитивизма, коннекционизма, а также комплекс предсказывающей обработки, о котором пойдет речь далее). Второй тип структурной организации – это неоднородная структура. Теории и модели комплекса с неоднородной структурой соотносятся друг с другом на манер семейных сходств по Л. Витгенштейну, т.е. на основе положений, пересекающихся между собой лишь частично. В первой работе, посвященной концепции теоретических комплексов, я выдвинул и обосновывал тезис, что неоднородную структуру имеет комплекс умеренного (т.е. не отрицающего участие репрезентаций в познании) воплощенного познания [Суцин, 2021, с. 13–15].

С помощью того же различения между абстрактными и конкретными компонентами отдельных теорий и моделей можно дать ответ и на вопрос о преимущественных функциях теоретических комплексов в когнитивных науках. В первой функции комплекс (в первую очередь в части его базовых онтологических и методологических положений) выступает в качестве своеобразного ориентира при создании новых, а также модификации уже имеющихся теорий и моделей (он определяет проблемное поле, понятийный аппарат и методологию исследования). Вторая ключевая функция имеет негативную окраску: в этой ипостаси комплекс рассматривается сторонниками других теоретических направлений в качестве отрицательного примера при создании и/или модификации теорий и моделей (именно таким образом комплекс классического когнитивизма рассматривался сторонниками коннекционизма и воплощенного познания). Впрочем, и в этой роли комплекс в определенной мере способствует созданию новых теорий, но уже в рамках альтернативных направлений, указывая их приверженцам, что было упущено и какие проблемы необходимо решить, какие методы при этом нужно использовать и на какие понятия

опираться и т.д. Таким образом, в обеих своих главных функциях комплексы способствуют созданию новых теорий и моделей. Компактно это можно выразить так: теоретические комплексы содействуют пролиферации научных теорий и моделей. Они представляют собой эффективный механизм пролиферации.

Необходимо особо подчеркнуть, что пролиферация как таковая, с этой точки зрения, не является самоцелью. Пролиферация и плюрализм теорий, по моему мнению, благодаря совершенствованию теорий через взаимную критику и сохранению альтернатив (на тот случай, если имеющиеся теории столкнутся с серьезными трудностями) призваны² содействовать прогрессу в отношении ряда так называемых достоинств научных теорий (также именуемых эпистемическими ценностями). К их числу обычно относят [McMullin, 2008]: (1) предсказательный успех, (2) способность давать непреднамеренные объяснения имеющимся значимым фактам, (3) способность объединять разрозненные предметные области (если данные области фундаментальным образом связаны между собой³), (4) способность успешно проходить серии экспериментальных испытаний, (5) способность объяснять уже имеющиеся данные (эмпирическая адекватность), (6) простоту конструкций теории, (7) широту предлагаемых объяснений, (8) способность предсказания новых непредвиденных фактов (которые, впрочем, еще не нашли подтверждения). (Первые четыре достоинства более труднодостижимы, прогресс в отношении этой группы достоинств связывают со способностью теории аппроксимировать истину.)

Соответственно, следует проводить различие между неумеренной пролиферацией (и неумеренным плюрализмом), не имеющей никакой другой цели, кроме создания новых теорий, увеличения их количества, и конструктивной пролиферацией (а также конструктивным плюрализмом), призванной содействовать теоретическому прогрессу.

В чисто количественном же плане для образования комплекса требуются всего-навсего две (сходные по своим центральным онтологическим и методологическим положениям) теории. С использованием так называемой формы записи множеств (set builder notation) это можно выразить так:

$$C = \{t_1, \dots, t_n\}, |C| \geq 2.$$

Итак, теоретические комплексы представляют собой группу (т.е. не менее двух) теорий, объединяемых на основе схожих абстрактных онтологических и методологических положений и имеющих однородную или неоднородную структуру, функционирующих для обеспечения пролиферации научных теорий и моделей. Я добавляю к этому нормативное методологическое утверждение, а именно что обеспечиваемая комплексами пролиферация должна в конечном счете способствовать прогрессу в отношении ряда достоинств научных теорий и теоретическому прогрессу как таковому. (Подробнее об отличиях предлагаемой концепции теоретических комплексов от близких ей концепций Лакатоса и Лаудана – см. таблицу.)

² Это требование носит не дескриптивный, а нормативный методологический характер.

³ Проблемы безоговорочного акцентирования ценности унификации в области когнитивных наук будут подробно обсуждаться ближе к концу статьи.

Таблица. Основные отличия концепции теоретических комплексов от концепции научно-исследовательских программ И. Лакатоса и концепции исследовательских традиций Л. Лаудана

	Положения из состава концептуального ядра	Отдельные научные теории и модели	Эпистемическая аксиология (в ее связи с постулируемыми функциями программ, традиций, комплексов, соответственно)
Научно-исследовательские программы по Лакатосу [1978]	Неизменяемые и однородно структурированные	Формируют последовательную цепочку версий научно-исследовательской программы	Основной акцент на ценности предсказательного успеха, т.е. монизм эпистемических ценностей
Исследовательские традиции по Лаудану [1977]	Могут изменяться радикально, однородно структурированные	Безусловный приоритет традиций по отношению к отдельным теориям, при этом последние могут переходить из одной традиции в другую	Основной акцент на ценности способности решения проблем, т.е. монизм эпистемических ценностей
Теоретические комплексы в предложенном смысле [Суцин, 2022]	Изменяемые нерадикально, могут структурироваться как в однородной, так и в неоднородной манере	Не формируют единообразной цепочки и не могут переходить из одного комплекса в другой	Акцент на необходимости достижения прогресса в отношении множества эпистемических ценностей, т.е. плюрализм эпистемических ценностей

Комплекс предсказывающей обработки: структура и функции

Теперь, имея в своем распоряжении идею теоретических комплексов в когнитивных науках, можно перейти к анализу направления, получившего известность как «предсказывающая обработка» (predictive processing). Данное направление получило широкое распространение и обсуждение в когнитивных науках в последние полтора десятилетия и уже может рассматриваться в качестве одного из основных когнитивистских теоретических направлений наряду с классическим вычислительным когнитивизмом, коннекционизмом и воплощенным познанием. К его предшественникам относят Альхазена (Ибн аль-Хайсама), И. Канта и в особенности Г. Фон Гельмгольца с его идеей бессознательных умозаключений. Изначально предложенные для понимания

восприятия и действия идеи предсказывающей обработки были расширены для понимания множества когнитивных и ментальных феноменов: эмоций, сознания, перцептивных иллюзий и др. Сторонники этого направления высказывают смелые и оптимистичные утверждения, что «предсказывающие модели подходят чрезвычайно близко к преодолению некоторых основных трудностей, препятствовавших предшествующим попыткам создать объединенную теорию разума, мозга и действия» [Clark, 2013, p. 200]. Оценим эти утверждения.

Итак, если предсказывающую обработку предлагается рассматривать как теоретический комплекс, то в первую очередь необходимо выявить ключевые положения данного комплекса и ответить на вопрос о его структуре. С этой целью я буду опираться на работы трех авторов, оказавших наибольшее влияние на его развитие: нейробиолога К. Фристана [Friston et al., 2018], философов Я. Хохви [Hohwy, 2013; Hohwy, 2020] и Э. Кларка [Clark, 2013]. К примеру, в сжатом виде основные идеи этого направления приводятся в недавней статье Хохви [Hohwy, 2020].

Так, пожалуй, (1) первой ключевой идеей предсказывающей обработки оказывается идея обновляемой внутренней модели мира – точнее говоря, модели подлежащих установлению причин сенсорного входа (стимулов). Внутренняя модель мира не имеет прямого доступа к этим стимулам, но может генерировать (2) предсказания (гипотезы) касаясь возможных причин сенсорного входа. Предсказания представляют собой сообщения (сигналы), передающиеся по слоям (3) иерархически организованной нервной системы, на самом верхнем уровне которой реализована внутренняя модель мира, а на самом нижнем – осуществляется регистрация непосредственно сенсорных данных. Каждый уровень этой системы может взаимодействовать только с соседними уровнями (в случае промежуточных уровней только с одним уровнем выше и одним уровнем ниже). Предсказания передаются от модели «сверху вниз» и сопоставляются с сенсорными данными на нижнем уровне системы. Если имеется несоответствие, так называемая (4) ошибка в предсказании (prediction error), определяемая как

$$\text{ошибка} = \text{сенсорный вход} - \text{предсказание},$$

то соответствующий сигнал передается вверх по иерархии для обновления внутренней модели мира (понятно, что если значения для сенсорного входа и предсказания равны, то, согласно уравнению выше, вычитая второе из первого, ошибка будет равна нулю). Обновленная внутренняя модель далее генерирует новое предсказание возможных причин сенсорного входа. Таким образом, уравнение для нового предсказания будет выглядеть следующим образом:

$$\text{новое предсказание} = \text{старое предсказание} + \text{ошибка}.$$

При этом предполагается, что ошибка в предсказании принимается в расчет также и на основе (5) оценки интенсивности обучения системы: если интенсивность была высокой (с увеличением коэффициента π_L , указывающего на информативность текущего сенсорного входа), то ошибке придается более значимый вес. Если же интенсивность была невысокой (с увеличением

коэффициента π_p , описывающего, сколько системе уже известно), то и ошибке не придается большого веса:

$$\text{новое предсказание} = \text{старое предсказание} \left(\frac{\pi_L}{[\pi_p + \pi_L]} \times \text{ошибка} \right).$$

По мере того как внутренняя модель системы будет становиться все более аккуратной, в долгосрочной перспективе система будет совершать все меньше ошибок в предсказании. По мнению Хохви, если система оказывается в состоянии минимизировать ошибки в предсказании на протяжении определенного времени, то на этой основе можно заключить, что данная система

...аппроксимирует Байесовский вывод. Таким образом, ПО [предсказывающая обработка] фокусируется на системах, минимизирующих ошибку в предсказании и, соответствующим образом, аппроксимирующих [эмпирический] Байесовский вывод... Это важно по той причине, что легче представить (биологически правдоподобные) системы, минимизирующие ошибку в предсказании и, соответствующим образом, аппроксимирующие Байесовский вывод, нежели системы, реализующие точный Байесовский вывод [Hohwy, 2020, p. 3].

Приведенные выше положения формируют достаточно аккуратное представление о центральных допущениях комплекса предсказывающей обработки. Соответственно, можно говорить о том, что данный комплекс имеет однородную структуру: относящиеся к предсказывающей обработке теории и модели эксплицитно или имплицитно разделяют указанные допущения. При этом необходимо отметить, что обозначенные допущения формируют лишь общее представление об исследуемой предметной области и не формируют детальной онтологической картины. В отношении более детальной онтологии среди сторонников предсказывающей обработки имеются расхождения. Так, с одной стороны, Хохви трактует указанные допущения в интерналистском, нейрентрическом ключе: «Нет необходимости, так сказать, каким-то невероятным образом выпрыгивать за пределы черепной коробки и сравнивать перцепты с причинными состояниями мира. Эти причины остаются скрытыми за вуалью сенсорного входа и могут быть определены только с помощью вывода» [Hohwy, 2013, p. 50]. Хохви полагает, что перспективы предсказывающей обработки и воплощенного познания могут быть совместимы, но с учетом акцента на скрытость мира за вуалью сенсорного входа. С другой стороны, Э. Кларк утверждает, что идеи предсказывающей обработки естественным образом сочетаются и дополняют представления о познании, развитые в рамках комплекса воплощенного познания. Пафос и мотивация последнего во многом состояли в критике репрезентационных и вычислительных подходов к пониманию познавательных процессов (в первую очередь классического когнитивизма), критике идеи чрезмерно детализированных внутренних репрезентаций и моделей мира. Тем не менее, по мнению Кларка, предсказывающая обработка является идеальным объяснительным партнером для «недавних подходов, подчеркивающих воплощенные, определенные средой (environmentally embedded), измерения психики (mind) и разума (reason). Если эти подходы верны, восприятие, действие и внимание все вместе вовлечены в одно и то же общее дело:

минимизацию сенсорных ошибок в предсказании, являющихся результатом нашего взаимодействия со средой» [Clark, 2013, p. 201]. Как бы там ни было, в настоящий момент эта дискуссия не затрагивает непосредственно центральные допущения комплекса предсказывающей обработки. Как отмечалось ранее, для функционирования комплекса как такового достаточно, чтобы положения из состава концептуального ядра формировали лишь общее представление об исследуемой предметной области.

Ответив на вопрос о структуре предсказывающей обработки как теоретического комплекса в когнитивных науках, мы можем перейти ко второму ключевому вопросу, а именно: обеспечивал ли этот комплекс (в первую очередь при помощи допущений из состава концептуального ядра) пролиферацию теорий и моделей? То есть функционировал ли он так, как это предполагается в рамках концепции теоретических комплексов? На этот вопрос однозначно стоит дать утвердительный ответ. Совершенно нетрудно увидеть, как ключевые идеи предсказывающей обработки, развитые в первую очередь работами К. Фристана, Я. Хохви и Э. Кларка, стали использоваться для создания дальнейших теорий и моделей комплекса.

Как отмечалось выше, первые теории и модели комплекса предсказывающей обработки были развиты для объяснения восприятия и действия. В этом плане для предсказывающей обработки важнейшую роль играет не только идея устранения ошибок и обновления модели мира в рамках перцептивного вывода, но и понятие активного вывода (*active inference*). Последнее означает, что ошибки в предсказании могут быть минимизированы также и с помощью действий, когда система осуществляет двигательную активность, чтобы устранить имеющиеся ошибки (скажем, некто, не найдя чашки под рукой, идет на кухню и находит ее там). Эти два понятия тесно связаны друг с другом и формируют то, что Хохви называет «арсеналом предсказывающей обработки» (*PP toolbox, predictive processing toolbox*) [Hohwy, 2020, p. 4]. Между тем за последние полтора десятилетия было предпринято множество попыток расширить ключевые идеи комплекса предсказывающей обработки для понимания широкого круга когнитивных (и ментальных) феноменов и способностей, включая бинокулярное соперничество, Я и самосознание, эмоции, внимание, сознание и др. (О разнообразии областей, в которых были выдвинуты соответствующие теории и модели предсказывающей обработки, можно судить по вспомогательной таблице из работы Хохви.)

Так, в совместной работе Я. Хохви, А. Ропсторффа и К. Фристана была предпринята попытка объяснения давно известного феномена бинокулярного соперничества (попеременной смены перцептов при восприятии неоднозначных стимулов наподобие полупрозрачных изображений, наложенных друг на друга) в терминах предсказывающего кодирования [Hohwy et al., 2008]. С этой точки зрения, выбранная мозгом гипотеза в данный момент может объяснить только часть неоднозначного стимула, предъявленного субъекту, – допустим, полупрозрачное изображение лица. Вторая же часть изображения (скажем, полупрозрачное изображение дома) в это время посылает сигнал об ошибке, в результате чего мозг оставляет предшествующую гипотезу (гипотезу лица, присутствующего на картинке) и выбирает новую гипотезу

в качестве создаваемого стимула. Через некоторое время ситуация разворачивается в обратную сторону и т.д.

А. Сетом была предложена предсказывающая модель interoцепции и эмоций, в соответствии с которой эмоции следует понимать как результат предсказаний причин interoцептивных афферентных (идущих от периферии к мозгу) сигналов [Seth, 2013]. С позиций комплекса предсказывающей обработки было развито несколько альтернативных моделей Я, были предложены модели мультисенсорной интеграции (т.е. интеграции мозгом информации, поступающей от разных органов чувств), чувства собственного тела, были сформулированы достаточно интересные объяснения ряда психических расстройств. Была предпринята попытка связать феномен внимания с упомянутым выше механизмом уделения веса сигналам об ошибке в предсказании в зависимости от интенсивности предшествующего обучения⁴.

Сейчас нет необходимости делать подробный общий обзор множества работ, которые были предложены для понимания самых разнообразных когнитивных и ментальных феноменов с позиции основных принципов предсказывающей обработки. Сказанного вполне достаточно, чтобы понять, что данный комплекс в самом деле выступал и продолжает выступать в качестве основы для создания новых теорий и моделей когнитивных и/или ментальных феноменов.

Проблемы комплекса предсказывающей обработки

Поскольку в рамках концепции теоретических комплексов подчеркивается, что пролиферация научных теорий должна способствовать прогрессу в отношении ряда достоинств научных теорий, тотчас же встает вопрос о прогрессивности комплекса предсказывающей обработки. Только что мы пришли к выводу, что данный комплекс в самом деле обеспечивал дальнейшую пролиферацию теорий и моделей. Однако привела ли пролиферация теорий и моделей предсказывающей обработки к прогрессу в отношении таких достоинств научных теорий, как доказанная предсказывающая плодотворность (подтверждение в опытах предсказаний новых неожиданных фактов), способность давать непреднамеренные объяснения имеющимся значимым фактам и т.п.? На этот вопрос можно ответить следующим образом: несмотря на то, что сторонниками комплекса предсказывающей обработки периодически высказываются утверждения о выдвигании новых эмпирических предсказаний, значимого теоретического прогресса (прогресса в отношении первых четырех упомянутых в начале статьи достоинств научных теорий) к настоящему моменту достигнуто не было.

Смелые утверждения сторонников комплекса предсказывающей обработки о том, что с помощью их теорий были сформулированы новые эмпирические предсказания, навлекли на себя критику со стороны философа когнитивных наук М. Милковского. В нескольких недавних работах Милковский и его соавтор П. Литвин осуществили достаточно подробное критическое рассмотрение

⁴ См. таблицу, прилагаемую к работе Я. Хохви: [Hohwy, 2020].

недавних теорий и моделей предсказывающей обработки [Litwin, Miłkowski, 2020; Miłkowski, Litwin, 2022]. В адрес последних было высказано несколько критических замечаний. Во-первых, по замечанию Милковского, сторонники предсказывающей обработки часто не вполне обоснованно переносят формальные понятия (вроде понятия «точности» – precision) в область описания когнитивных феноменов наподобие внимания и т.п. Так, доверие, уверенность, заостренное внимание –

...все связываются с «усиленной точностью» (enhanced precision): однако подобного рода попытки охватить различные психологические феномены с помощью одного-единственного вычислительного термина не могут привести к информативным объяснениям... Если убеждения, определенность, уверенность, ожидания или способность строить планы на будущее предполагается объяснять в терминах ПО [предсказывающей обработки], модели соответствующих феноменов, отличающие их от других аспектов нашей ментальной жизни, должны быть предоставлены. Нельзя полагаться только на интуитивные аргументы, основывающиеся на простой терминологической близости... [Litwin, Miłkowski, 2020, p. 7].

Во-вторых, многие модели предсказывающей обработки оказываются столь размытыми, что под них можно подвести практически любые эмпирические данные (похожая критика ранее высказывалась в отношении байесовских моделей в психологии и нейронауке [Bowers, Davis, 2012]). Соответственно, то, что сторонники предсказывающей обработки называют новыми эмпирическими предсказаниями, на поверку может оказаться, к примеру, открытыми исследовательскими вопросами.

Корень проблем предсказывающей обработки Милковский видит в нефальсифицируемости теорий и моделей предсказывающего кодирования. Милковский, как и некоторые другие современные когнитивисты, в особенности подчеркивает значимость унификации в науке (как достоинства научных теорий, или эпистемической ценности). Он утверждает, что для унифицированной теории чрезвычайно важно оставаться единой, однородной теорией – такая теория не должна быть монструозной конъюнкцией не связанных между собой объяснительных паттернов:

Научная репрезентация является монструозной, когда она содержит части, которые невозможно опровергнуть при помощи единственного факта, что значит, что она содержит нечто подобное изолированным островам – собрания утверждений, рассматриваемых в качестве истинных в одно и то же время, без каких-либо более глубоких связей [Litwin, Miłkowski, 2020, p. 4].

Милковский полагает, что в настоящий момент предсказывающая обработка не способна выступить в качестве основания для великой унификации в когнитивных науках, она не может рассматриваться как теория великого объединения (grand unified theory). Таким образом, предсказывающую обработку, по его мнению, следует понимать в качестве универсального инструмента моделирования с неограниченными степенями свободы, исследовательской традиции в смысле Лаудана, а не претендента на статус единой теории когнитивных процессов.

И хотя следует согласиться с вердиктом Милковского в отношении выдвигаемых в рамках комплекса предсказывающего кодирования «новых предсказаний», некоторые другие его выводы и утверждения вызывают вопросы. Так, остается неясным его выбор предпочтительной единицы анализа в философии когнитивных наук: если предсказывающую обработку предлагается понимать в терминах концепции исследовательских традиций Лаудана, то как с этим могут быть совместимы поиски некой теории великого объединения (напоминающей парадигму в смысле Куна)? Здесь же возникает вопрос и о том, почему столь значительное внимание (не только со стороны Милковского, но и некоторых других когнитивистов) уделяется именно унификации как достоинству научных теорий (эпистемической ценности) и цели теоретического развития когнитивных наук.

Выше уже говорилось о том, что пролиферация научных теорий за счет взаимного критицизма и сохранения альтернатив должна вести к прогрессу в отношении целого ряда достоинств научных теорий: предсказательного успеха, непреднамеренного объяснения имеющихся значимых фактов и т.п. Действительно, традиционно в философии науки способность объединять разрозненные предметные области, унификацию, относят к числу такого рода эпистемических ценностей (приводя в качестве примеров объединение И. Ньютоном земной и небесной механики, объединение Дж. Максвеллом исследований электричества и магнетизма). Однако в области когнитивных наук и анализа их прогресса акцент на ценности унификации не может быть безоговорочным. В другой своей недавней работе я рассматривал противостояние плюрализма и монизма в философии когнитивных наук в трех важнейших контекстах: теоретическом, методологическом и объяснительном. Я отстаивал преимущества и необходимость плюрализма в теоретическом и методологическом отношениях (теоретический и методологический плюрализм предстают как необходимое следствие концепции теоретических комплексов). Между тем аналогичное заключение о безусловной необходимости объяснительного плюрализма сделано быть не может, т.к. характер объяснения обуславливается самим предметом исследования, а именно его однородностью или неоднородностью. Этот вопрос не может быть решен за письменным столом в кабинете теоретика, только непосредственные эмпирические исследования могут показать, какая объяснительная стратегия окажется наиболее уместной для когнитивных наук, плюралистическая или монистическая. Авторы, *ex cathedra* подчеркивающие значимость объяснительного монизма и унификации для когнитивных наук, не должны упускать из виду имеющиеся свидетельства неоднородности самого предмета когнитивистских исследований (по крайней мере, значительной их части), т.е. человеческого разума.

Так, одной из наиболее интересных и перспективных для когнитивных наук представляется гипотеза о гибридной последовательно-параллельной организации человеческого разума, представленная в работе французского нейрочеловека С. Деана [Dehaene, 2014]. В соответствии с данной гипотезой⁵, разум

⁵ Имеющей определенные сходства с более ранней гипотезой Д. Деннета [Dennett, 1991, p. 209–225].

(взрослого, овладевшего языком) человека предстает в качестве гибридной последовательно-параллельной машины, «в которой стадии массивных параллельных вычислений чередуются со стадией сознательного принятия решения и координации информации» [Dehaene, 2014].

С одной стороны, бессознательные (в особенности перцептивные и моторные) механизмы разума можно уподобить

...статистикам, использующим каждый малейший перцептивный признак (hint) – слабое движение, тень, пятно света – чтобы вычислить вероятность того, что данное свойство имеется в мире. Подобно тому, как метеобюро объединяет десятки метеорологических наблюдений, чтобы определить вероятность дождя в течение нескольких следующих дней, наше бессознательное восприятие использует входящие сенсорные данные для вычисления вероятностей цветов, форм, животных, присутствующих в нашем окружении [Ibid.].

Как представляется, именно в этой области, области объяснения работы бессознательных механизмов восприятия и действия, с лучшей стороны способны показать себя теории и модели предсказывающего кодирования, опирающиеся на указанные выше общие принципы работы механизма минимизации ошибки в предсказании.

С другой стороны, сознающий разум обладает своими особыми компетенциями, которые не могут быть реализованы механизмами, отвечающими за вероятностную бессознательную обработку. Одной из таких функций является функция организации и связи отдельных шагов в цепях многошаговых вычислительных операций.

К примеру, умножая числа 13 и 12, люди обычно проделывают в уме следующие операции. Ясно определив вычислительную задачу, число 12 возводят в квадрат, получая в качестве промежуточного результата число 144. После этого к 144 еще раз прибавляют 12 и получают искомое произведение, число 156. В этом случае (как и в других аналогичных ситуациях), указывает Деан, сознающий разум выполняет следующие присущие исключительно ему обязанности: сперва он формулирует общую вычислительную задачу (перемножить числа 13 и 12), после чего определяет первую подзадачу (возведение числа 12 в квадрат) и делегирует ее выполнение бессознательным механизмам. Сохраняя промежуточный результат (число 144), он снова определяет требующуюся вычислительную подзадачу (сложить числа 144 и 12), направляет ее для исполнения бессознательным системам и получает искомый результат (в данном примере число 156).

И хотя данный пример является весомым свидетельством в пользу плюралистической стратегии объяснения работы человеческого разума, я обратился к нему сейчас только с целью показать, что вопрос о необходимости объяснительного монизма и объяснительной унификации для когнитивных наук не должен решаться исключительно в теоретическом ключе. Касаемо же поиска некой «теории великого объединения» (по-видимому, заимствованный из физики термин) и вопросов о пригодности той или иной теории на эту роль, то с точки зрения концепции теоретических комплексов это просто-напросто ложная цель для когнитивных наук. В настоящий момент нет возможности

подробно рассматривать основные аргументы в пользу теоретического плюрализма и методологического плюрализма для когнитивных наук, это было сделано в других работах [Сущин 2022a]. Кратко: действительный, претворенный в практику отказ от теоретического плюрализма, от связанного с ним критицизма и сохранения альтернатив имел бы в долгосрочной перспективе негативное воздействие не только на когнитивные науки, но и на любую область научного познания (здесь можно вспомнить, что один из критиков Куна уподобил образ «нормальной науки» школам ортодоксальной теологии [Watkins, 1970]). При этом следует внести оговорку о том, что речь идет о действительных попытках отказа от плюрализма, т.к. ясно, что, даже руководствуясь мистическими идеалами, ученые на практике могут иметь (и в большинстве случаев имеют) плюрализм разных точек зрения. Я не возражаю против плюрализма «великих объединенных теорий» в когнитивных науках (если только будущие эмпирические исследования покажут, что этим теориям в самом деле лучше быть унифицированными в объяснительном плане).

Заключение

Итак, в этой статье я стремился показать, что не так давно возникшее, но уже ставшее влиятельным направление предсказывающей обработки следует понимать как теоретический комплекс в когнитивных науках. Были выявлены ключевые положения предсказывающей обработки и показано, что данный комплекс имеет однородную структуру. Кроме того, было показано, что ключевые положения комплекса действительно обеспечивают дальнейшую пролиферацию предсказывающих теорий и моделей когнитивных и/или ментальных процессов, как и предполагается в рамках концепции теоретических комплексов.

В то же самое время было сделано заключение, что пролиферация теорий предсказывающей обработки пока не привела к прогрессу в отношении наиболее значимых достоинств научных теорий (или эпистемических ценностей, если угодно), подобных формулированию и подтверждению предсказаний новых неожиданных фактов или непреднамеренному объяснению имеющихся фактов. Не стоит связывать это с нефальсифицируемостью теорий предсказывающей обработки, как это делает М. Милковский. Прежде всего, ввиду известной проблемы Дюгема фальсифицируемость можно рассматривать как важный, но, по выражению Г. Максвелла, чаще всего недостижимый регулятивный идеал для теорий. Возможно, сторонникам предсказывающей обработки следует сузить широту охвата их моделей до той области, где они показывают себя с лучшей стороны (т.е. до области исследований восприятия и действия), занять плюралистическую позицию, рассмотреть вопрос о более совершенных альтернативах для объяснения других когнитивных феноменов (и, конечно, о том, насколько предсказывающая обработка с ними совместима) и только в случае отсутствия таких альтернатив пытаться выдвигать предсказывающие объяснения соответствующих феноменов. В этом отношении упомянутая выше гипотеза смешанной последовательно-параллельной организации человеческого разума выступает хорошим примером.

Список литературы

- Сущин, 2021 – Сущин М.А. Когнитивная наука: от парадигм к теоретическим комплексам // *Философия науки и техники*. 2021. Т. 26. № 1. С. 5–22.
- Сущин, 2022а – Сущин М.А. Плурализм в когнитивных науках: теоретический, методологический или объяснительный? // *Философия и культура*. 2022. № 10. С. 117–131.
- Сущин, 2022б – Сущин М.А. Теоретические комплексы в когнитивных науках // *Вопросы философии*. 2022. № 12. С. 40–51.
- Bowers, Davis, 2012 – *Bowers J.S., Davis C.J.* Bayesian Just-So Stories in Psychology and Neuroscience // *Psychological Bulletin*. 2012. Vol. 138. No. 3. P. 389–414.
- Clark, 2013 – *Clark A.* Whatever next? Predictive brains, situated agents, and the future of cognitive science // *Behavioral and brain sciences*. 2013. Vol. 36. No. 3. P. 181–204.
- Dehaene, 2014 – *Dehaene S.* Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts [E-book]. N.Y.: Viking Press, 2014.
- Dennett, 1991 – *Dennett D.C.* Consciousness Explained. N.Y.: Little, Brown and Company, 1991. 528 p.
- Friston, Fortier, Friedman, 2018 – *Friston K.J., Fortier M., Friedman D.A.* Of woodlice and men: A Bayesian account of cognition, life and consciousness. An interview with Karl Friston // *ALIUS Bulletin*. 2018. Vol. 2. P. 17–43.
- Hohwy, 2013 – *Hohwy J.* The Predictive Mind. Oxford: Oxford University Press, 2013. 292 p.
- Hohwy, 2020 – *Hohwy J.* New directions in predictive processing // *Mind & Language*. 2020. Vol. 35. No. 2. P. 209–223.
- Hohwy, Roepstorff, Friston, 2008 – *Hohwy J., Roepstorff A., Friston K.* Predictive coding explains binocular rivalry: An epistemological review // *Cognition*. 2008. Vol. 108. No. 3. P. 687–701.
- Kuhn, 1996 – *Kuhn T.* The Structure of Scientific Revolutions. Chicago: The University of Chicago Press, 1996. 212 p.
- Lakatos, 1989 – *Lakatos I.* Falsification and the Methodology of Scientific Research Programmes // *Imre Lakatos: the methodology of scientific research programmes*. Philosophical Papers. Vol. 1. N.Y.: Cambridge University Press, 1989. P. 8–101.
- Lakatos, Zahar, 1989 – *Lakatos I., Zahar E.* Why did Copernicus’s research programme supersede Ptolemy’s // *Imre Lakatos: the methodology of scientific research programmes*. Philosophical Papers. Vol. 1. N.Y.: Cambridge University Press, 1989. P. 168–192.
- Laudan, 1978 – *Laudan L.* Progress and Its Problems: Towards a Theory of Scientific Growth. Berkeley and Los Angeles: University of California Press, 1978. 272 p.
- Litwin, Miłkowski, 2020 – *Litwin P., Miłkowski M.* Unification by Fiat: Arrested Development of Predictive Processing // *Cognitive Science*. 2020. Vol. 44. No. 7. P. 1–27.
- McMullin, 2008 – *McMullin E.* The virtues of a good theory // *The Routledge Companion to Philosophy of Science* / Ed. by S. Psillos, M. Curd. N.Y.: Routledge, 2008. P. 498–508.
- Miłkowski, Litwin, 2022 – *Miłkowski M., Litwin P.* Testable or bust: theoretical lessons for predictive processing // *Synthese*. 2022. Vol. 200. Article number: 462. DOI: 10.1007/s11229-022-03891-9.
- Seth, 2013 – *Seth A. K.* Interoceptive inference, emotion, and the embodied self // *Trends in Cognitive Sciences*. 2013. Vol. 17. No. 11. P. 565–573.
- Von Eckardt, 1995 – *Von Eckardt B.* What is Cognitive Science? Cambridge, Massachusetts: MIT Press, 1995. 477 p.
- Watkins, 1970 – *Watkins J.W.N.* Against “Normal Science” // *Criticism and the Growth of Knowledge* / Ed. by I. Lakatos, A. Musgrave. Cambridge: Cambridge University Press, 1970. P. 25–37.

Predictive processing as a theoretical complex in the cognitive sciences

Mikhail A. Sushchin

Institute of Scientific Information for Social Sciences, Russian Academy of Sciences. 51/21 Nakhimovskiy av., Moscow, 117997, Russian Federation; e-mail: sushchin@bk.ru

The article continues the series of papers devoted to the conception of theoretical complexes in the cognitive sciences. The key theses of the conception of theoretical complexes are used for the assessment of theories and models of predictive processing. The author proposes considering predictive processing as a theoretical complex in the cognitive sciences. The questions about the structure and functions of the complex of predictive processing have been addressed. The author singles out the main assumptions of this complex and points out that predictive processing has a homogenous structure. It has been shown that the main assumptions of the complex promoted further proliferation of predictive theories and models of cognitive and/or mental phenomena, in full accordance with the key theses of the conception of theoretical complexes. At the same time the author concludes that no significant theoretical progress has been made within the complex of predictive processing. It has also been shown that theoretical progress in the cognitive sciences should not be primarily related to the value of unification and that only future empirical studies can show which explanatory strategy (monistic or pluralistic) will turn out to be preferable for the cognitive sciences.

Keywords: theoretical complexes, cognitive sciences, predictive processing, theoretical progress, virtues of a good theory

References

- Bowers, J.S., Davis, C.J. "Bayesian just-so stories in psychology and neuroscience", *Psychological bulletin*, 2012, vol. 138, no. 3, pp. 389–414.
- Clark, A. "Whatever next? Predictive brains, situated agents, and the future of cognitive science", *Behavioral and brain sciences*, 2013, vol. 36, no. 3, pp. 181–204.
- Dehaene, S. *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts* [E-book]. New York: Viking Press, 2014.
- Dennett, D.C. *Consciousness Explained*. New York: Little, Brown and Company, 1991. 528 pp.
- Friston, K.J., Fortier, M., Friedman, D.A. "Of woodlice and men: A Bayesian account of cognition, life and consciousness. An interview with Karl Friston", *ALIVUS Bulletin*, 2018, vol. 2, pp. 17–43.
- Hohwy, J. *The Predictive Mind*. Oxford: Oxford University Press, 2013, 292 pp.
- Hohwy, J. "New directions in predictive processing", *Mind & Language*, 2020, vol. 35, no. 2, pp. 209–223.
- Hohwy, J., Roepstorff, A., Friston, K. "Predictive coding explains binocular rivalry: An epistemological review", *Cognition*, 2008, vol. 108, no. 3, pp. 687–701.
- Kuhn, T. *The Structure of Scientific Revolutions*. Chicago: The University of Chicago Press, 1996. 212 pp.
- Lakatos, I. "Falsification and the Methodology of Scientific Research Programmes", *Imre Lakatos: the methodology of scientific research programmes. Philosophical Papers. Vol. 1*. New York: Cambridge University Press, 1989, pp. 8–101.

Lakatos, I., Zahar, E. "Why did Copernicus's research programme supersede Ptolemy's", *Imre Lakatos: the methodology of scientific research programmes. Philosophical Papers. Vol. 1.* New York: Cambridge University Press, 1989, pp. 168–192.

Laudan, L. *Progress and Its Problems: Towards a Theory of Scientific Growth.* Berkeley and Los Angeles: University of California Press, 1978. 272 pp.

Litwin, P., Miłkowski, M. "Unification by fiat: arrested development of predictive processing", *Cognitive Science*, 2020, vol. 44, no. 7, pp. 1–27.

McMullin, E. "The virtues of a good theory", *The Routledge Companion to Philosophy of Science*, ed. by S. Psillos, M. Curd. New York: Routledge, 2008, pp. 498–508.

Miłkowski, M., Litwin, P. "Testable or bust: theoretical lessons for predictive processing", *Synthese*, 2022, vol. 200, article number: 462. DOI: 10.1007/s11229-022-03891-9.

Seth, A.K. "Interoceptive inference, emotion, and the embodied self", *Trends in cognitive sciences*, 2013, vol. 17, no. 11, pp. 565–573.

Sushchin, M.A. "Kognitivnaya nauka: ot paradigm k teoreticheskim kompleksam" [Cognitive science: from paradigms to theoretical complexes], *Filosofiya nauki i tekhniki / Philosophy of Science and Technology*, vol. 26, no. 1, pp. 5–22. (In Russian)

Sushchin, M.A. "Plyuralizm v kognitivnykh naukakh: teoreticheskiy, metodologicheskiy ili ob"yasnitel'nyy?" [Pluralism in the Cognitive Sciences: Theoretical, Methodological or Explanatory?] *Filosofiya i kul'tura*, 2022, no. 10, pp. 117–131. (In Russian)

Sushchin, M.A. "Teoreticheskiye komplekсы v kognitivnykh naukakh" [The Theoretical Complexes in the Cognitive Sciences]. *Voprosy filosofii*, 2022, no. 12, pp. 40–51. (In Russian)

Von Eckardt, B. *What is Cognitive Science?* Cambridge, Massachusetts: MIT Press, 1995. 477 pp.

Watkins, J.W.N. "Against 'Normal Science'", *Criticism and the Growth of Knowledge*, ed. by I. Lakatos, A. Musgrave. Cambridge: Cambridge University Press, 1970, pp. 25–37.