



С.Ю. СТЕПАНОВ*

Рефлексия и искусственный интеллект: от психологии рефлексии к рефлексивно- цифровым практикам развития человека

Статья прослеживает эволюцию представлений о рефлексии от её истоков у Дж. Локка и в интроспективной психологии до становления самостоятельной области психологоческой науки в конце XX в. Автор подчёркивает важность институционализации психологии рефлексии в 1990-е гг. для разработки рефлексивных практик и психотехник для консультирования, тренингов, психотерапии и образования. Основной смысловой фокус статьи сделан на революционном воздействии цифровых технологий и искусственного интеллекта (ИИ) на выявление новых возможностей рефлексии в XXI в. Показано, что интеграция ИИ в сферу жизнедеятельности человека трансформирует рефлексивные процессы и связанные с ними практики. Возникает феномен «би-рефлексии» при взаимодействии человеческого сознания и машинного интеллекта, выражающийся в появлении «цифровых кентавров» и новых рефлексивно-цифровых практик. Эти практики расширяют когнитивные и личностные ресурсы, одновременно ставя вопросы о меняющейся субъектности, идентичности и этике, а также стимулируют «очеловечивание» ИИ через наделение его функциями креативности и рефлексивности. Автор представляет две перспективные концептуальные разработки. Первая – Цифровой Ангел (DA): персональный ИИ-агент, выступающий как психологический буфер и защита от рисков цифрового мира (киберугроз, конкуренции с AGI и т.п.). DA обеспечивает безопасное цифровое пространство, анализирует данные пользователя для усиления его рефлексии и саморазвития. Вторая – Цифровое Альтер-Эго (DAE): технология для «конструирования» и «проживания» новых личностных амплуа в виртуальных/дополненных реальностях. DAE позволяет актуализировать у человека неочевидные способности и ресурсы, создавая уже не «зону ближайшего развития» (по Л.С. Выготскому), а «зону отдалённого развития» – перспективы самореализации через взаимодействие с ИИ. В статье утверждается, что технологии DA и DAE, основанные на отечественных разработках психологии рефлексии, определяют новый вектор для цифровой гуманистики и гибридных рефлексивных практик.

Ключевые слова: рефлексия, искусственный интеллект, рефлексивно-цифровые практики, цифровой ангел, цифровое альтер-эго

* Степанов Сергей Юрьевич — доктор психологических наук, профессор Московского государственного педагогического университета, руководитель Центра психологии метакогнитивных процессов и рефлексивных практик Института психологии РАН.

E-mail: parusnik1@ya.ru

Гетерохронность проблематики рефлексии и искусственного интеллекта

Представления о рефлексии и искусственном интеллекте (ИИ) формировались в ходе человеческой истории в разные времена, в разных социокультурных контекстах и с разным темпом. Если на формирование представлений о первой ушли века, то о втором – лишь несколько десятилетий. Несмотря на это, и для первой, и для второго современность является моментом чрезвычайно интенсивного развития, во многом обусловленным тем, что они стали активно взаимодействовать как в теоретической плоскости профессионального сознания, так и в технологическом и практическом своём воплощении. В данной статье попробуем раскрыть коллизии их взаимообусловленности и перспективы применения в контексте проблематики развития человека. Однако прежде чем непосредственно приступить к выполнению этой задачи, обозначим для начала трек становления исследований рефлексии до начала разработок ИИ и цифровых технологий.

Философско-психологический трек изучения рефлексии

В конце XVII в. Дж. Локк в своём знаменитом труде «Опыт о человеческом разумении» впервые ввёл термин «рефлексия» как важнейший источник знаний, приобретаемых на основе наблюдения человеком за деятельностью своего ума и души [7]. Данное представление о рефлексии воплотилось в главном гностическом инструменте возникшей на рубеже конца XIX – начала XX в. молодой психологической науки – в методе интроспекции. Позже, однако, метод самонаблюдения под давлением позитивистической методологии был раскритикован и элиминирован в психологии, а вслед за ним – и понятие «рефлексия».

До начала 70–80-х прошлого столетия рефлексия изредка использовалась в психологии и только лишь отдельными авторами как объяснительный принцип, позво-

ляющий интерпретировать и трактовать различные феномены, связанные с процессами психической саморегуляции человека и развития теоретического мышления. В работах В.В. Давыдова и его последователей рефлексия определялась как психический механизм осознания умственных действий и основа теоретического мышления [3]. У Я.А. Пономарёва, Н.Г. Алексеева, И.Н. Семёнова и В.К. Зарецкого рефлексия фигурировала в качестве уровневого компонента, регулирующего процесс творческого мышления [8].

Благодаря работам, опубликованным в 1982 [17] и 1983 г. [10], рефлексия стала обретать статус особой эмпирической реальности и предметно-психологической области исследований, что привело к формированию отдельной области науки – психологии рефлексии [18]. В рамках этой области удалось концептуально обосновать, методически выделить, эмпирически различить и экспериментально верифицировать в изначально синкетичном представлении о рефлексии не только её отдельные уровневые компоненты (например, такие виды, как ретроспективная, ситуативная, перспективная, экстенсивная, интенсивная, конструктивная рефлексия), но и различные типы рефлексии: интеллектуальная и личностная, коммуникативная и кооперативная, экзистенциальная и культуральная [18]. Рефлексия стала трактоваться не просто как сугубо интеллектуальный процесс аутоперцепции, осознания и ретроспективного анализа оснований мыслительных действий и переживаний человека, но как личностно обусловленный процесс осмыслиения и переосмыслиения стереотипов его опыта, а также содержаний (структур и процессов) сознания, мышления и поведения людей, лежащий в основе саморегуляции и самоорганизации, самодетерминации и саморазвития человека и приводящий к творческим открытиям и новациям в его деятельности и взаимодействии с другими людьми [18].

Осуществлённые в 90-е годы институционализация психологии рефлексии, разра-



Ил. 1. Рефлексирующий ИИ
(здесь и далее помещены иллюстрации,
созданные ИИ)

ботка метаметода сценарирования и проведения рефлексивных практик, а также отдельных рефлексивных психотехник (полилог, позиционная дискуссия, «коктейлирование», полипленум, композиция «проблемного поля», версификация символических инверсий, аукцион идей и т.д. и т.п.) [14] стали отправной точкой для их широкого применения в психологическом, управлении, организационном и политическом консультировании [13], при подготовке и проведении рефлексивных тренингов [9] и инновационных практик, направленных на модернизацию деятельности образовательных организаций [14]. Вместе с Е.П. Варламовой нам удалось создать концептуальные основы и разработать методические процедуры рефлексивной диагностики и самодиагностики, обладающие высоким практическим потенциалом для исследования и проектирования

процессов самопознания и саморазвития человека и организаций [2].

Вслед, а затем и параллельно с нашими разработками по психологии рефлексии уже в начале XXI века начали интенсивно разворачиваться исследования других отечественных психологов: А.В. Карпова [5], Д.А. Леонтьева, Е.Н. Осина [6], В.Д. Шадрикова, С.С. Кургиянья [19] и др. Помимо отечественных разработок по проблематике рефлексии, с 80-х гг. XX в. её интенсивная проработка осуществлялась и за рубежом в рамках психологии метакогнитивизма, прежде всего в исследованиях Ж. Пиаже, Дж. Дьюи, Дж. Флейвелла, Д. Дёрнера, Д. Шёна, Т. Бортона, Д. Колба и других.

Рефлексивный гребень цифровой волны

Однако все описанные достижения меркнут на фоне кардинальных подвижек в понимании рефлексии и её возможностей, которые стали происходить уже в XXI в., т.е. с началом создания компьютерно-цифровых технологий машинного обучения, обработки больших данных и нейронных сетей. Тогда специалисты в области ИТ-индустрии и Интернета стали активно асимилировать результаты исследований в области психологии мышления, творчества и других когнитивных процессов, что привело к революционным подвижкам в области ИИ, к созданию генеративных нейросетей и больших языковых моделей (LLM) (ил. 1).

На фоне активно развивающихся технологий ИИ, инкорпорируемых практически во все сферы психического бытия современного человека, с неизбежностью начинают трансформироваться опосредуемые ими психические процессы и структуры, в том числе и рефлексия. Это, в частности, подтверждают эмпирические исследования Г.У. Солдатовой с коллегами, в которых показано, что «гиперподключённость» человека к взаимодействию с цифровыми ресурсами [12] меняет структуру самоидентичности и самоотношения, а значит, и характер личностной рефлексии чело-



Ил. 2. Цифровой двойник

века [20]. Опираясь на возможности ИИ, человек открывает средства к усилению и изменению рефлексии, направленной не только на самого себя (т.е. интеллектуальной, личностной и экзистенциальной), но и на переосмысление действий и поведения других людей и их сообществ (т.е. коммуникативной, кооперативной и организационной).

Цифровые платформы, такие как чат-боты или системы анализа поведения, становятся «внешними агентами рефлексии», предлагая альтернативные интерпретации действий и мыслей пользователей [11]. Это создаёт феномен бирефлексии, где человеческое сознание и машинный интеллект взаимодействуют и взаимообуслов-

ливают друг друга, формируя гибридные когнитивно-цифровые модели и «цифровые кентавры» [12]. Тем самым сейчас мы становимся свидетелями и участниками формирования новых видов практик, а именно рефлексивно-цифровых практик, которые, основываясь на взаимодействии человека с ИИ, с одной стороны, открывают новые возможности для расширения когнитивно-интеллектуальных и личностно-эмоциональных ресурсов личности, а также ставят вопросы о переосмыслинии природы субъектности, формах человеческой идентичности, этических границах технологий и, в том числе, трансформации существующих моделей образования и терапии. С другой же сторо-

ны, трансформации затрагивают не только психическую сферу человека, но параллельно инициируют попытки разработчиков гуманизировать цифровую метавселенную и «очеловечить» искусственный интеллект, программируя его в направлении ресурсов и функций, которые ещё до недавнего времени считались прерогативой человеческой психики, а именно креативность, рефлексивность, диалогичность и т.п. (ил. 2).

К примеру, уже сейчас создаются человекоподобные голограммические аватары на основе моделирования чувствительного цифрового двойника, которые представляют собой сочетание эмоциональных вычислений, аффективной обратной связи и технологий Интернета вещей, которые предназначены для создания «чувствительной цифровой среды». Эта модель собирает и интегрирует мультимодальные данные в режиме реального времени в систему, которая может интерпретировать поведение реального человека и прогнозировать его операционные планы, синхронизируя все эти данные с поведением чувствительного цифрового двойника, который может выполнять роль, например, интерактивного «экскурсовода» в музее или «учителя» во время урока [21]. Очевидно, что здесь речь идёт об «овладении» искусственным интеллектом в образе «чувствующего» аватара функцией коммуникативной и кооперативной рефлексии.

Разработки отечественной психологии рефлексии также находят своё воплощение в области рефлексивно-цифровых практик и бирефлексивных моделей взаимодействия человека с ИИ.

От цифрового ассистента к рефлексивно-цифровому ангелу

Так, размышляя на рубеже 2018–2019 гг. вместе с Д.В. Ушаковым над перспективами развития технологии ИИ применительно к образовательным и психотерапевтическим задачам, мы предположили, что ближайшим и ключевым трендом её прогресса станет переход от разрабо-

ток Общего (Глобального и/или Супер) ИИ к созданию технологий Индивидуализированного (персонального) ИИ [22]. Этот прогноз основывался на рефлексии того факта, что подобное уже произошло трёх-четырьмя десятилетиями раньше с компьютерами, главный технологический сдвиг в разработке которых был сделан при переходе от больших вычислительных центров к персональным компьютерам и цифровым гаджетам.

Сегодня идеей персональных цифровых ассистентов, помощников, агентов и т.п. на основе ИИ уже никого не удивишь. Вместе с тем, анализируя предложенные на сегодняшний день модели цифровых ассистентов, можно сказать, что практически ни одна из этих цифровых «поделок» пока не дотягивает до предложенного нами концепта «Цифрового Ангела», поскольку их разработчики стремятся удовлетворить сиюминутные, по большей части утилитарные и весьма прагматические потребности людей, использующих возможности ИИ, которые связаны с решением, как правило, бытовых, развлекательных, сервисных, бизнесовых или узкопрофессиональных задач (ил. 3).

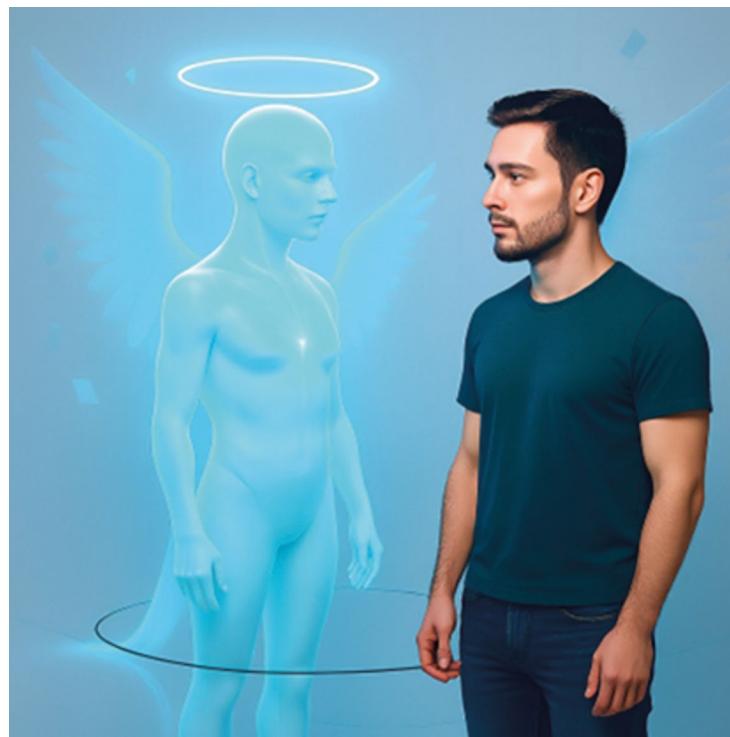
Концепция же ДА предполагает, что его возможности не просто ограничиваются вспомогательно-сервисными функциями, направленными на разгрузку человека от утилитарных и рутинных задач интеллектуального характера, но прежде всего на обеспечение его большей рефлексивности и, следовательно, большей защищённости от рисков и напастей, нарастающих в лавинообразно цифровизирующемся мире с разгулом кибермошенничества, индоктринационных и манипулятивных технологий, кибербуллинга, сетевого троллинга фейков и т.п. цифровых угроз. Усилившиеся возможности ИИ, в контексте вероятной конкуренции с естественным интеллектом человека и человечества в целом, грозящие обрушением рынков труда, также являются причиной всё возрастающей тревоги не только со стороны общественности, но даже самих IT-разра-

ботчиков. Из-за указанных рисков разработка технологии DA становится насущной необходимостью, поскольку она и должна стать своего рода рефлексивно-цифровым волнорезом перед нарастающей экспансией «Общего искусственного интеллекта» (Artificial General Intelligent – AGI) в различных его видах и формах.

В контексте разворачивающихся сегодня дискуссий о понимании позиционирования AGI относительно человека либо как «конкурирующего» (competitive), либо как «усиливающего» (augmenting), либо как «дополняющего» (complementary) разработка технологии DA как раз и может стать ключом к тому, чтобы первый вариант не стал доминирующим, а тем более единственным. А об опасности такого рода перспективы заявил Илон Маск – один из основных инициаторов и инвесторов разработки ChatGPT. В частности, в апреле 2024 г. он написал, что к концу 2025 г. возможности новых моделей искусственного интеллекта превзойдут человеческий разум, и в течение следующих пяти лет возможности AGI, способного выполнять умственную работу на человеческом уровне, превзойдут возможности всех людей [4].

DA – это технология, всецело выстроенная в интересах персоны конкретного человека, являющаяся психологическим буфером между ним и в первую очередь AGI, а также другими различными цифровыми сущностями (неважно, «добрьими» или «злыми», в силу относительности этих оценочных категорий в различных контекстах). Тем самым DA должен встать «на границе» цифрового мира и человека с его естественным интеллектом, обеспечив его безопасным личным цифровым пространством жизнедеятельности.

По нашему замыслу DA должен стать личным конфидентом человека, его



Ил. 3. «Цифровой Ангел» — персонализированный ИИ

рефлексивно-цифровым двойником, обеспечивающим не только обратной связью относительно интеракций с цифровым миром, но также прогнозами и оценками рисков и возможностей, которые они сулят ему в доступном для прогнозирования пространстве его жизнедеятельности. В этом смысле одной из важных составных частей данной технологии станет цифровой двойник, создаваемый и непрерывно обновляющийся за счёт накопления «больших данных» в виде цифровых следов жизнедеятельности человека, а также аналитики, которая будет материалом для его когнитивной, социальной, интеллектуальной и экзистенциальной рефлексии, а значит, основой для дальнейшего само-проектирования им своих жизненных треков развития.

Цифровое Альтер-Эго – мультипликатор человеческих возможностей и способностей

Важным добрачиванием технологии Цифрового Ангела может стать ещё одна предлагаемая нами технология – Цифро-



Ил. 4. Цифровое Альтер-Эго человека

вого Альтер-Эго (Digital Alter Ego – DAE). Она может выступать в качестве источника неведомых и неочевидных ранее психологических возможностей, ресурсов и способностей человека (ил. 4). DAE – это персонализированная технология ИИ, которая позволит человеку погружаться и действовать не просто в компьютерно-игровой виртуальной и/или дополненной реальности в качестве «аватара», но конструировать и примерять на себя новые личностные, интеллектуальные и коммуникативные амплуа с самым разнообразным набором мыслительных действий, ролей, компетенций, характерологических особенностей, потребностей и мотивов. И как актёр благодаря драматургу и режиссёру перевоплощается в определённую роль или сценического героя, проживая все коллизии и перипетии их судьбы, точно так же и любой человек с помощью DAE сможет разрабатывать различные личностные амплуа, вживаться в них и реализовывать себя в новых ролях в самых разнообразных сценариях с набором заранее запланированных и/или, наоборот, не-

предсказуемых жизненных ситуаций и коллизий, профессиональных профилей, образовательных и карьерных треков. Такие сценарии, генерируемые совместно человеком и ИИ, будут содержать различные проблемные, конфликтные и даже экзистенциальные коллизии. Тем самым человек сможет в проактивном режиме актуализировать, реализовывать и рефлексировать с помощью DAE различные свои «дремлющие» и наличные психологические задатки и способности, в том числе креативного характера.

В этом контексте может получить своё дальнейшее развитие весьма популярная в отечественной образовательной практике концепция «зоны ближайшего развития», предложенная Л.С. Выготским, когда коммуникация и взаимодействие человека с меньшей компетентностью (как правило, таковым является ребёнок) с человеком, обладающим большей компетентностью (здесь обычно подразумевается взрослый), создаёт для первого «зону ближайшего развития» (ЗБР). Интеракции же человека с ИИ, который уже сейчас аккумулирует практически весь доступный на данным момент опыт человечества за всю историю его существования, в рамках технологии DAE смогут, с нашей точки зрения, создавать для конкретного человека зону его отдалённого развития (ЗОР). Ведь вживание в различные комбинации психологических амплуа, неожиданных для человека, их поведенческое апробирование в предлагаемых виртуальных или дополненных реальностях при взаимодействии с другими подобными же персоналиями в ситуациях с высоким уровнем неопределённости и новизны условий самореализации должно будет актуализировать новые для самого человека ценности, мысли и действия, а также мобилизовать силы, о которых он мог не догадываться или даже не смел помышлять.

Возможность создания эффекта «зоны отдалённого развития» была проверена нами на практике. В качестве концептуально-психологического прообраза DAE выступила одна из психотехник, которые разрабатывались и успешно апробировались нами с конца 90-х годов в целом ряде рефлексивно-креативных практик в работе с командами и коллективами образовательных, государственных и бизнес-организаций целого ряда регионов [15]. Изначально эта психотехника была названа «методом рефлексивных инверсий и с творческих импровизаций». С её помощью создались условия, когда участники могли актуализировать, переосмысливать и преодолевать свои внутри- и межличностные проблемы и конфликты, производные от тех стереотипов поведения, которые выступают для них барьерами в личностном росте, в образовании, в развитии их профессиональных способностей и компетенций.

Процедурно это достигалось тем, что участникам рефлектируемости предлагалось: 1) вспомнить своих любимых сказочных, литературных или киногероев; 2) в ходе коллективного обсуждения произвести инверсии имён и психотипов последних (например, за счёт изменения нескольких букв в именах этих героев так, чтобы оставалось созвучие с исходным прототипом, но изменялся смысловой и психологический профиль персонажа); 3) дофантазировать (досочинить) их жизненные треки на основе осуществлённых инверсий; 4) вжиться в полученные таким образом инвертированные психологические амплуа; 5) реализовать их в смоделированных проблемно-конфликтных ситуациях в режиме психологической и/или социально-ролевой импровизации перед другими командами участников, выступающих в качестве рефлексивных супервизоров; и, наконец, 6) осуществить специально организованную личностную, интеллектуальную, коммуникативную, кооперативную и даже экзистенциальную рефлексию с помощью инструментов психо- и социографирования.

С одной стороны, такая рефлексивно-практическая проработка способствовала актуализации различных личностных стереотипов поведения (комплексов, по З. Фрейду) и социокультурных стереотипов (архетипов коллективного бессознательного, по К. Юнгу), а также их проживанию и изживанию в психологически защищённой среде и, соответственно, переосмыслению и креативной трансформации в новые степени личной свободы, в новые способности и горизонты возможностей развития. Как правило, изумлению и радости от открываемых как в себе, так и в других участниках неожиданных возможностей, талантов и способностей не было пределов. Глубокая рефлексия проживания такого опыта открывала удивительные и неожиданные для самих участников личностные качества креативной направленности, а также приводила к рекреативным и/или психотерапевтическим эффектам, к существенным, как правило позитивным, изменениям в жизненных и карьерных треках [1].

Интересно, что после уже многих лет реализации психотехники и рефлексивных практик создания «зон отдалённого развития» в пространстве социокультурного освоения появился художественный аналог технологии DAE в фантастическом фильме Джеймса Кэмерона «Аватар», вышедшем на экраны в 2009 г., где главный герой имел возможность за счёт особой ментально-цифровой и генно-биологической технологии далёкого будущего, созданной людьми, «переселяться» из своей телесной оболочки человека-инвалида в оболочку фантастического существа с планеты Пандора. Это открыло для него совершенно новую, необычную и даже невероятную по своей экзистенциальной полноте и насыщенности жизнь, в событийных перипетиях которой он попутно овладел целым рядом возможностей и способностей, не свойственных для человека, но присущих инопланетным аборигенам Пандоры.

В связи с этим важно отметить: опыт разработки и внедрения цифровых и генных технологий, основанных на ИИ, дока-

зывает: то, что вчера казалось несбыточной фантазией, завтра может стать обычной и повседневной реальностью.

И действительно, современные технологии на базе ИИ – генеративные нейросети (для создания текстового, аудио- и видеоконтента, голограмм), big-data, 3D-моделирование, цифровые двойники, чат-боты, компьютерные игры и квесты, дополненная и виртуальная реальность и т.д. и т.п. – уже вплотную подводят нас к возможности создания технологии DAE. Встраивание в контекст этих технологических новаций концепций DA и DAE позволяет целенаправить их дальнейшее совершенствование и развитие по абсолютно новому и весьма перспективному пути, придав им особо ценный психологический смысл и социо-

культурное значение в контексте цифровой гуманитаристики.

Для психологии же разработка технологии DAE наряду с технологией DA актуализирует как целый ряд проблем, так и новых возможностей для фундаментальных и прикладных исследований. В первую очередь это вопросы соотношения рациональных и иррациональных, рефлексивных и бессознательных, эксплицитных и имплицитных, лингвистических и паралингвистических процессов и феноменов в цифровой реальности и взаимодействии человека и человеческих сообществ с ИИ. Но это уже темы для дальнейших размышлений и разработок в области рефлексивно-цифровых технологий и моделей бирефлексивных практик.

ЛИТЕРАТУРА

1. Варламова Е.П., Степанов С.Ю. Психология творческой уникальности человека: рефлексивно-гуманистический подход. М.: ИП РАН, 2002.
2. Варламова Е.П., Степанов С.Ю. Рефлексивная диагностика в системе образования // Вопросы психологии. 1997. № 5. С. 28–43.
3. Давыдов В.В. Теория развивающего обучения. М.: Педагогика, 1996.
4. Илон Маск заявил о возможности ИИ стать умнее человека уже в 2025 г. [Электронный ресурс] // Forbes. 2025. 4 апр. URL: <https://www.forbes.ru/milliardery/509969-ilon-mask-zaavil-o-vozmozhnosti-ii-stat-umnee-celoveka-uze-v-2025-godu> (дата обращения: 25.05.2025).
5. Карпов А.В. Рефлексивность как психическое свойство и методика её диагностики // Психологический журнал. 2003. Т. 24, № 5. С. 45–57.
6. Леонтьев Д.А., Осин Е.Н. Рефлексия «хорошая» и «дурная»: от объяснительной модели к дифференциальной диагностике // Психология. Журнал Высшей школы экономики. 2014. Т. 11, № 4. С. 110–135.
7. Локк Дж. Сочинения: В 3 т. Т. 1. М.: Мысль, 1985.
8. Пономарёв Я.А., Семёнов И.Н., Зарецкий В.К. [и др.] Исследование проблем психологии творчества / Под ред. Я.А. Пономарёва. М.: Наука, 1983.
9. Растворников А.В., Степанов С.Ю., Ушаков Д.В. Рефлексивное развитие компетентности в совместном творчестве. М.: PerSe, 2002.
10. Семёнов И.Н., Степанов С.Ю. Проблема предмета и метода психологического изучения рефлексии // Исследование проблем психологии творчества / Под ред. Я.А. Пономарёва. М.: Наука, 1983. С. 154–182.
11. Скурут А.И. ИИ как расширение рефлексивных способностей [Электронный ресурс]. URL: https://www.b17.ru/blog/ai_and_reflection (дата обращения: 25.06.2024).
12. Солдатова Г.У., Чигарькова С.В., Илюхина С.Н. Метаморфозы идентичности человека достроенного: от цифрового донора к цифровому кентавру // Социальная психология и общество. 2024. Т. 15, № 4. С. 40–57. DOI: 10.17759/sps.2024150404.
13. Степанов С.Ю. Развитие рефлексивной культуры госслужащих инновационно-гуманистическими методами: Дисс. ... докт. психол. н. М.: РАГС, 1999.

14. Степанов С.Ю. Рефлексивная практика творческого развития человека и организаций. М.: Наука, 2000.
15. Степанов С.Ю. Рефлексивно-гуманистическая психология створчества. М.; Петрозаводск, 1996.
16. Степанов С.Ю., Оржековский П.А., Ушаков Д.В. [и др.] Цифровизация образования: психолого-педагогические и валеологические проблемы / Под редакцией С.Ю. Степанова. М.: МГПУ, 2021.
17. Степанов С.Ю., Семёнов И.Н. Проблема формирования типов рефлексии в решении творческих задач // Вопросы психологии. 1982. № 1. С. 99–104.
18. Степанов С.Ю., Семёнов И.Н. Психология рефлексии: проблемы и исследования // Вопросы психологии. 1985. № 3. С. 31–40.
19. Шадриков В.Д., Кургянин С.С. Исследование рефлексии деятельности и её диагностика через оценку конструктов психологической функциональной системы деятельности // Экспериментальная психология. 2015. Т. 8, № 1. С. 106–126.
20. Belk R.W. Possessions and the Extended Self // Journal of Consumer Research. 1988. Vol. 15, № 2. Pp. 139–168.
21. Fominska I., Di Tore S., Nappi M., Iovane G., Sibilio M., Gelo A. Approaches to identifying emotions and affections during the museum learning experience in the context of the future internet // Future Internet. 2024. Vol. 16, № 11. Art. 417. DOI: 10.3390/fi16110417.
22. Stepanov S.Yu., Ushakov D.V. Artificial intelligence and digital angel technology in education // Восьмнадцатая Национальная конференция по искусственному интеллекту с международным участием КИИ–2020. М.: Рос. ассоц. искусств. интеллекта, 2020. С. 105–112.

Reflection and Artificial Intelligence: From the Psychology of Reflection to Reflexive Digital Practices of Human Development

Sergey Yurievich Stepanov – Doctor of Sciences in Psychology, Professor at the Moscow State Pedagogical University, Head of the Center for the Psychology of Metacognitive Processes and Reflexive Practices of the Institute of Psychology of the Russian Academy of Sciences.

E-mail: parusnik1@ya.ru

This article traces the evolution of the concept of reflection from its origins with J. Locke and introspective psychology to its establishment as a distinct field within psychological science by the end of the 20th century. The author emphasizes the importance of the institutionalization of the psychology of reflection in the 1990s for developing reflective practices and psycho-techniques used in counseling, training, psychotherapy, and education.

The main semantic focus of the article is on the revolutionary impact of digital technologies and artificial intelligence (AI) on uncovering new possibilities for reflection in the 21st century. It demonstrates that the integration of AI into human activity transforms reflective processes and related practices. This gives rise to the phenomenon of "bi-reflection" during the interaction between human consciousness and machine intelligence, manifesting in the emergence of "digital centaurs" and new reflective-digital practices.

These practices expand cognitive and personal resources while simultaneously raising questions about changing subjectivity, identity, and ethics. They also stimulate the "humanization" of AI by endowing it with functions of creativity and reflectivity.

The author presents two promising conceptual developments. The first is the Digital Angel (DA): a personal AI agent that acts as a psychological buffer and protection against the risks of the digital world (such as cyber threats and competition with AGI). The DA provides a se-

ture digital space and analyzes user data to enhance their reflection and self-development. The second is the Digital Alter-Ego (DAE): a technology for “constructing” and “experiencing” new personal roles in virtual and augmented realities. The DAE allows individuals to activate latent abilities and resources, creating not a “zone of proximal development” (per L.S. Vygotsky) but a “zone of distal development” – offering prospects for self-realization through interaction with AI. The article argues that DA and DAE technologies, grounded in domestic advancements within the psychology of reflection, define a new vector for the digital humanities and hybrid reflective practices.

Keywords: reflection, artificial intelligence, reflexive digital practices, digital angel, digital alter ego

REFERENCES

1. Varlamova E.P., Stepanov S.Yu. Psikhologiya tvorcheskoi unikal'nosti cheloveka: refleksivno-gumanisticheskii podkhod. M.: IP RAN, 2002 (in Russian).
2. Varlamova E.P., Stepanov S.Yu. Refleksivnaya diagnostika v sisteme obrazovaniya // Voprosy psichologii. 1997. № 5. S. 28–43 (in Russian).
3. Davydov V.V. Teoriya razvivayushchego obucheniya. M.: Pedagogika, 1996 (in Russian).
4. Ilon Mask zayavil o vozmozhnosti II stat' umnee cheloveka uzhe v 2025 g. [Elektronnyi resurs] // Forbs. 2025. 4 apr. URL: <https://www.forbes.ru/milliardery/509969-ilon-maskzaavil-o-vozmozhnosti-ii-stat-umnee-celoveka-uze-v-2025-godu> (data obrashcheniya: 25.05.2025) (in Russian).
5. Karpov A.V. Refleksivnost' kak psikhicheskoe svoistvo i metodika eyo diagnostiki // Psichologicheskii zhurnal. 2003. T. 24, № 5. S. 45–57 (in Russian).
6. Leont'ev D.A., Osin E.N. Refleksiya «khoroshaya» i «durnaya»: ot ob"yasnitel'noi modeli k differentsial'noi diagnostike // Psichologiya. Zhurnal Vysshei shkoly ekonomiki. 2014. T. 11, № 4. S. 110–135 (in Russian).
7. Lokk Dzh. Sochineniya: V 3 t. T. 1. M.: Mysl', 1985 (in Russian).
8. Ponomaryov Ya.A., Semyonov I.N., Zaretskii V.K. [i dr.] Issledovanie problem psichologii tvorchestva / Pod red. Ya.A. Ponomaryova. M.: Nauka, 1983 (in Russian).
9. Rastyannikov A.V., Stepanov S.Yu., Ushakov D.V. Refleksivnoe razvitiye kompetentnosti v sovmestnom tvorchestve. M.: PerSe, 2002 (in Russian).
10. Semyonov I.N., Stepanov S.Yu. Problema predmeta i metoda psikhologicheskogo izucheniya refleksii // Issledovanie problem psikhologii tvorchestva / Pod red. Ya.A. Ponomaryova. M.: Nauka, 1983. S. 154–182 (in Russian).
11. Skurtul A.I. II kak rasshirenie refleksivnykh sposobnostei [Elektronnyi resurs]. URL: https://www.b17.ru/blog/ai_and_reflection (data obrashcheniya: 25.06.2024) (in Russian).
12. Soldatova G.U., Chigar'kova S.V., Ilyukhina S.N. Metamorfozy identichnosti cheloveka dostroenennogo: ot tsifrovogo donora k tsifrovomu kentavru // Sotsial'naya psichologiya i obshchestvo. 2024. T. 15, № 4. S. 40–57. DOI: 10.17759/sps.2024150404 (in Russian).
13. Stepanov S.Yu. Razvitiye refleksivnoi kul'tury gossluzhashchikh innovatsionno-gumanisticheskimi metodami: Diss. ... dokt. psikhol. n. M.: RAGS, 1999 (in Russian).
14. Stepanov S.Yu. Refleksivnaya praktika tvorcheskogo razvitiya cheloveka i organizatsii. M.: Nauka, 2000 (in Russian).
15. Stepanov S.Yu. Refleksivno-gumanisticheskaya psichologiya sotvorchestva. M.; Petrozavodsk, 1996 (in Russian).

16. Stepanov S.Yu., Orzhekovskii P.A., Ushakov D.V. [i dr.] Tsifrovizatsiya obrazovaniya: psichologo-pedagogicheskie i valeologicheskie problemy / Pod redaktsiei S.Yu. Stepanova. M.: MGPU, 2021 (in Russian).
17. Stepanov S.Yu., Semyonov I.N. Problema formirovaniya tipov refleksii v reshenii tvorcheskikh zadach // Voprosy psichologii. 1982. № 1. S. 99–104 (in Russian).
18. Stepanov S.Yu., Semyonov I.N. Psichologiya refleksii: problemy i issledovaniya // Voprosy psichologii. 1985. № 3. S. 31–40 (in Russian).
19. Shadrikov V.D., Kurginyan S.S. Issledovanie refleksii deyatel'nosti i eyo diagnostika cherez otsenku konstruktov psichologicheskoi funktsional'noi sistemy deyatel'nosti // Eksperimental'naya psichologiya. 2015. T. 8. № 1. S. 106–126 (in Russian).
20. Belk R.W. Possessions and the Extended Self // Journal of Consumer Research. 1988. Vol. 15, № 2. Pp. 139–168.
21. Fominska I., Di Tore S., Nappi M., Iovane G., Sibilio M., Gelo A. Approaches to identifying emotions and affections during the museum learning experience in the context of the future internet // Future Internet. 2024. Vol. 16, № 11. Art. 417. DOI: 10.3390/fi16110417.
22. Stepanov S.Yu., Ushakov D.V. Artificial intelligence and digital angel technology in education // Vosemnadtsataya Natsional'naya konferentsiya po iskusstvennomu intellektu s mezhdunarodnym uchastiem KII–2020. M.: Ros. assots. iskusstv. intellekta, 2020. S. 105–112.