
КОЛОНКА РЕДАКТОРА NOTES FROM EDITOR

Введение

Карданова Е.Ю.,

*кандидат физико-математических наук, доцент, ординарный профессор,
директор Центра мониторинга качества образования, Институт образования, Национальный исследовательский
университет «Высшая школа экономики», Москва, Россия,
ekardanova@hse.ru*

Орел Е.А.,

*кандидат психологических наук, научный сотрудник, Институт образования,
Центр мониторинга качества образования, Национальный исследовательский университет
«Высшая школа экономики», Москва, Россия,
eorel@hse.ru*

Современная психодиагностика — это область исследований и практики, которая давно уже выросла за пределы собственно психологии. Начавшись во второй половине XIX века с экспериментов Гальтона, психодиагностика за последние 150 лет прошла очень большой путь. Сегодня психодиагностические исследования немислимы вне кооперации специалистов в разных областях: психологов, математиков, социологов, биологов. Ее необходимость связана с тем, что с развитием технологий, культуры, сильно изменился класс задач, которые решает психодиагностика. Благодаря компьютерному тестированию, мы можем охватить все большие выборки; различные мультимедиа форматы позволяют расширить горизонты измеряемых показателей, знания о различных культурных практиках ложатся в основу межстрановых сравнительных исследований.

Разговор об отечественной психодиагностике принято начинать с печально известного постановления 1936 года «О педологических извращениях в системе наркомпросов», которое на полвека затормозило развитие количественных исследований. Однако после того, как официальные запреты на применение тестирования в массовой практике были сняты, психодиагностическая культура в нашей стране начала формироваться довольно быстро. Возникают локальные сообщества разработчиков и пользователей тестов, растет понимание того, что такое качественный психодиагностический инструмент, появляются стандарты его разработки и администрирования, а вместе с этим — целые учебные программы, которые готовят специалистов нового поколения. Одновременно с этим укрепляются и международные связи: наши специалисты проходят стажировки в ведущих мировых психометрических центрах, успешно выступают на международных конференциях, участвуют в летних школах. Российские научные журналы также хорошо отражают этот тренд: специальные выпуски, публикующие статьи о разработке и вали-

дизации стандартизованных психодиагностических методик в последнее время выходили во многих значимых изданиях. Текущий выпуск продолжает и укрепляет эту традицию.

Мы не ставим своей целью охватить в этом выпуске весь спектр психодиагностических исследований. К примеру, мы не затрагиваем здесь активно развивающуюся и очень интересную область аппаратурной психодиагностики. Современные технологии исследования мозга позволяют существенно уточнить наши знания об индивидуальных особенностях в сфере способностей, восприятия, личности.

Современные технологии позволяют собирать и анализировать объективные данные о человеке не только когда он лежит в МРТ-сканере или подключен к ай-трекеру. Многочисленные устройства, которые мы используем в повседневной жизни (смартфоны, планшеты, компьютеры, фитнес-браслеты), собирают и хранят самую разнообразную информацию о нас, которая также несет отпечаток нашей индивидуальности. На этом принципе построен анализ цифрового следа — то есть особенностей поведения человека в цифровой среде. В статье И.Л. Углановой, И.В. Брун и Г.М. Васина, описывающей приложение систематического подхода к разработке тестов (Evidence-Centred Design) к измерению сложных конструктов, в том числе, показано, каким образом поведение респондента в рамках работы со специально сконструированным компьютерным сценарием позволяет получить свидетельства об уровне развития его креативности и критического мышления. Эти свидетельства — не результат решения искусственной задачи, не самоотчет респондента о том, как он вел себя в ситуациях, требующих креативности и критического мышления, а фиксация его действий и продуктов деятельности в рамках решения задачи, максимально приближенной к тому, с чем он сталкивается в повседневной жизни.

В статье А.А. Марголиса с соавторами также описывающей потенциал современных цифровых технологий для психодиагностики, показано, как поведение в среде компьютерной игры может стать индикатором уровня интеллекта и креативности игрока.

Другой аспект объективности оценивания — это обеспечение равных возможностей респондентов, проходящих тест. При разработке измерительного инструмента необходимо учитывать индивидуальные особенности, которые влияют на результаты тестирования, но при этом не являются объектом оценки. Так, в статье Е.И. Брагинец показывается, как универсальный дизайн тестирования позволяет учесть особые образовательные потребности респондентов (на примере нарушений чтения), не меняя при этом природу измеряемого конструкта, и тем самым обеспечить равные возможности для всех категорий респондентов. Этот подход особенно востребован в оценке образовательных результатов в условиях инклюзивной школы.

Однако и классические самоотчетные методы продолжают занимать важное место в современной психодиагностике. Их особенности и сферы применения были хорошо описаны, и в рамках этих ограничений они продолжают оставаться востребованными, так как, с одной стороны, они требуют относительно небольшого количества ресурсов для разработки, а с другой стороны, легко масштабируются для разных условий тестирования и позволяют работать там, куда современные технологии еще не успели проникнуть. В текущем выпуске представлено сразу несколько подобных инструментов: мониторинг развития социально-эмоциональных навыков в начальной школе (статья Е.А. Орел и А.А. Куликовой), опросник для измерения самооценки (статья Н.В. Лебедевой и Ю.В. Кузьминой), батарея методик исследования суицидального поведения (статья Т.В. Журавлевой, С.Н. Ениколопова, Н.В. Дворянчикова, Ю.А. Бубеева).

Большое внимание в данном сборнике уделено относительно новому для российской психодиагностики подходу к построению инструментов (тестов, опросников, шкал и т. д.) и психометрическому анализу отдельных вопросов и инструментов в целом. Традиционно качество инструментов оценивалось в рамках классической теории тестирования (КТТ). Эта теория была создана еще в первой половине прошлого века, но и в настоящее время активно используется специалистами во всем мире. Однако иногда в практике тестирования возникают задачи, которые невозможно решить в рамках КТТ.

Перечислим основные ограничения КТТ применительно к проблемам тестирования в психологии. Главный ее недостаток — отсутствие свойства инвариантности, что подразумевает, что тестовые баллы респондентов зависят от конкретного инструмента, который использовался. Другой тест, измеряющий тот же конструкт, может привести к другим результатам. Это в полной мере касается и оценки заданий теста: в классической модели результат их выполнения соот-

носим только с конкретной выборкой респондентов. Именно поэтому на смену классической теории пришла новая теория, получившая название современной теории тестирования или Item Response Theory (IRT). При использовании современной теории тестирования показатели респондентов не зависят от параметров заданий, при выполнении которых они получены, равно как оценки параметров заданий не зависят от конкретной выборки испытуемых. Более того, результаты оценки параметров (и заданий, и респондентов) отображаются на общей метрической (интервальной) шкале и сопровождаются индивидуальными характеристиками точности оценивания. Наличие общей метрической шкалы позволяет сопоставлять (с помощью специальных процедур выравнивания) результаты тестирования, полученные по отчасти различным тестам (и даже в различное время), измеряющим один конструкт.

Более того, современная теория тестирования дает дополнительные возможности анализа функционирования шкал и ответных категорий. Так, например, помимо индексов, аналогичных классической надежности, существует ряд дополнительных характеристик точности измерений — например, индекс отделимости (separation index), показывающий, на сколько статистически различных с учетом ошибки измерения групп может быть поделена выборка всех респондентов. Для психологического тестирования важно и то, что IRT позволяет на стадии анализа данных тестирования оптимизировать, в случае необходимости, число ответных категорий. Разработчики иногда закладывают в шкалу много категорий в надежде усилить различительную силу инструмента. Однако респонденты не всегда в состоянии различить категории, близкие по смыслу (например, категории «обычно неверно» и «верно очень редко»), что приводит к неправильному функционированию шкалы. Возможности анализа в рамках IRT позволяют выявить такие ответные категории и проверить функционирование шкалы, если, например, две соседние, плохо функционирующие категории объединить в одну. Интересно, что сделать это можно на стадии анализа, даже не предъявляя тест еще одной выборке респондентов (хотя для валидации новой шкалы желательно провести повторное тестирование). Можно смело сказать, что применение IRT обогащает классический анализ, предоставляя массу дополнительных возможностей для проверки психометрических свойств, как отдельных заданий, так и всех инструментов в целом.

Применительно к моделям психологического тестирования IRT моделирует вероятность ответа респондента в данной ответной категории как функцию измеряемой характеристики респондента (уровня измеряемой способности, степени выраженности измеряемой черты и т. д.) и трудности выбора ответа в данной категории. При этом измеряемых характеристик респондентов может быть одна или несколько (модели, соответственно, называются одномерными

или многомерными). Наиболее часто в психологическом тестировании используется модель рейтинговых шкал — Rating Scale Model (RSM), специально разработанная для шкал типа Ликерта, имеющих единую шкалу ответных категорий для всех вопросов теста (например, от «полностью не согласен» до «полностью согласен» в классической шкале Ликерта).

В общей сложности, 5 статей сборника используют IRT для анализа и валидации различных методик и шкал. Так, в статье И.В. Антипкиной анализируется опросник дошкольной родительской вовлеченности; в статье А.Е. Ивановой — опросник для учителя, который нацелен на оценку поведения ребенка в условиях школы; в статье Н.В. Лебедевой и Ю.В. Кузьминой — опросник для измерения самооценки. В статье Е.А. Орел и А.А. Куликовой описывается процесс валидации инструмента мониторинговой оценки социально-эмоциональных навыков в начальной школе, и, наконец, в статье Т.Н. Канонир, И.Л. Углановой, Д.А. Федерякина рассматривается процесс адаптации и валидации русскоязычной версии шкал удовлетворенности учащихся школой и удовлетворенности учащихся друзьями.

В целом, анализ психометрических свойств любого опросника в рамках IRT предполагает следующую последовательность шагов.

1. Выбор модели тестирования.
2. Исследование факторной структуры опросника.
3. Исследование согласованности эмпирических данных с выбранной моделью измерения.
4. Анализ психометрических свойств отдельных вопросов.
5. Исследование карты переменных.
6. Исследование функционирования ответных категорий шкалы.
7. Выявление вопросов, по-разному функционирующих для респондентов из различных групп (например, для респондентов различного пола).

8. Выявление респондентов, профили ответов которых имеют аберрантный характер и вносят искажения в функционирование шкалы.

9. Исследование психометрического качества опросника в целом.

Все эти этапы нашли отражение в статьях сборника, описывающих использование IRT. Наиболее подробно методология исследования представлена в статье И.В. Антипкиной и в статье Т.Н. Канонир, И.Л. Углановой, Д.А. Федерякина.

Разумеется, для проведения анализа с применением IRT необходимо специализированное программное обеспечение. Одной из самых популярных программ для анализа с помощью модели RSM является программа Winsteps. Однако в настоящее время исследователи все чаще для анализа используют бесплатный статистический пакет R. При наличии минимальных навыков программирования этот пакет позволяет проводить самые разные виды анализов, используя готовые скрипты, разработанные другими исследователями.

Следует отметить, что IRT-моделирование — молодое направление в психометрике, и основная литература представлена на английском языке. Поэтому русскоязычный тезаурус еще не устоялся; в основном, специалисты используют англоязычные термины. Этим объясняется тот факт, что в некоторых статьях сборника авторы используют собственные переводы некоторых терминов (которые могут не совпадать в разных статьях), сопровождая их английскими оригиналами.

В целом, в данном тематическом номере мы постарались показать разнообразие современных исследований в области психодиагностики и особенно нам хотелось обратить внимание уважаемых читателей на возможности IRT-моделирования при конструировании и анализе психодиагностических методик.