

## К ПРОБЛЕМЕ КОМПЬЮТЕРНОЙ ДИАГНОСТИКИ В ОБРАЗОВАНИИ

В статье рассматриваются возможности и ограничения использования применения компьютера в психодиагностике. Автор анализирует пути становления компьютерной психодиагностики в зарубежной и отечественной психологии. Отмечаются целесообразность использования компьютерных диагностических программ для получения психологического знания. Описывается собственный опыт применения компьютерной психодиагностики. Предлагаются алгоритмы создания адаптивных тестов как разновидности компьютерной диагностики.

Внедрение информационных технологий в психодиагностику имеет свою историю. С начала 60-х годов прошлого столетия компьютер выполнял функции фиксации элементарных реакций и статистической обработки данных. Первоначально компьютер был вспомогательным инструментом исследования, выполнявшим трудоемкие операции.

С развитием информационных технологий и острой потребности в оперативном анализе массивов данных компьютер приобретает самостоятельную роль психодиагностических процедур в образовании.

В начале 70-х годов процедура обследования приобретает форму «диалога» с компьютером. С появлением обратной связи в «диалоге» с компьютером изменяется стратегия исследования. Появляются первые компьютерные тесты. Компьютеризированные тесты – это тесты, приспособленные к условиям компьютера.

Идеально для компьютеризированного психодиагностического исследования в образовании подходит зарубежное адаптивное тестирование. В различных психологических источниках адаптивное тестирование называется последовательным, разветвленным, специализированным, индивидуализированным, программируемым, динамическим или зависящим от ответа тестированием (А. Анастаси, Cleary Linn Rock, 1968, Lord, 1971).

Адаптивное тестирование может быть представлено множеством процедурных моделей, (De Witt, Weiss, 1974, Lachin, Weiss, 1974, Weiss, Betz, 1973). Один из вариантов адаптивного тестирования – двустадийное тестирование с тремя уровнями измерения. В целом тест содержит 70 заданий, из которых 10 заданий самой разной степени трудности, 20 заданий – высокой трудности, 20 заданий – средней трудности, 20 заданий – легкой степени трудности. Испытуемому предлагается 10 заданий самой разной степени, в зависимости от успешности выполнения этих заданий предлагается один из трех последующих вариантов. Испытуемый выполняет только 30 заданий из теста, содержащего 70 заданий. Выполнение пирамидального адаптивного теста начи-

нается с заданий средней трудности. Если испытуемый справляется с этим тестом, ему предлагается следующее по степени трудности задание, если не справляется, то предлагается более легкое задание. Процедура повторяется после каждого ответа, пока испытуемый не даст 10 ответов. Такая процедура является примером 10-стадийного теста. Во время этой процедуры испытуемому предъявляется 10 из 55 входящих в тест заданий (А. Анастаси).

Зарубежная образовательная практика располагает двумя видами адаптивных тестов: двустадийный тип исполнения и пирамидальный тип исполнения. Адаптивные тесты реализуются как в форме «карандаш – бумага», так и на базе компьютера.

По утверждению А. Анастаси, более сложные модели, не предусматривающие заранее установленного, фиксированного порядка предъявления заданий, допустимы для реализации только в форме компьютеризированного адаптивного тестирования (Embretson, 1992, BF Green, 1983, Wainer et al, 1980).

В основу этих процедур компьютеризированного адаптивного тестирования (КАТ) положены определенные алгоритмы теории «задание-ответ» (IRT), которые используются для составления комплекта заданий, проведения тестирования конкретных испытуемых и подсчета индивидуальных показателей.

Для каждого задания теста определяется оценка способности, выводимая для его выполнения с вероятностью 0,50. Эта оценка способностей является показателем, который индивидуум получает за правильное выполнение данного задания. Этот показатель является индикатором сложности задания, различительной способности и вероятности угадывания правильного ответа для данного задания. Каждое задание обладает индексом точности измерения, или информационной функцией теста. Этот показатель представляет собой сумму информационных заданий и исполняет роль традиционной ошибки измерения (SEM).

После ответа испытуемого на каждое задание компьютер предоставляет для него следу-

ющее задание с учетом всех полученных ранее ответов. Добавление новых заданий в процессе тестирования продолжается до тех пор, пока информационная функция теста не достигнет заранее установленного стандарта.

При обследовании всех испытуемых выборки достигается одинаковый уровень точности измерения. Показатель отдельного испытуемого способен определить не только количество правильно выполненных заданий, но и уровень трудности, а также другие психометрические характеристики этих заданий.

Совокупный тестовый показатель выводится на основе оценок способностей, соответствующих каждому выполненному заданию. Эта оценка способностей исправляется и уточняется по мере добавления каждого нового задания до тех пор, пока не достигается заданная точность измерения.

Такие показатели обладают универсальным качеством сопоставимости у всех испытуемых независимо от специфических форм набора заданий. Зарубежная образовательная практика имеет опыт упрощения процедуры конструирования инструмента КАТ за счет использования ряда доступных компьютерных программ. Исследования показали, что индивидуализированное компьютерное адаптивное тестирование может давать такие валидные и надежные результаты, как и общепринятые тесты, при наличии гораздо меньшего числа заданий. Эти исследования показали, что адаптивное исследование обеспечивает высокую точность измерения, что корреляции между правильно сконструированными КАТ формами и их бланковыми аналогами («карандаш - бумага») почти такие же высокие, как и коэффициенты надежности большинства тестов.

Результаты этих исследований убеждают в том, что одни и те же конструкты можно измерить при помощи обеих форм тестов.

В отдельных ситуациях тестирования КАТ не подходит (тесты скорости и скрининг-тесты). Разновидности этих тестов распределяют испытуемых по группам на основе критического показателя (Wainer, 1993). КАТ же имеют более сложные психометрические характеристики.

Исследователи Green, Bosh, Humphreys, Linn, Reckase (1984) разработали техническое руководство по оцениванию инструментов КАТ.

Адаптивное тестирование наиболее уместно для использования индивидуализированных программ обучения. Учащиеся и студенты проходят учебный предмет в удобном для себя темпе, затем определяется степень усвоения содержания учебной дисциплины, предполагается, что выполнение тестовых заданий по трудно-

сти может значительно отличаться от всех представителей выборки испытуемых.

Компьютеризированное тестирование по типу КАТ позволяет прекращать проверку, как только ответы испытуемых дают достаточно информации для принятия решения об уровне овладения предметом (А. Анастаси).

В зарубежной образовательной практике, в частности в США, интенсивно исследуются возможности адаптивного тестирования (Smittle, 1990, Ware, Kline, Flaugh, 1986). Подбор заданий индивидуализирован, исполнение его не занимает много времени и позволяет сразу же получить оценку. Этот тест можно проводить в ходе регистрации поступивших и сразу же распределять студентов по курсам или группам, соответственно полученным результатам испытаний.

Американская психодиагностическая практика использует КАТ в крупномасштабных программах отбора и распределения персонала в промышленности, государственных учреждениях, армии. КАТ хорошо подходит для этих целей, так как дает возможность обследования значительных групп кандидатов, охвата широкого разброса уровней способностей. Тест имеет высокую степень защиты, так как каждый кандидат получает разный набор заданий из соответствующего банка заданий, хранящегося в памяти компьютера.

Дифференциальные тесты способностей (Differential aptitude Tests – DAT). DAT включает в себя восемь субтестов: вербальное рассуждение, числовое рассуждение, абстрактное рассуждение, перцептивная скорость и точность, пространственные отношения, механическое рассуждение, орфография и словоупотребление. DAT используется в профессиональном и информационном консультировании учащихся 8–11 классов, на сегодняшний день создан тренировочный тест, охватывающий все восемь областей. Создана специальная форма DAT дифференциальных тестов способностей для оценки персонала и карьеры, где каждый из восьми субтестов сокращен и представлен в отдельной брошюре. DAT можно эффективно использовать в компьютеризированной версии.

Более современной разработкой этой батареи является форма компьютеризированного адаптивного тестирования (КАТ – ДАТ) (1987). В этой версии каждый испытуемый получает только те задания, которые соответствуют его уровню выполнения теста. В этой версии КАТ использованы задания более ранней версии DAT, которые были проанализированы на основе модели РАША – упрощенной, однопараметрической модели теории «задание – ответ».

Версии КАТ дифференциального теста имеют явные преимущества. Экономичность во времени, способность исследовать большие количественные группы испытуемых, возможность получения индивидуальных показателей – все это, несомненно, делает данное направление психодиагностики очень перспективным. В исследованиях Embretson (1992) проанализированы несомненные преимущества описанного выше методологического подхода.

Отечественная компьютерная психодиагностика как направление исследований оформляется в начале 80-х годов.

Ее развитие достаточно противоречиво. Психологическая наука не является активным потребителем информационных технологий. О.К. Тихомировым в начале 90-х годов XX века было проведено впервые в нашей стране исследование об отношении различных социальных групп к компьютеризированному тестированию.

Результатом исследования явился тот факт, что позитивно к такому виду исследований относятся руководители высшего уровня, разработчики компьютеризированных тестов, лица, имеющие навыки в работе с компьютерами, а также лица, имеющие развитые мотивы самопознания.

Негативное отношение к компьютерному тестированию выразили руководители среднего звена, а также лица, усматривающие зависимость личного благополучия в профессиональной деятельности и карьере от результатов тестирования.

Компьютеризированные методики в сравнении с традиционными имеют ряд достоинств. Их использование позволяет обеспечить постоянство условий тестирования, точность и однозначность регистрации множества возможных реакций испытуемых. Экспериментатор имеет возможность восстановить и проследить последовательность действий испытуемого. Единые банки эмпирических данных позволяют создать нормативы для различных групп обследуемых.

Психолог освобождается от сложной рутинной работы при конструировании теста и при проведении обследования. В практической деятельности психолога открываются возможности расширения группового тестирования и применения мощного статистико-математического аппарата при обработке банка данных.

Степень сохранения конфиденциальности результатов тестирования, применение экспресс-методик значительно облегчается и упрощается. Возможность устранения негативных воздействий тестовых исследований, вызванных неадекватными межличностными взаимо-

действиями экспериментатора и испытуемых, при компьютеризации процедуры тестирования выше. Возможности компьютерного тестирования расширяются за счет предъявления стимульного материала в виде обучающей игры. Обратная связь устанавливается достаточно быстро. Многие практические задачи возможно решить только через подобную процедуру.

Важно отметить, что исследования Хортмана (1986), О.В. Доротиной (1993) подтвердили то, что компьютерные исследования в психодиагностике приводят к появлению психологических новообразований в системе взаимодействия «экспериментатор – тест – испытуемый». Психологическим новообразованием явился «феномен компьютерной тревожности» у 30% обследуемых (причем у 5% обследуемых встречается состояние, схожее с фобией). Ряд тестов невозможно перевести в компьютерный режим. Любая автоматизированная версия может рассматриваться как новая психодиагностическая методика, которая в свою очередь нуждается в процедуре стандартизации. Надежность компьютерной версии необходимо доказать соответствующими выборками, в связи с тем, что стабильность результата как основной признак надежности диагностических методик значительно снижается при компьютерном предъявлении стимульного материала (Л.Ф. Бурлачук).

Исследования О.К. Тихомирова и Л.П. Гурьевой (1999) позволили выявить степень психологических последствий компьютеризации психодиагностики на структурных компонентах деятельности (мотивов, целей, операций) и сделать сопоставление с традиционными ее формами.

Авторы исследования разработали критерии дифференцированной оценки психологических последствий компьютеризированной психодиагностики в образовании, обладающие противоречивыми характеристиками по отношению к основному критерию значимости последствий для повышения продуктивности деятельности.

Эти последствия классифицированы авторами на реальные и потенциальные, позитивные и негативные, управляемые и неуправляемые.

Параметрами исследования явились мотивация, целеполагание и операционный состав компьютеризированной психодиагностической деятельности. Авторами выделены основные детерминанты: особенности социального заказа, развитие научного знания, специфика составления алгоритмов и программирования деятельности, техническое обеспечение работы компьютеров, личностные особенности психодиагностов и обследуемых. По мнению авто-

ров, изменение детерминант выступает в качестве принципа управления последствиями компьютеризации.

Широкое распространение доступных систем психологического тестирования на основе персональных компьютеров и доступа в сети Интернет по результатам многих исследований не способствует преодолению сложившегося в психодиагностике разрыва между теорией и практикой. По-прежнему теоретические проблемы психодиагностики не являются предметом исследования, очень медленно разрабатываются исследовательские процедуры, связанные с компьютеризированной психодиагностикой.

Основным итогом современной психодиагностики является интенсивное развитие известных формализованных методик. В отечественной образовательной практике компьютерная психодиагностика выполняет единственную функцию экстенсивного воспроизводства формализованных методик с целью итогового контроля знаний при отсутствии профессионально-этических требований за этим процессом.

Валидность известных высокоэффективных методик резко снижается, они фактически выводятся из профессиональной практики. Диагностика личностных особенностей в режиме диалога между испытуемым и компьютером остается проблематичной. Необходимость внедрения современных информационных технологий в психодиагностику образования очевидна. Однако при этом необходим углубленный психологический анализ последствий компьютеризации психодиагностической деятельности в образовательных целях.

В отечественной психодиагностике и образовательной практике отсутствуют теоретические исследования и практические разработки компьютеризированных версий адаптивного тестирования.

Мы полагаем, адаптивное тестирование способно создать в отечественной психодиагностике и образовательной практике более объективные модели контроля знаний учащихся, студентов, слушателей.

Наши собственные разработки адаптивных тестов по ряду учебных дисциплин позволили сделать вывод о том, что этот подход позволяет на начальном этапе осуществлять дифференцированное обучение. Наше исследование

убеждает также в том, что адаптивное тестирование более экономично не только по времени его проведения, но и по времени, затраченному на его разработку. Времени, необходимого для создания входного, рубежного и выходного контроля, требуется в 2–3 раза больше, чем времени, затраченного на создание самого адаптивного теста.

Автоматизированные обучающие системы позволяют управлять процессом тестирования в форме диалога «ЭВМ - учащийся», собирать, адаптировать и хранить разнообразную информацию о работе каждого испытуемого, проводить оперативно статистико-математическую обработку полученных результатов. С помощью компьютерной психодиагностики можно получить исключительно важные данные: время получения правильного или ошибочного ответов, время обдумывания решения, время ввода ответа, количество отказов от решения заданий теста, обращения за помощью. При компьютерном тестировании исследователь может получить объективную информацию не только об особенностях выполнения определенных заданий группой испытуемых, но и об их индивидуальных характеристиках (показателях умственного развития, скоростных особенностях деятельности, динамики работоспособности, вербальных и невербальных способностей).

В отечественной практике существует некоторый опыт компьютерной диагностики особенностей умственного развития, восприятия, внимания. Наибольший интерес представляют работы по созданию диалоговых программ специально для целей психодиагностики, точнее, диагностики становления смыслового и ценностного сознания.

Преимущества компьютерной диагностики в образовании бесспорны. Это позволяет ставить вопросы о необходимости разработки методов компьютерного тестирования. Однако существует ряд нерешенных вопросов. Не разработаны механизмы переноса разрабатываемых и оправдавших свое применение целого ряда диагностических методик в образовательной практике по компьютерному обеспечению. Вопрос, определяющий степень сравнения автоматизированного и традиционного тестирования, представляет большой интерес для исследователей компьютерного тестирования.

#### Список использованной литературы:

1. Абульханова-Славская К.А., Воловикова М.И., Елисеев В.А. Проблемы исследования индивидуального создания // Психологический журнал. 1999.
2. Анастаси А., Урбина С. Психологическое тестирование. 7-е международное издание. СПб.: Питер, 2003.
3. Дюк В.А. Компьютерная диагностика. СПб.: Братство, 1994.
4. Забродин Ю.М. Методологические проблемы психологического анализа и синтеза человеческой деятельности // Вопросы кибернетики: Эффективность деятельности оператора, М., 1982.
5. Клайн П. Справочное руководство по конструированию тестов. Киев, 1994.