

История

ПЕДАГОГИЧЕСКАЯ ДИАГНОСТИКА ЗА РУБЕЖОМ ВО ВТОРОЙ ПОЛОВИНЕ XX ВЕКА. Расширение применения тестов в образовании

Е.А. Михайлычев, Г.Ф. Карпова, Е.Е. Леонова

С 60-х годов в Западной Европе и в США ряд левых организаций требуют отмены обязательного тестирования в школах (выступают против использования тестов интеллекта, определяющих интеллектуальный коэффициент для отбора детей по потокам). Тесты критикуются в основном за их недостаточную валидность и социокультурную зависимость от первоначальных элитарных выборов, на которых они были изданы. Такая критика учитывается при валидации тестов, но она не мешает массовому тестированию и в системе образования, и при профотборе, аттестации кадров во всех сферах социальной и профессиональной деятельности — от конторских клерков до менеджеров.

В США более полувека существует служба педагогического тестирования. В её задачи входит обобщение и расширение практики тестирования, создание банков диагностических данных об учащихся и анализ этих данных, разработка стандартизированных тестов и их апробация. Служба устанавливает правила тестирования и следит за их выполнением. Она расположена в университетском центре Принстона, имеет в штате около 2300 человек. Из них уже в конце 70-х годов 150 имели степень доктора и около 200 магистра по различным дисциплинам — философии, педагогике, психологии, статистике. Служба обеспечивает компьютерную обработку и оценку результатов тестирования, имеет сеть из

1
Тесты и вопросы по географии. Под ред. К.В. Пашканга. М.: Просвещение, 1995.

2
Параил В.А. Высшее техническое образование в США: Критический анализ. Киев; Одесса: «Высш. школа», 1980. С. 44–45.

3
Ингенкамп К. Педагогическая диагностика. М.: Педагогика, 1991; *Ковалёва Т.Г.* Экспериментальная общая школа в ФРГ / Проблемы обучения и воспитания в зарубежных странах. М.: НИИ общ. пед. АПН СССР, 1972. С. 342–356.

почти 5000 пунктов тестирования по всей стране. Консультанты службы «Чайдж» в округах и школах активно участвуют в её работе¹.

Тесты играют огромную роль в высшем образовании США (наряду с другими методами контроля знаний), в которые они довольно органично вписываются. В подавляющем большинстве высших учебных заведений США принята особая система учёта учебной работы студента — зачётная единица или зачётный час. Это означает в среднем три часа учебной работы студента каждую неделю всего семестра или, учитывая, что семестр состоит из 15 недель, 45 часов учебной работы. В это время включаются как часы аудиторных занятий (лекции, практические и лабораторные), так и время самостоятельной внеаудиторной работы. При планировании учебного процесса исходят из того, что на час лекции или практических занятий студенту для освоения материала надо затратить два часа самостоятельной работы, а лабораторный практикум домашней работы не требует. Итого получается, что в одну зачётную единицу входит либо час лекции (практического занятия) в неделю, либо три часа лабораторного практикума, либо домашней работы. На основе такого подхода определяется объём учебного курса, а при установлении требований к получению той или иной учёной степени — минимальное число единиц, которые должен набрать кандидат. В

зачётных единицах измеряют нагрузку преподавателей и департамента (кафедры), её учитывают при финансировании вуза². Широко используются тесты не только в традиционном предметном обучении, но и в интегративных курсах.

Тесты контроля знаний и учебных достижений проводятся на Западе не только в учебных заведениях. Так, в 1966 г. Германская торгово-промышленная палата провела специальное тестирование 2000 учеников на предприятиях, окончивших полгода тому назад основную школу для выявления реального уровня образованности. Полученные результаты показали, что один из пяти имеет неудовлетворительные знания по правописанию, один из шестерых вполне владеет четырьмя арифметическими действиями, а знания каждого четвёртого по арифметике просто неудовлетворительны³. Результаты подобных исследований привели к созданию при Германском совете по вопросам образования специальной подкомиссии по проблемам способностей, развития и отбору особо способных детей.

В Европе тесты занимают в педагогической практике не столь могущественные позиции, как в США. К. Ингенкамп отмечает, что в 1960 г. школе было предложено всего 74 индивидуальных и групповых теста, однако в большинстве своём они не удовлетворяли методическим критериям. В 1970г.

«вся немецкоязычная территория обладала 122 тестами и анкетами, определяющими школьную успеваемость, подготовленность детей к занятиям в школе, умственное развитие, пригодность к профессиональному обучению, способность концентрироваться. Это довольно мало, если учесть, что они обслуживали все 13 школьных классов общеобразовательной школы, все учебные предметы и прочие учебные курсы. Использование, достойное упоминания, эти тесты получили лишь в двух названных выше сферах, а также в экспериментальных школах. На занятиях в школах второй ступени, прежде всего в гимназиях, тесты применялись крайне редко». К 80-м годам количество тестов и анкет, применявшихся в Германии, возросло до 222, но из-за критики тестологии, частота их применения резко падала⁴.

В Англии в массовой школе ведётся курс «Трудовое обучение для профессиональной карьеры» профориентационного характера. Его преподаватели получают на специальных отделениях педагогических институтов хорошую практическую диагностическую подготовку, обеспечены специальным кабинетом со средствами наглядности и оборудованием. В их задачи входит разработка и ведение диагностической документации об интересах и склонностях учащихся, их психофизических возможностях. На еженедельных (1–2 часа) обязательных занятиях

учащиеся заполняют различные тесты и анкеты, участвуют в дидактических играх, выбирают предметы для дальнейшего обучения⁵. Обработка проводится оперативно, с помощью компьютеров, наличие которых в учебных классах позволяет выполнять тесты и заполнять анкеты прямо на компьютеризированном рабочем месте с последующей автоматической обработкой, предъявляя учащимся их результаты и рекомендации по корректировке «западающих» качеств.

Дидактическое тестирование широко распространено не только на Западе, но и в странах «третьего мира» с западной ориентацией систем образования.

Согласно закону 1968 г. в Египте в первых трёх классах начальной школы проводятся ежемесячные устные и письменные тесты по всем предметам и в конце учебного года выводятся средние оценки, на основе которых учащиеся переводятся в следующий класс. Начиная с 4-го класса, ученики в дополнение к ежемесячным тестам сдают экзамены в середине и в конце учебного года и по их результатам переводятся в следующий класс — минимальный переводной балл по всем предметам — 40%, за исключением искусства — 20%, основ религии и арабского языка — по 50%⁶. Даже в Республике Чад в начале 70-х годов отбор в старшие (4–6-е) классы начальной школы осуществлялся с помощью тестов,

4

Ингенкам К. Педагогическая диагностика. М.: Педагогика, 1991. С. 15–16.

5

Саймон Брайан. Общество и образование. Пер. с англ. / Под. ред. В.Я. Пилиповского. М.: Прогресс, 1989; *Сергеева Н.И.* Обучение одаренных детей в школах Великобритании // Сов. педагогика. 1990. № 6. С. 137–144.

6

Мельник В.В. Современная общеобразовательная школа Арабской Республики Египет / Проблемы обучения и воспитания в зарубежных странах. М.: НИИ общ. педагогики АПН СССР, 1972. С. 110–111.

причём тестирование проводили не специалисты, а учителя начальных школ, многие из которых сами имели лишь начальное образование⁷.

Тесты широко распространены в Турции, где создан центр тестирования, аналогичный США; широко они используются в Израиле на всех стадиях обучения. Тестирование распространено в странах Балтии и в Узбекистане, где созданы научно-методические центры тестирования и ведётся активная их разработка.

Типология тестов и тенденции их разработки

Большинство современных авторов выделяют следующие группы тестов:

- психодиагностические тесты (диагностики функциональных состояний, диагностики психических процессов, сенсомоторные, тесты способностей, характерологические);

- дидактические тесты, среди которых выделяются: тесты обучаемости, тесты учебных достижений. Все тесты подразделяются на нормативные — ориентированные на сопоставление данных обследованного ученика с группой подобных ему учеников и критериальные — характеризующие степень достижения поставленной учебной цели (знания, умения) каждым учеником. Грань между труппами тестов довольно подвижна,

она во многом определяет и технологию разработки методики, её специфические возможности и ограничения — ведь тесты отнюдь не всеильны. Как всякая методика, они хороши при уместном и умелом применении. Западные исследователи и разработчики это осознали и доказали в многочисленных профессионально поставленных экспериментах, чей опыт для нас, отставших более чем на полвека, весьма ценен. Предпочтение в последнее время отдаётся критериальным как ориентированным на учебную цель. Они сложнее в составлении и обработке, но — более диагностичны и информативны для учителя, а при включении в диагностико-коррекционную программу могут играть тренинговую роль.

Немецкие специалисты по психолого-педагогической диагностике считают, что «тесты, ориентированные на учебную цель, обеспечивают сопоставление достижений индивида с учебной целью. Они построены на основе идеальной нормы, причём идеальная норма идентична учебной цели»⁸. Каковы основы их построения?

В критериальных тестах, диагностирующих овладение навыком, оценка выполнения заданий проводится по принципу «Все или ничего», то есть итоговый показатель фиксирует либо наличие, либо отсутствие навыка на заранее установленном уровне (обычно в 80–85% всех заданий, а не 70, как

7

Борисенков В.П. Актуальные проблемы начальной школы в странах Центральной Африки (Народная Республика Конго, Республика Габон, Центральная Африканская и Чад) / Проблемы обучения и воспитания в зарубежных странах. М.: НИИ общей педагогики АПН СССР, 1972. С. 148.

8

Ингенкамп К. Педагогическая диагностика. М.: Педагогика, 1991. С. 42.

предлагает В.П. Беспалько). Применяется и трёхступенчатая шкала: положительная, отрицательная и промежуточная оценка (вроде привычных для советской педагогики высокого, средней и низкого уровней — использовавшуюся в педагогических консилиумах Ю.К. Бабанского). Внутренняя структура, композиция теста может быть различна. Но вариации выбора у разработчика невелики. Специфичны по своему построению и весьма разнообразны очень популярные у педагогов тесты обучаемости, обладающие нередко большими прогностическими возможностями.

Выделяются две стратегии, имеющие несколько вариаций — обусловленные ответом: секвенционная и адаптивная. В секвенционных тестах задания зависят от предыдущего ответа и предлагаются из гомогенного (по замыслу) фонда, в принципе неограниченного. Их последовательность определяется выполнением заданий, предложенных в последнюю очередь. Это происходит до тех пор, пока с помощью заранее вычисленной вероятности не будет установлено, что необходимый уровень развития диагностируемых знаний, умений или качеств личности, определённый экспертами, достигнут. Такая тестовая стратегия используется для отбора по принципу «цель достигнута — не достигнута» и не позволяет проводить ранжирование обследуемых.

В адаптивных тестах, также зависящих от ответа обследуемого, помимо количества выполненных заданий учитывается дополнительно их качество, необходимое для дифференцированной оценки тестируемого. Задания имеют различную степень сложности и располагаются на различных ступенях пирамиды. Только от предыдущих успехов, подготовленности тестируемого зависит, с какой выборкой заданий он будет работать — при встрече со слишком сложными для него заданиями он может переключиться на другую программу, другой, более приемлемый уровень тестирования. В результате появится возможность более дифференцированно определить достижения, позицию тестируемого по диагностируемым показателям. На этом принципе строятся профессионально сделанные компьютерные диагностико-коррекционные программы⁹.

На практике реально доминируют уже наработанные, привычные, имеющие разветвлённые нормативы и прошедшие стандартизацию нормативные тесты, ориентированные, преимущественно, на тестовую модель Раша. Это стохастическая модель теста, разработанная в 1960 году Г. Рашем, в соответствии с которой считается, что вероятность ответа испытуемого на тестовую задачу является аддитивной или мультипликативной функцией трудности задания («сила») и интенсив-

ности свойства (способности) — т.е. разности уровня способности испытуемых решать задания. Шкалы Раша представлены как шкалы отношений, т.е. самые высокие уровни шкал, позволяющие производить все измерительные операции. Модель Раша является первичной моделью IRT — теории конструирования тестов, исходящая из методологии латентно-функционального анализа. Модель Раша однопараметрическая. В ней отношение трудности заданий не зависит от способности испытуемых. При реализации этой, весьма распространённой модели, неоднократно проводится эмпирическая проверка теста. Главным недостатком модели теста считается недостаточное внимание дифференцирующей силе тестовых заданий и то, что она не учитывает субъективные факторы тестирования.

Современная зарубежная тестология и то, что имелось в довоенный период, могут сравниваться только как ребёнок и взрослый молодой человек (резервы роста велики). По сравнению с 30-ми годами, когда выдвигались предложения разрабатывать тесты обучаемости, построенные на подражании и повторении, в 70–80-е годы произошёл качественный рывок в теории тестологии. Предпринимались попытки «выявлять обучаемость за счёт создания стандартизированного обучающего импульса в ходе процесса тестирования и последующей регистра-

ции обучающего эффекта»¹⁰. Суть современного подхода к тестам обучаемости заключается в том, что эти тесты «должны не только констатировать выигрыш за счёт обучения, но и прояснить процессы учения и мышления, лежащие в основе этого выигрыша»¹¹. Ю. Гутке и У. Волраб¹² выделяют две большие группы тестов обучаемости — кратковременные и долговременные.

Ведущие исследователи в области педагогической и психологической диагностики считают, что отдельные тесты и даже батареи тестов — вчерашний день психолого-педагогической диагностики, пройденный этап. Будущее, по их мнению, за диагностико-коррекционными программами, в которых тест является составной частью. Он выполняет функции входной, текущей и итоговой диагностики. Причём нередко — это тесты разного типа, разного уровня. Между ними «проложены», включены корректирующие воздействия педагога или компьютерной программы (либо же какого-то тренингового устройства, учебных пособий с заданиями для самостоятельной работы).

Такая технология, ведущая из опыта программированного обучения и контроля (с одной стороны) и опыта терапевтической помощи и коррекции, в т.ч. — психокоррекции — с другой строится на междисциплинарной (медицина-психология-педагогика-эти-

10

Психодиагностика: теория и практика. Пер. с нем. Общ. ред. Н.Ф. Талызиной. М.: Прогресс, 1986. С. 153.

11

Там же. С. 168.

12

Мельникова Е.А. Меч и лира: Англосаксонское общество в истории и эпосе. М.: Мысль, 1987. С. 154.

ка) методологии человековедения. Опыт возвращает нас к мысли о необходимости и актуальности возрождения педологии как целостной междисциплинарной области познания, образования и развития ребёнка. (Разумеется, с учётом ранее допущенных ошибок и профилактикой вульгаризации этой области познания и деятельности неопитами-энтузиастами).

Диагностические программы в принципе могут быть построены по-разному. Но обычно в диагностических программах соблюдается определённая системная последовательность предъявления задач — с очень простых, становящихся всё более сложными и комплексными (нередко предлагаются после разъяснения или демонстрации способов решения задач экспериментатором. Не исключён здесь и противополож-

ный подход, когда «вначале предъявляется комплексная задача, а систематизация задач производится сверху вниз»¹³. Это распространённая практика композиции, хотя хороших программ, апробированных и широко адаптированных по всей строгости правил стандартизации диагностических подходов, весьма мало. Призывов к их созданию много (как в Век Просвещения к изучению ребёнка). Реально сделано намного меньше.

В современной педагогической диагностике (и, особенно, в более развитой её части — педагогической психодиагностике) накопилась большая группа хорошо осознанных вопросов методологии педагогической диагностики, есть обширная зона для поисков и отработки различных вариантов технологии педагогической диагностики и коррекции.