

ОБ ОСОБЕННОСТЯХ КОНСТРУИРОВАНИЯ СОДЕРЖАТЕЛЬНО-ОРИЕНТИРОВАННЫХ ТЕСТОВ

В статье рассматриваются основные проблемы нормативности содержательной диагностики. Анализируются основные алгоритмы конструирования тестов. Представлены основные линии анализа выполнения теста. Изложен авторский вывод ограниченности содержательной диагностики.

Содержательная диагностика не так нова и не так отлична от диагностики, ориентированной на статистические нормы. Оценки индивидуального выполнения теста, выраженные в абсолютных единицах или процентах правильных ответов, несомненно, намного старше нормативной интерпретации. Еще до введения термина «критериально-ориентированное тестирование» Ибель (1962), Jitko (1984) описывали выполнение теста с точки зрения содержания. Ибель (Ebel, 1972) заметил, что в педагогике понятие овладения (masteru) чем-либо по принципу «все или ничего» достигло значительной популярности в 1920-1930-х гг., но позднее от него отказались.

Нормативная основа имитационно присутствует во всяком тестировании, независимо от того, как выражаются показатели теста (Angoff, 1974, Jitkj 1984). Выбор содержания, навыков, умений, которые подлежат измерению, определяется знанием специалиста, каких результатов можно ожидать от людей на определенном уровне их развития или обучения. Такой выбор обусловлен наличием знаний о том, как другие действовали в подобных ситуациях.

Но устанавливая граничные показатели на континууме умения, содержательная диагностика обладает возможностью получать значения об индивидуальных различиях.

Содержательная диагностика должна быть тесно связана с педагогическим процессом: педагог и психолог должны быть связаны общими интересами, общей целью – поисками наилучшего пути для оптимального развития учащегося, студента. Необходим выбор таких задач для создания содержательных тестов, которые поддаются четкому определению, чтобы можно было установить, какие мыслительные операции должен выполнить учащийся, студент; в каком месте встретились трудности, которые привели к ошибкам, причины и значения этих ошибок.

Если учесть, что один и тот же результат может быть следствием различных процессов мышления, то оказывается очень сложным на основе анализа ошибок судить о процессах, которые привели именно к данному ответу. А значит существует

необходимость в точных знаниях объективной структуры требований к предлагаемым в тесте задачам. Задания должны соответствовать нормативным требованиям базового уровня программы. Отбор критериального содержания мыслительной деятельности должен начинаться на уровне проекта: анализируются требования к сформированности отдельных мыслительных действий, предъявляемых обществом к изучаемому вопросу.

Задания должны представляться в определенной системной последовательности (по степени усложнения), т. к. известно, что сложность одной и той же задачи меняется в зависимости от ее места в тесте, отсюда выстраивается ступенчатая структура субтеста.

Во время диагностики и анализа результатов необходимо учитывать активность ориентировки, контактность, заторможенность, восприимчивость к помощи, работоспособность (утомляемость) каждого учащегося. Необходимо учитывать явления «горизонтального деколяжа» (Н.Ф. Талызина, В.Е. Карпов), суть которого в том, что ребенок может решать задачи высокого уровня трудности и не справиться с более легкими задачами (возможно, в их решении участвуют разные познавательные процессы и действия). Специальный подбор заданий позволяет выделить динамику продвижения обучающегося. Следует учитывать закономерность, что задание обладает наибольшей диагностичностью в том случае, если оно настолько сложно, что может быть решено лишь половиной представитель референтной группы, причем половиной, заведомо более способной к выполнению подобных заданий.

Инструкция к каждому заданию должна быть четкой. Необходимо убедиться, что учащийся верно понял инструкцию, иначе снижается надежность и важность результатов. Ответы должны быть однозначными. Исключение составляют задания, изучающие креативность мышления. Необходимо принимать во внимание ситуацию обследования (В.Н. Дружинин): во время тестирования обстановка не должна быть напряженной, а наоборот, доверительной. Необходимо избегать слов «тест», «проверка». Их можно заменить словами «упражнения».

У каждого учащегося, студента должны быть тестовые задания, бланки ответов и ручки. Оптимальный состав 15-36 человек. Если обследуемых больше, необходима помощь ассистента (ассистентов).

При разработке содержательно-ориентированного теста важно учитывать и большой опыт, накопленный в рамках широких формирующих экспериментов (П.Я. Гальперин, В.В. Давыдов, Л.Н. Занков, Н.Ф. Талызина, Д.Б. Эльконин), где в течение длительного времени выявляются и диагностируются характеристики учебной и умственной деятельности (в различных предметных областях).

Близость тестов к содержанию образовательных программ, использование учебного материала позволяют получить объективные данные об умственном развитии учащихся, студентов.

Нормативная модель мыслительной деятельности по решению критериальных задач строится в результате эмпирической и теоретической рефлексии специалистов (методистов, психологов, педагогов) по поводу анализа различных путей их решения. Нормативная сложность задачи определяется методом экспертных оценок. Параллельно с этим целесообразно использовать алгоритмический подход, который является для психологических исследований в этом плане достаточно популярным (Г.А. Балл, О.К. Тихомиров), т. е. нормативная сложность задачи определяется количеством операций в квазиалгоритмическом способе ее решения.

Субтесты должны содержать не индивидуальные, а родовые задачи, каждой из которых соответствует некоторый класс индивидуальных задач, что позволяет выявить общий принцип решения (алгоритм, алгоритмические предписания задач данного класса).

Все задания в содержательно-ориентированном тесте должны образовывать определенную систему, репрезентативную всему предметному содержанию конкретной области познания (в рамках конкретного возраста). Это во многом обусловлено тем, что «теоретическое мышление имеет свое объективное содержание, отличное от содержания эмпирического мышления».

При составлении тестов в диагностике развития необходимо учитывать несколько видов норм: предметную, возрастную, индивидуальную, социальную.

Вопрос о разработке одной общей нормы для оценки умственного развития не должен стоять, так как умственное развитие зависит от социально-культурных условий, в которых развивается каждый отдельный человек. Каждая норма имеет свои

преимущества. Нормы взаимно дополняют друг друга. В мотивационно-психологическом плане наиболее значима индивидуальная норма.

Предметная логика заданий выстраивается на анализе предметного содержания задачи.

Возрастная логика становления умственного развития – на основе периодизации возрастного развития. Возрастной норматив зависит от общественно-исторических условий, прежде всего от сложившегося типа обучения и воспитания.

В каждом обществе явно или неявно присутствуют определенные стандарты психического облика умственного и личностного развития детей различных возрастов. Причем диагностика должна быть направлена на установление как наличия социально-нормативных видов деятельности, так и их состояния.

Конкретно все сопоставления результатов тестирования следует проводить с тем максимумом, который представлен в самом тесте, а это есть общее число помещенных в нем заданий. Требования к умственному развитию, которые закладываются в такой тест, могут рассматриваться как модель норматива.

В диагностике развития необходимо учитывать особенности системы критериальных оценок.

Принятие нормативов не означает, что все сопоставления будут проводиться только в сравнениях того, насколько та или иная группа обследуемых близка к нормативам.

Сохраняют свое значение и другие количественные показатели.

Однако при разработке концепции тестов умственного развития придется пересмотреть систему критериальных оценок. В частности, нужно пересмотреть традиционное представление, что результаты психологического тестирования больших выборок якобы должны быть распределены по кривой Гаусса, т. е. нормально. Нужно отметить, что нормальное распределение имеет место, когда на случайную величину действует большое число разнообразных факторов и доля воздействия каждого из них одинаково мала по сравнению с их числом. В данном случае на распределение влияет наряду с множеством разнообразных факторов также один мощный фактор – фактор культуры. Распределение результатов тестирования будет в таком случае зависеть от того, в каких долях представлены в данной выборке лица с разными степенями приобщения к данной культуре, как она отражена в тесте. Так как подбор испытуемых не может быть заранее предсказан, то и о характере распределения заранее ничего сказать нельзя. Воз-

можно, что удастся когда-то получить нормальное распределение, но это будет не правилом, а исключением. Отсюда появляются трудности, например нет оснований всегда применять параметрические статистические методы. Причем распределение, характеризующее критериальный тест, должно постоянно изменяться, т. к. нужно учитывать результаты коррекционных занятий.

Нельзя более удовлетворяться и сложившимся пониманием критерия надежности, по которому качества теста тем выше, чем больше совпадений между первым и вторым тестированием (тест – ре-тест). Этот критерий несет в себе идею метафизической неизменности содержания психики, не допускает возможности ее развития. Получение высокого коэффициента надежности при повторном тестировании должно, скорее всего, восприниматься как сигнал неблагополучия: либо тест не отражает перемен, либо такие перемены и на самом деле не произошли, что свидетельствует о паузе в развитии.

Важным этапом процедуры построения теста, диагностирующего определенную стадию сформированности умственного развития, является определение его теоретической валидности, которая связана с выбором экспериментатором варианта обобщенной модели теста. В нашем случае теоретическая валидность должна определяться тем, насколько соответствует тест умственной деятельности, характерной для определенных предметов.

Не менее важным является и определение внешней валидности критериального теста. В качестве внешнего критерия должна и может выступать только социально обусловленная и социально актуальная умственная деятельность субъекта. В нашем случае такой деятельностью является учебная деятельность. Ее результаты отражают при всей их относительности школьные оценки. В качестве критерия используются не все оценки, а только те, которые содержательно валидны данному виду продуктивной умственной деятельности учащихся, студентов (т.е. оценки, поставленные за выполнение самостоятельных работ, связанных с определенной темой, соответствующей содержанию субтестов).

При определении содержательной валидности тест должен ориентироваться на внешние социально обусловленные и педагогически адаптированные требования. Однако сам теоретический критерий валидности должен быть психологическим, а также динамичным (т. е. различным для дифференцированных моделей обучения, реализуемых в рамках различных образовательных систем).

Мы считаем, что пересмотр методического аппарата в диагностике развития необходим. Поиск правил подбора материала для составления заданий в тестах находится в прямой зависимости от психологической интуиции авторов теста.

При конструировании, проведении и обработке критериально-ориентированных тестов существенное и принципиальное значение имеет «точка отсчета» в оценке достижений учащихся, студентов. За нее принимается полное овладение всеми умениями, необходимыми для выполнения тестового задания. Таким образом, диагностика умственного развития ориентирована на критерий, конкретизированный по отношению к определенному учебному содержанию, определенному классу задач, предъявляющих требования к умственному развитию учащихся. Критериальные тесты строятся на основе логико-функционального анализа содержания заданий, при этом анализируется некоторая предметная область (учебный предмет), выбираются базовые, необходимые способы действия, фундаментальные умения, формирование которых определяет специализацию умственного развития.

Предметом анализа становится содержание деятельности, т. е. то, что индивид знает или умеет делать по отношению к конкретным требованиям, которые предъявляет учащимся учебно-образовательная программа. Одним словом, критериально-ориентированный тест определяет уровень индивидуальных достижений относительно некоторого объективного эталона на конкретном этапе обучения.

Критерий (уровень выполнения, который должен быть воспроизведен каждым учащимся), или стандарт выполнения, всегда специфичен в силу специфичности содержания самих задач и умений. В качестве критерия обычно рассматриваются конкретные знания (система понятий, совокупность фактов), умения, навыки, необходимые для успешного выполнения критериальной задачи. Тем самым содержание критерия задается объективно, сложившейся конкретно-исторической системой требований к обучению.

Таким образом, критериально-ориентированные тесты своими содержательными и структурными характеристиками соответствуют конкретной ситуации обучения, т. е. определяются целями обучения и выступают в качестве оперативного средства контроля и оценки его результатов.

Разрабатываемая тестовая структура должна быть связана с диагностикой развития теоретического мышления как процесса. При этом диагнос-

тируется промежуточный результат в умственном развитии, который является характеристикой самого процесса мышления (для раскрытия закономерностей мышления используется процесс решения задач).

Причем необходимо отметить, что критериальное задание носит конкретно-исторический характер: с изменением областей человеческой деятельности (в том числе и учебной) изменяются как критериальные задания, так и подходы исследователя к их анализу.

Отсюда диагностика развития позволяет своевременно выявлять учащихся, студентов, чьи способности учебной работы не отвечают социально-нормативным требованиям умственного развития.

В зависимости от психолого-педагогических целей и конкретных принципов конструирования используют две формы критериально-ориентированных тестов:

1. Задания гомогенны, т. е. сконструированы на одной и той же или подобной содержательной и логической основе. Каждое направлено на выявление владения конкретным навыком усвоения определенной единицы предметного содержания, того или иного показателя умственного развития. Такие тесты имеют ярко выраженную дидактико-методическую направленность. Разработанные на материале образовательных программ, они являются эффективными инструментами контроля за формированием учебных знаний и навыков, оценки эффективности обучающих воздействий, стимулирования познавательной активности учащихся.

2. Задания гетерогенны и существенно различаются по логической структуре. Тест обычно имеет ступенчатую структуру. Каждая ступень характеризуется своим уровнем сложности, который определяется логико-функциональным анализом содержания, относящегося к критериальной области поведения. По каждому заданию на основании результатов можно предложить образец ответа или даже определить причины, обусловившие определенный характер ответа. Такие критериально-ориентированные тесты являются надежными инструментами диагностики и коррекции специфических трудностей в обучении. Следует обратить особое внимание на то, что если в таких тестах моделируется структура умственных действий или навыков, каждый из которых имеет свое особенное содержание и форму проявления в реальной учебной деятельности, индекс внутренней согласованности может быть нулевым или даже отрицательным, а тест тем не менее может оказаться диагностичным.

При критериальном подходе каждая из задач теста содержательно самоценная и не может быть заменена задачей из другого содержательного класса (хотя по степени сложности задачи одного уровня должны быть эквивалентными). Поэтому-то в подобных тестах пересматривается проблема валидности в ее традиционном тестологическом понимании: задания рассматриваются не как обезличенный внешний критерий, а приобретают новый статус и направлены на анализ реального совпадения заранее установленных логико-психологических показателей и результатов учебной деятельности. Таким образом, в основе конструирования критериально-ориентированных тестов лежат процедуры, связанные с предварительным отбором содержания, внутренне связанного с критериальным заданием. В связи с этим установление содержательного и структурного соответствия заданий теста и реального учебного задания выступает как решающий этап разработки.

Для этого составляется таблица спецификации, которая становится основанием для установления соответствия формы и содержания тестового задания конкретному содержанию и структуре учебной задачи.

В качестве приложения к спецификации могут быть включены описание внешних показателей критериального поведения, варианты тестовых заданий для разных выборок испытуемых, описание возможных альтернативных ответов на тестовые задания с их психологическим анализом и др.

Предварительно моделируются идеальные приемы познавательной деятельности (это проводится для того, чтобы в последующем можно было сравнить реальные учебные и умственные действия с нормативными действиями, идеальными приемами, определяемыми логикой учебного предмета, например математики).

После моделирования приема познавательной деятельности возможна диагностика усвоения приема в целом и отдельных его действий.

Принципиальным моментом в адекватном выборе критериальной области является тот факт, что умение, подлежащее психологическому анализу в тесте, находится у учащихся в стадии становления. Именно на этой стадии активно формируются и приобретают произвольность те умственные действия, которые опосредуют данное учебное умение. Здесь впервые возникают способы компенсации, которые позволяют учащимся справиться с определенными заданиями в том случае, если у них еще не сложился целостный комплекс необходимых умственных действий. Именно на начальном

этапе формирования умения закладывается база для таких способов решения, которые не относятся к мыслительной деятельности в исследуемой предметной области, например штампы, механическое воспроизведение заученного алгоритма решения определенного типа задач.

Существует две основные линии анализа выполнения теста:

1) анализ выполнения с точки зрения специфических операций, его составляющих: внимание уделяется составу умения и навыков, необходимых для решения конкретной учебной задачи, их описанию и систематизации. Однако выполнение учебных задач зависит не только от знаний и умений, но и от определенного уровня умственного развития;

2) анализ выполнения с точки зрения умственных операций, уровень владения каждой операцией проверяется отдельным субтестом.

В среднем в критериально-ориентированном тесте субтестов может быть 4-6. Ступенчатая структура субтестов и теста в целом позволяет не только выявить, владеет ли учащийся, студент определенной операцией, но определить конкретный уровень такого владения: насколько он обусловлен содержательными и структурными характеристиками предметного материала.

Кроме того, степень успешности выполнения субтестов позволяет судить об индивидуальном характере затруднений учащихся в актуализации той или иной умственной операции, а также приблизить исследователя к пониманию присущей каждому учащемуся субъективной логики, обусловившей способ выполнения. Способы выполнения следует изучать с точки зрения как соответствия логике возрастного развития, так и соответствия их конкретно-историческому содержанию мышления в данной предметной области.

Исследователь получает возможность наблюдать и анализировать затруднения учащихся, студентов, которые последние испытывают при выполнении заданий возрастающего уровня сложности. Причем одной из причин затруднений выступает несформированность требуемых умственных действий.

Содержательно-ориентированные тесты ставят учащихся в такую ситуацию, когда для того, чтобы выполнить требуемое, необходимо найти признаки, предполагаемые заданием, и логически их обработать. Причем когда учащиеся находятся на стадии неокончательной сформированности требуемого выбора, то возникают прочные ассоциации, всплывают признаки, существенные для иного по

специфике материала, происходит автоматизированное воспроизведение какого-то заученного штампа или этапа решения.

Исследования Е.И. Горбачевой (1991) показали, что несформированность у учащихся отдельных умственных действий еще не предопределяет роковым образом их неуспешность в выполнении учебных заданий. Все зависит от того, удастся ли учащемуся на основе наличествующих в его опыте умственных действий (или даже одного действия) найти способы компенсации, позволяющие справиться с заданием (например, использует неэкономный в логическом плане способ решения). Сочетание умственных действий не функционирует как простая арифметическая сумма. Сочетание действий – это их взаимосвязь и взаимообуславливание при доминировании одного из них.

Таким образом, проблема заключается в том, какое из умственных действий выступает как доминирующее, а возможно, и структуро образующее.

Качественный анализ актуализируемых в тесте умственных действий, анализ успешности выполнения заданий, относящихся к исследуемой критериальной области, позволяет судить о релевантности критериального задания реальному процессу выполнения учебных задач. В основу конструкции критериально-ориентированного теста могут быть положены два существенных показателя сложности учебных задач.

Сложность рассматривается как объективная категория, определяемая числом и характером свойств и отношений между элементами задачной ситуации и имеющая следующие показатели:

1) количественный (число конкретных данных, типов связей между величинами задачи, существенных взаимосвязей между данным и искомым);

2) качественный (характер зависимости между величинами, который может быть ярко выражен, т. е. открыт для прямого перевода на язык математических знаков, или не выражен вовсе, в последнем случае необходимы определенные опосредования).

Показатели сложности определяют ступенчатую структуру каждого субтеста:

1) простейшая структурная организация элементов задачи, число зависимостей между величинами минимально, и они однородны или почти однородны;

2) структурная организация элементов задачи достаточно осложнена дополнительными вводными данными, часть условий вынесена в вопрос задачи и др.

Ступенчатая структура заданий в тесте позволяет установить возможности учащихся в владении умственным действием на материале равной степени сложности.

Сформированность умственных операций определяется качеством выполнения (правильностью ответов) и устойчивостью (учащийся устойчиво может выделять общее и особенное в различных задачных ситуациях).

Суммарный балл характеризует интенсивность выраженности свойства, а не вероятность его наличия у обследуемого.

Наши исследования показали, что за пределами исследования уровня базовых знаний, умений и навыков содержательная диагностика может быть неприменима и недостаточна, ей могут быть недоступны более сложные и менее структурированные области знаний.

Конкретный человек может неограниченно совершенствовать такие компетенции, как понимание, критическое мышление, предчувствие, оригинальность, умение руководить, умение подчиняться, быть инициативным и др. Такие слагаемые знаний не всегда возможно подвергнуть измерению средствами содержательной диагностики.

Видимо, содержательная диагностика способна дополнять, делать более прогностическими показатели тестов интеллекта. Комбинирование показателей интеллектуальных тестов, ориентированных на статистические нормы, с показателями содержательной диагностики, ориентированной на социально-психологические и индивидуальные нормы, позволяет получить более адекватную информацию о мыслительной деятельности учащегося, студента, о причинах пробелов в учебной деятельности, о возможных вариантах коррекционной работы.

Список использованной литературы:

1. Анастаси А., Урбина С. Психологическое тестирование. Спб.: Питер, 2002.
2. Берулава Г.А. Психология естественного мышления. – Бийск, 1991.
3. Гуревич К.М., Горбачева Е.И. Умственное развитие школьников: критерии и нормативы. – М., 1992. – С.12-49.
4. Епанчинцева Г.А. Возможности содержательной диагностики // Одаренность: теория и практика. Материалы Российской конференции. Москва, 2001.
5. Ингенкамп К. Педагогическая диагностика / Пер. с нем. – М., 1991.
6. Равен Дж. Педагогическое тестирование. Проблемы. Заблуждения. Перспективы. М., 1999.
7. Психологическая диагностика / Под ред. К.М. Гуревича. Москва – Бийск, 1993. – С.7-117, С.241-285.
8. Beck I.L., McCaslin E.S., McKeown M.G. The Rationale and Design of a Program to Teach Vocabulary to Fourth-Grade Students. Pittsburgh, PA, 1980.
9. Clay M. Reading: The Patterning of Complex Behavior. Auckland, 1979.