

МЕТОДОЛОГИЯ ИСПОЛЬЗОВАНИЯ ТЕСТОВЫХ ЗАДАНИЙ, ПРИМЕНЯЕМЫХ ДЛЯ ОЦЕНКИ ЗНАНИЙ ОБУЧАЮЩИХСЯ

Г.А. Бейсенова

Мақалада тест тапсырмаларының әдіснамасы, тесті құру әдісінің тарихи дамуы қарастырылған, тесті құру кезеңдерінің сипаттамасы, сондай-ақ тест тапсырмасының сапасындағы критерийлердің бағалау әдісі беріледі.

This article touches upon the methodology of tests, historical development of testing methods, the characteristic of test stages, estimation methods of criteria quality of testing as well.

Тест – система заданий возрастающей трудности специфической формы, позволяющей объективно оценивать уровень и структуру знаний студентов. Единицами тестирования служат тестовые задания – это одна единица контрольного материала, сформулированная в виде утверждения, предложения с неизвестным, удовлетворяющая ряду требований. Тестовое задание – задание в тестовой форме, прошедшее электронную проверку и обладающее определенными оценками. [1]

Для научного создания тестов необходимы:

1. Методология, то есть теория методов создания тестов. Традиционно в тестировании сложились две методологии:
 - методология создания тестов до 20-60 г.г. 20 в.;
 - современная методология, называемая латенто – структурный анализ.
2. Теория педагогических измерений. Для выделения предмета измерения вначале необходимо концептуально решить – «что такое знание».
3. Педагогическая теория теста (некий понятийный аппарат, формы задания, методы отбора содержания и т.д.).

Предтечей теста школьной успеваемости были появившиеся в 1864 г. «scale books» (шкалированные книги) англичанина Джорджа Фишера. В 1894 г. американец Дж. М. Райс применял свои таблицы по проверке знаний орфографии для изучения эффективности дидактических приемов. В 1908 г. ученик Торндайка Стоун опубликовал первый тест по арифметике. После того, как Союз американских школьных советов в 1914 г. отстоял

объективные методы оценки, США захлестнул бум тестов по школьной успеваемости. [2, с. 13]

До тех пор, пока отправление ведущих функций зависело не от приобретенной той или иной личностью квалификации, а от признаков, которые характеризуют с момента рождения (принадлежность к тому или иному сословию), а также от религиозных убеждений и покупки должностей, до тех пор не было необходимости в особых диагностических процедурах для определения квалификации. В последние два столетия индивидуальные успехи в учебе становились все более важными для личной карьеры. Инструментарий для определения приобретенной личностью квалификации соответствовал неточным общим требованиям. Оценка стала доминирующим средством определения успеваемости, не смотря на то, что она содержала слишком мало информации для совершенствования учебного процесса. Первые отдельные возражения против принятой оценочной практики были обобщены в докладе Цилинского в 1961 г. На международных конференциях, проходивших в конце 20-х – начале 30-х г.г., был дан толчок критическому анализу существовавшей оценочной практики, который нашел свое отражение в работах английских, американских и немецких ученых (Hartog, Rhodes, Kandel, Bobertag). Лишь с появлением тестов, определяющих уровень физического и умственного развития ребенка, последовало их более активное внедрение. Важную роль в этот период сыграли первое после войны руководство по созданию тестов Бельзера, классический труд Линерта «Построение и анализ тестов». Международная конференция в Берлине 1967 г. дала новый толчок развитию прежде всего неформальных тестов и тестов, ориентированных на определенный критерий. [2, с. 15-16]

На протяжении многих десятилетий тесты школьной успеваемости создавались по правилам классической теории теста. Результаты тестирования сообщались таким образом, чтобы их можно было сравнить с результатами тестирования всей соотносительной группы (чаще всего классов того или иного выпуска). Этот вид тестов чаще всего назывался «ориентированным на норму» или «стандартизованным». Позднее возникла потребность в тестах, результаты которых позволяли бы определить, каких учебных целей достиг тот или иной учащийся. Этот вид тестов получил название «ориентированного на учебные цели или критерии». [2, с. 125]

Кюффнер [3] различает три основных вида тестов школьной успеваемости:

А. Методы тестирования, фиксирующие успеваемость в целом при завершении какого – либо периода обучения (например, для выставления отметки или получения сертификата).

Б. Методы тестирования, фиксирующие успеваемость в более или менее обширных сегментах учебных целей и предлагающие в качестве

результата информацию о том. Достигнуты ли учебные цели в достаточной мере (методы, ориентированные на критерии).

В. Методы тестирования, фиксирующие успеваемость на каком – либо отрезке занятия и информирующие о том, где следует искать сильные и слабые стороны ученика или где могут возникнуть трудности с усвоением материала (диагностические методы).

Тесты школьной успеваемости, ориентированные на критерии, - это тесты, сообщающие индивидуальный результат, соотнося его с заранее установленным критерием. Для большей наглядности обратимся к дефиниции Фрике: «Тест, ориентированный на критерии, является обычным научным методом, с помощью которого определяется возможность и степень достижения конкретной учебной цели. Используемые при этом задания не идентичны учебной цели, а только представляют ее и служат для сравнения фактических индивидуальных способностей учащегося с желаемой степенью способностей. Для проведения этого сравнения необходимы квантификация учебной цели, количественная фиксация успеваемости и измерительная модель для принятия решения о том, достигнута ли учебная цель. Для вычисления критериев качества теста необходим специальный, ориентированный на критерии анализ теста» [4].

Различают следующие фазы создания тестов: а) подготовительное планирование; б) куррикулярный анализ (куррикулум (*curriculum*) – общий план проведения занятия, включающий содержание занятия, его цели, методы и результаты); в) анализ заданий; г) выверка.

На стадии *подготовительного планирования* необходимо проанализировать, кто, что и почему подлежит исследованию. Чем глубже размышления о целях тестирования, тем планомернее и продуманнее могут стать последующие этапы создания теста. На этой стадии обсуждается вопрос о педагогическом значении теста, его дидактической корректности и диагностической функции.

Куррикулярный анализ в создании теста относится, прежде всего, к анализу учебных планов и учебников. Обычно отфильтровывают из учебников важнейшие учебные цели и заносят их в матрицу, в которой на одной оси расположены содержательные структуры, а на другой – уровни требований (или категории: «понимание», «обдумывание», «восстановление в памяти»).

Конструируя задания, необходимо прежде всего взвесить, какой тип заданий лучше всего соответствует цели, содержанию и условиям тестирования. В зависимости от целей исследования уже при конструировании заданий предпринимаются попытки настроиться на требуемый уровень трудности. В соответствии с приобретенным опытом авторы теста создают на 50 или 100%

заданий больше, чем того требует окончательная форма теста, ибо только часть заданий выдерживает проверку.

Эмпирический *анализ заданий* должен предоставить сведения о степени сложности каждого задания, его селективности и – при выборе ответа из нескольких предложенных вариантов – о выверке дистракторов. Эти величины испытывают на себе воздействие выборки, с помощью которой проверяется опытный вариант теста. По этой причине группы учащихся, подобранные для анализа заданий, должны по своему составу быть максимально приближены к тем, на кого ориентирована окончательная форма теста. Кроме того, должны быть выдержаны сроки проведения исследования. [2, с. 136-137]

Степень сложности заданий определяется процентом учащихся, получивших верный результат. Если 80% учеников, обрабатывающих задание, решили его правильно, то в таком случае степень сложности этого задания составляет 80.

Степень сложности задания можно вывести с помощью следующей формулы:

$$P = 100 \frac{\text{количество тестируемых, правильно решивших задание}}{\text{количество тестируемых, обработавших задание}}$$

Степень сложности **P** в числовом выражении тем больше, чем легче задание. *Коэффициент селективности задания* определяет взаимосвязь между решением этого задания и решением всех заданий в опытной форме теста. Эта взаимосвязь выражается с помощью коэффициента корреляции, который теоретически может варьировать между -1 и +1. Если учащиеся, решившие данное задание, выполнили также максимально возможное количество заданий во всем опытном варианте теста и, напротив, учащиеся, не сумевшие решить это задание, не смогли выполнить и остальные, то в таком случае мы получаем коэффициент селективности, который приближается к величине +1.

Если не контролировать число легких и трудных заданий, то возникает так называемая проблема сбалансированного теста по трудности, суть которой в том, что обычно в тесте должно быть больше заданий средней трудности и меньше совсем легких и трудных. В процессе конструирования теста мера трудности регулярно проверяется на случайной выборке испытуемых из того контингента, для которого тест предназначается. В сбалансированном тесте авторы обычно легко добиваются нормальности распределения результатов. Дальнейшее совершенствование теста идет по пути замены тех заданий, ответы на которые нарушают нормальность распределения. [2, с. 127-131]

В ходе анализа качества тестов должны быть установлены валидность и надежность тестов. [5, с. 126-127]

Неформальные тесты, или так называемые тесты, созданные учителями («teacher made tests»), отличаются от формальных тестов тем, что ориентированы на педагогические потребности одной группы учащихся. Неформальные тесты школьной успеваемости являются методом педагогической диагностики; они конструируются преподавателями с целью максимально объективно зафиксировать результаты планируемого ими процесса обучения в своем классе (группе) и использовать их в своей педагогической деятельности. [2, с. 144-145]

К сожалению, не всегда тесты находят применение на этапах текущего и рубежного контролей в вузе. Это объясняется тем, что преподавателей отпугивают затраты, связанные с их созданием, а также недооценка возможностей тестовой формы контроля объективно оценивать знания студентов. При этом надо отметить, что более высокие затраты при конструировании тестов окупаются значительной экономией времени при обработке данных, объективность оценки знаний может быть достигнута при правильном построении тестовых заданий, с учетом всех требований к их созданию. Тесты, построенные на тщательном изучении учебных программ и применении методик измерения, могут заменить многочисленные письменные работы и устные задания. Система банка тестов по определенной дисциплине должна постоянно дополняться новыми заданиями при одновременном повышении качества тестовых заданий.

Концепция учебного тестирования в немецкоязычных странах была представлена прежде всего работами Гутке, который усматривал в ней «попытку диалектического отрицания теста интеллекта, т.е. преодоление основных недостатков традиционных тестов интеллекта при сохранении позитива психометрического начала» [6]. Традиционный тест интеллекта получил определение «статус - теста» («Statustest»), проверяющего точно готовую информацию о прошлой успеваемости. «В отличие от этого метода с помощью концепции учебного тестирования можно было бы стимулировать учебную активность благодаря включению по возможности наиболее стандартизированной педагогизированной фазы, чтобы получить дополнительную и/или специальную информацию об индивидуальных предпосылках учебного процесса, формах обучения, процессах решения заданий, а также успеваемости» [7]. Учебное тестирование со своими основными компонентами – предтест, стандартизированная учебная фаза, пост – тест – было опробовано в различных вариантах, в которых варьировались продолжительность учебной фазы и ее оформление, например, программированное или нет. Более высокую прогностическую валидность в плане диагностики интеллектуального развития учебные тесты пока еще не доказали. С точки зрения индивидуальной диагностики трудностей обучения они, очевидно, обладают большими возможностями. Гроффманн приходит к следующему выводу: «Учебные тесты не внесли революционных изменений в диагностику интеллекта. Возможно, эти методы имеют некоторые специфические преимущества для определенных областей

диагностики, которые, однако, должны быть еще убедительно доказаны. Несмотря на ранее упомянутые нерешенные методические проблемы диагностики с помощью учебных тестов, было бы преждевременным пренебрегать проверкой учебных тестов только потому, что они еще не дали ожидаемых результатов» [8].

Таким образом, основными этапами разработки тестовых заданий являются:

1. Выбор содержания заданий;
2. Выбор формы заданий;
3. Создание инструкции и описания теста;
4. Апробация теста (сбор эмпирических данных);
5. Обработка эмпирических данных;
6. Интерпретация результатов обработки;
7. Экспертиза качества теста.

Классическая теория тестов опирается на теорию корреляции, главными параметрами которой являются надежность и валидность.

Если раньше данные тестирования подвергались статистической обработке, то в настоящее время конечный результат достигается с помощью математического структурного анализа. Теория математического структурного анализа основывается на том, что человеческие знания имеют некоторую латентную переменную. Идея этой переменной (параметра) и легла в основу тестирования. Основное предположение математического структурного анализа состоит в том, что между результатами теста и измеряемыми параметрами есть функциональная зависимость, позволяющая с помощью анализа и оценки этих параметров создавать представление о знании человека. [5, с. 130-131]

Результативность в процессе обучения во многом зависит от тщательно отработанной методики контроля знаний. Проверка и оценка знаний, умений и навыков обучающихся – важное звено учебного процесса. Необходимость контроля объясняется потребностью в получении информации об эффективности функционирования системы обучения.

Тестовый контроль – это оперативная проверка качества усвоения знаний, немедленное исправление ошибок и восполнение пробелов. Тестовый контроль помогает педагогу оперативно проверить уровень формирования представлений и понятий обучающихся, определить их продвижение в обучении. Использование тестов для проверки знаний обучающихся повышает их объективность, позволяет определить уровень самостоятельной работы.

В структуре проверки знаний тесты могут быть предварительными, текущими, тематическими, обобщающими, итоговыми.

Литература:

1. Педагогика и психология высшей школы. М., 1998, с. 126-127, 130-131.
2. Ингенкамп К. Педагогическая диагностика. – М., «Педагогика», 1991, с. 144-145, с. 13, 15 - 16, 124-125, 127-131, 136-137.
3. Küffner H. Fehlerorientierte Tests: Konzept und Bewährungskontrolle. Weinheim: Beltz, 1980.
4. Fricke R., Lühmann R. Kriteriumsorientierte Tests – Theorie und Praxis. Psychologische Arbeiten aus dem FB Erziehungswissenschaften 1 der Universität Hannover. Arbeit Nr. 1. 1982. Juni.
5. Аванесов В.С. Вопросы объективизации оценки результатов обучения. – Отдел научной информации, М., 1976. с. 51-53.
6. Guthke J. Zur Diagnostik der intellektuellen Lernfähigkeit. Stuttgart: Klett, 1977, с. 9.
7. Kornmann A. Möglichkeiten von Lerntests für Diagnose und Optimierung von Lernprozessen // In: Ingenkamp K., Horn R., Jäger R.S. Tests und Trends 1982. Weinheim: Beltz, 1982, с. 99.
8. Groffmann K.-J. Die Entwicklung der Intelligenzmessung // In: Groffmann K.-J., Michel L. Intelligenz- und Leistungsdiagnostik. Enzyklopädie der Psychologie B. II, 2. Göttingen: Hogrefe, 1983, с. 70.