

МОЕ ПОНИМАНИЕ СИСТЕМЫ ОЦЕНИВАНИЯ В ПРОЦЕССЕ ПРОВЕДЕНИЯ ТЕСТОВ

Александр Васильевич ГАЛУЩАК, преподаватель географии

I дидактическая степень. Гимназия с. Каларашовка, Окница

Изучив вопрос, я пришел к мнению, что применение тестов в зарубежных школах имеет давнюю историю. Признанный авторитет в области педагогического тестирования Э. Торндайк (1874—1949) выделяет три этапа внедрения тестирования в практику американской школы [4]:

1. Период поисков (1900—1915). На этом этапе происходило осознание и первоначальное внедрение тестов памяти, внимания, восприятия и других, предложенных французским психологом А. Бинэ.

2. Последующие 15 лет — годы «шума» в развитии школьного тестирования, приведшие к окончательному осмыслению его роли и места, возможностей и ограничений. Ч. Спирмен предложил общие основы использования корреляционного анализа для стандартизации тестов [4].

3. С 1931 г. начинается современный этап развития школьного тестирования. Поиски специалистов направляются на повышение объективности тестов, создание непрерывной (сквозной) системы школьной тестовой диагностики, подчиненной единой идее и общим принципам, созданию новых, более совершенных средств предъявления и обработки тестов, накопления и использования диагностической информации.

В школах развитых стран внедрение и совершенствование тестов шло быстрыми темпами. Широкое распространение получили диагностические тесты успеваемости, использующие форму альтернативного выбора правильного ответа из нескольких правдоподобных, написания очень краткого ответа (заполнения пропусков), дописывания букв, цифр, слов, частей формул и т. п. С помощью этих несложных заданий удается накапливать значительный статистический материал, подвергать его математической обработке, получать объективные выводы в пределах тех задач, которые предъявляются к тестовой проверке. Тесты печатаются в виде сборников, прилагаются к учебникам, распространяются на цифровых носителях, кроме того, тесты — это один из самых дешевых форм оценивания и контроля уровня знаний учащихся и системы образования.

С моей точки зрения, прежде чем определиться с темой и типом теста необходимо:

1. Определить какова будет функция оценивания теста (глобальная, учебное заведение, дидактические кадры, ученики). Через функцию оценивания мы выясняем уровень образования, систему подготовки учебного заведения и дидактических кадров, уровень навыков, умений и знаний учащихся, видим как меняется менталитет нации, её место в цивилизации, можем воздействовать на их изменение. Благодаря этому, и исходя из уровня проверки тестирования, мы можем менять цель оценивания. Поэтому можно констатировать, что оценивание – это процесс, имеющий функции, задачи, уровни, внутренние и внешние операции.

2. Определить уровень тестирования (элементарный, средний, углубленный, профессиональный).

3. Определить время оценивания.

4. Целевую аудиторию, т.е. кто будет оцениваться.

5. Какой тип оценивания и какие формы оценивания будут применяться (как будут оцениваться).

6. Определить задачи (что и почему я хочу оценить).

7. Для кого производится оценивание (уровень).

8. Зачем производится оценивание (цели).

9. Определить систему и критерии оценивания (составить борем, определить итем, выработать матрицу спецификаций).

10. Провести анализ проведенного теста для целевой аудитории (за что получили те, или иные баллы).

Во-вторых, необходимо учитывать, что если в основу классификации тестов положить различные аспекты (компоненты) развития и формирования человеческих качеств, то она будет выглядеть следующим образом:

1. Тесты общих умственных способностей, умственного развития.

2. Тесты специальных способностей в различных областях деятельности.

3. Тесты обученности, успеваемости, академических достижений.

4. Тесты для определения отдельных качеств (черт) личности (памяти, мышления, характера и др.).

5. Тесты для определения уровня воспитанности (сформированности общечеловеческих, нравственных, социальных и других качеств).

При разработке тестов важно, насколько они соответствуют запроектованным целям обучения, образования, развития обучаемых. Важнейшими критериями диагностических тестов обученности являются действительность (валидность, показательность), надежность (вероятность, правильность), дифференцированность (различимость).

В-третьих, тесты обученности применяются на всех этапах дидактического процесса. С их помощью обеспечивается предварительный, текущий, тематический и итоговый контроль знаний, умений, учет успеваемости, академических достижений.

Предварительный контроль. Успех изучения любой темы (раздела или курса) зависит от степени усвоения тех понятий, терминов, положений и т.д., которые изучались на предшествующих этапах обучения. Если информации об этом у преподавателя нет, то он лишен возможности проектирования и управления в учебном процессе, выбора оптимального его варианта. Необходимую информацию преподаватель получает, применяя предварительный контроль (учет) знаний. Последний необходим еще и для того, чтобы зафиксировать (сделать срез) исходный уровень обученности. Сравнение исходного начального уровня обученности с конечным (достигнутым) позволяет измерять «прирост» знаний, степень сформированности умений и навыков. Если известны входные и выходные характеристики системы, проблемы ее оптимизации считаются во многом решенными.

Собрать максимальный объем информации о входных характеристиках обученности и оценить их в количественных показателях удастся с помощью тестирования, которое осуществляется с помощью специально разработанных для этой цели заданий.

Текущий контроль. Он необходим для диагностирования хода дидактического процесса, выявления динамики последнего, сопоставления реально достигнутых на отдельных этапах результатов с запроектированными. Кроме собственно прогностической функции текущий контроль и учет знаний, умений стимулирует учебный труд учащихся, способствует своевременному определению пробелов в усвоении материала, повышению общей продуктивности учебного труда.

Тематический контроль. Составление тематического тестового задания требует кропотливого и тщательного труда. Ведь речь идет не просто о проверке усвоения отдельных элементов, а о понимании системы, объединяющей эти элементы. Значительную роль при этом играют синтетические, комплексные задания, объединяющие вопросы об отдельных понятиях темы, направленные на выявление информационных связей между ними.

Итоговый контроль. Осуществляется во время заключительного повторения, а также в процессе экзаменов (зачетов). Именно на этом этапе дидактического процесса систематизируется и обобщается учебный материал. С высокой успешностью могут быть применены соответствующим образом составленные тесты обученности.

Естественно, не все необходимые характеристики усвоения можно получить средствами тестирования. Такие, например, показатели, как умение конкретизировать свой ответ примерами, знание фактов, умение связно, логически и доказательно выражать свои мысли, некоторые другие характеристики знаний, умений, навыков диагностировать тестированием

невозможно. Это значит, что тестирование должно обязательно сочетаться с другими традиционными формами и методами проверки.

Все выше сказанное, подтверждает точку зрения, высказанную президентом Всемирного Банка (2010): «Оценивание, имеет первостепенное значение в образовании: оно влияет не только на показатели количества учащихся, охваченных образованием, и показатели тех, кто окончил школу, но и на конечную цель обучения учащихся».

Теперь в рамках моего исследования становится понятно, что мы должны контролировать – оценивать овладение учеником новыми возможностями по каждой из пяти линий развития, фиксировать продвижение ученика от уровня «мини» к уровню «макси». Следовательно, к каждому учебному параграфу и в итоговых проверочных работах в виде тестов должны быть представлены соответствующие разноуровневые задания, позволяющие оценить уровень обучения и обученности.

И здесь применяются различные формы и виды педагогических тестов. Надежность - это характеристика теста, отражающая точность педагогического измерения, а также устойчивость результатов тестирования к воздействию посторонних (случайных) факторов.

Для оценки надежности нормативно-ориентированного теста используются две группы методов:

1. Двукратное тестирование:

- ретестовый метод;
- метод параллельных форм.

2. Однократное тестирование (метод расщепления теста).

Оценка надежности чаще всего строится на подсчете корреляции двух наборов результатов выполнения одного и того же теста или двух его параллельных форм. Чем выше корреляция, тем надежнее тест.

Ретестовый метод.

Данный метод оценки надежности предполагает двукратное проведение одного и того же теста в одной группе испытуемых. Повторное тестирование проводится примерно через 2-3 недели, чтобы учащиеся не слишком сильно забыли проверяемый материал и недалеко продвинулись в изучении нового, то есть уровень их учебных достижений остался примерно прежним. Ретестовый метод основан на подсчете корреляции индивидуальных баллов испытуемых по результатам выполнения первого и второго тестирования.

Метод параллельных форм.

Метод параллельных форм предполагает двукратное тестирование одной и той же группы испытуемых тестами, которые идентичны по содержанию, структуре, включают задания тождественные по трудности, дифференцирующей способности и др., то есть

параллельными формами теста. Главная трудность применения этого метода в том, что прежде чем использовать параллельную форму теста, разработчику надо создать новый тест, тождественный первому, и доказать идентичность, а это очень трудоемкая процедура. Поэтому этот метод, с практической точки зрения, малоэффективен.

Классификация педагогических тестов

По целям использования выделяют (классификация предложена Гронлундом, ее же придерживаются М.Б. Чельшкова, А.Н. Майоров):

- 1) *входное тестирование* (обеспечивает проверку знаний и умений в начале обучения);
- 2) *формирующее и диагностическое тестирование* (предполагает контроль формирования новых знаний и умений в процессе обучения, обеспечивает текущий контроль);
- 3) тематическое, итоговое, рубежное тестирование обеспечивает определение итоговых достижений [7].

Два подхода к разработке тестов

- 1) нормативно-ориентированный,
- 2) критериально-ориентированный.

Соответственно различают два вида тестов:

- *нормативно-ориентированный* (позволяет сравнивать уровень учебных достижений испытуемых друг с другом);
- *критериально-ориентированный* (служит для определения уровня освоения учащимся содержания какого-либо предмета, дисциплины, раздела, темы) [7].

Классификация видов тестов

По средствам предъявления педагогические тесты делятся на:

- 1) *бланочные* (испытуемые отмечают или вписывают правильные ответы на бланке);
- 2) *компьютерные* (задания высвечиваются на мониторе компьютера).

По степени однородности задач:

- 1) *гомогенные* (контроль знаний и умений по одному предмету, дисциплине);
- 2) *гетерогенные* (измерение уровня подготовленности по нескольким учебным предметам).
 - а) *полидисциплинарный* (проверка знаний по нескольким дисциплинам)
 - б) *междисциплинарный* (проверка знаний и умений в смежных дисциплинах (математика-физика, химия-биология, история-культурология)).

Типы тестов [7]:

I. *Открытые* (испытуемый сам записывает правильный ответ)

Формы тестовых заданий открытого типа

- а) задания на дополнение (краткий ответ: формула, числовое выражение, слово и др.).

б) задания со свободно конструируемыми ответами (ответ в свободной, развернутой форме: решение задачи с пояснениями, небольшое сочинение. Должна оговариваться полнота ответа для получения максимального балла).

II. *Закрытые* (предполагают выбор правильного ответа из предложенных вариантов. Оптимальное количество альтернативных ответов - 4-5. Неправильные, но правдоподобные ответы, то есть похожие на правильные, называются дистракторами (от англ. distract - отвлекать).

Формы тестовых заданий закрытого типа

а) задания альтернативных ответов (задания с двумя ответами: да - нет, правильно - неправильно). Имеют очень высокий процент угадывания (50%), их одиночное использование в тесте считается неэффективным;

б) задания с выбором нескольких правильных ответов;

в) задания на установление соответствия;

г) задания на установление правильной последовательности.

Функции компьютерного теста

– обучение (в качестве дидактического обучающего средства);

– контроль ЗУНов учащихся;

– самообучение (тренажёр) и самоконтроль;

– дистанционное обучение;

– корректировка учебного процесса.

Инструментальные тестовые оболочки

Для создания тестов по предметной области разрабатываются специальные инструментальные *программы-оболочки*, позволяющие создавать компьютерные тесты путем формирования базы данных из набора тестовых заданий.

Сбор и статистическая обработка результатов тестирования

Статистическая обработка полученных данных тестирования (результатов испытуемых) позволяет определить количественные характеристики тестовых заданий и всего теста, что даёт возможность использовать тест как научно обоснованный, объективный инструмент педагогических измерений. Статистическая обработка результатов тестирования проводится с помощью программного обеспечения.

1. Матрица тестовых результатов

Для удобства обработки тестовых данных результаты заносятся в матрицу тестовых результатов. Ниже приведён пример такой матрицы (один балл за правильный ответ, ноль баллов - за неправильный ответ или пропуск задания) [7].

Таблица 1 Матрица тестовых результатов

№ испытуемого i	№ задания j					Индивидуальный балл X_i
	1	2	3	4	5	
1	0	0	0	0	0	0
2	1	0	1	0	1	3
3	1	1	1	0	0	3
4	1	1	0	0	0	2
5	1	1	1	1	0	4
Число правильных ответов Y_j	4	3	3	1	1	12

Для наглядности и удобства дальнейшего использования необходимо упорядочить матрицу тестовых результатов. Для этого строки таблицы располагают таким образом, чтобы индивидуальные баллы располагались в порядке возрастания. Затем столбцы матрицы располагают в порядке убывания.

Таблица 2 Упорядоченная матрица тестовых результатов

№ испытуемого i	№ задания j					Индивидуальный балл X_i
	1	2	3	4	5	
1	0	0	0	0	0	0
4	1	1	0	0	0	2
2	1	0	1	0	1	3
3	1	1	1	0	0	3
5	1	1	1	1	0	4
Число правильных ответов Y_j	4	3	3	1	1	12

2. Трудность тестовых заданий

По результатам апробационного тестирования определяются характеристики тестовых заданий - трудность и дискриминативность.

Трудность задания – соотношение количества испытуемых, справившихся с данным заданием, и общего количества испытуемых. Трудность задания вычисляется по формуле:

$(r_{disc})_j = (p1)_j - (p0)_j$, где p_j - доля правильных ответов на j -ое задание; Y_j - количество испытуемых, выполнивших j - ое задание верно, N - число испытуемых в группе, j - номер задания.

Чем выше показатель трудности, тем задание легче, чем меньше показатель трудности задания, тем задание сложнее.

Показатель трудности важен для определения характеристики тестового задания и помогает проранжировать задания, входящие в тест по степени сложности. Благодаря этому можно определить место задания в тесте.

Исходя из вышеизложенного, могу сделать вывод, что грамотно составленные и апробированные тесты позволяют достичь коэффициента надежности 0,9. Установлено, что надежность теста повышается при увеличении количества тестовых заданий. Мною замечено, что тест, нацеленный на проверку усвоения конкретной темы, всегда будет более надежным, чем тест, направленный на проверку всего раздела (курса), охватывающий значительное количество материала — закономерностей, концепций, фактов. Происходит это именно

потому, что содержательное разнообразие последнего выше. Наибольшую практическую ценность имеют задания, на которые правильно отвечают 45—80% обучаемых.

Хотелось бы порекомендовать, при подготовке материалов для тестового контроля необходимо придерживаться таких основных правил:

1. Нельзя включать ответы, неправильность которых на момент тестирования не может быть обоснована учащимися.
2. Неправильные ответы должны конструироваться на основе типичных ошибок и должны быть правдоподобными.
3. Правильные ответы среди всех предлагаемых ответов должны размещаться в случайном порядке.
4. Вопросы не должны повторять формулировок учебника.
5. Ответы на одни вопросы не должны быть подсказками для ответов на другие.
6. Вопросы не должны содержать «ловушек».

Тесты обученности, например, мною применяются на всех этапах дидактического процесса. С их помощью обеспечивается предварительный, текущий, тематический и итоговый контроль знаний, умений, учет успеваемости, академических достижений.

Библиографический список

1. Национальный образовательный куррикулум. 2013.
2. Аванесов, В.С. Композиция тестовых заданий. - М., 2002.
3. Калужская, М. В. Рейтинговая система оценивания. Как? Зачем? Почему? –М.: Чистые пруды, 2006.
4. Подласый, И. П. Тестирование в учебном процессе: его история и возможности. Челябинск, 2006.
5. Шамова, Т.И. Современные средства оценивания результатов обучения в школе. М: Педагогическое общество России, 2007.
6. Формирующее оценивание – обучение для будущего: типы оценивания. <http://educate.intel.com/ru/AssessingProjects/OverviewAndBenefits/FormatieAssessment/>.
7. Чельшкова, М. Б. Разработка педагогических тестов на основе современных математических моделей. М., 2009.
8. Чернявская, А.П., Гречин, Б.С. Современные средства оценивания результатов обучения. Ярославль, 2008.