

О. А. Максимова, М. А. Червонный

ОСОБЕННОСТИ РАЗРАБОТКИ И КОМПЛЕКСНОЙ ЭКСПЕРТИЗЫ КАЧЕСТВА ПЕДАГОГИЧЕСКИХ ТЕСТОВ, ИСПОЛЪЗУЕМЫХ ДЛЯ КОНТРОЛЯ И ОЦЕНКИ КАЧЕСТВА ОБУЧЕНИЯ СЕЛЬСКИХ ШКОЛЬНИКОВ

Рассматриваются научно-методические основы разработки и технологии комплексной экспертизы педагогических тестов для контроля и оценки качества обучения сельских школьников, особенности применения тестовых технологий в учебном процессе. Предлагаются подходы к повышению качества педагогических тестов.

Ключевые слова: педагогический тест, тестовые технологии, критерии качества педагогического теста, комплексная экспертиза качества теста, таксономия учебных целей.

Обучение в сельской малокомплектной школе характеризуется рядом специфических условий [1–3]. В этих условиях контроль является не только средством обратной связи качества усвоения учебного материала, но и зачастую практически единственным средством контроля, позволяющим эффективно организовать диагностическую функцию при разновозрастном обучении. В последние десятилетия в отечественной педагогике вопрос об эффективности и целесообразности применения тестовых технологий считается практически решенным как педагогами-исследователями, так и педагогами-практиками ввиду того, что контроль с использованием тестовых технологий в полной мере соответствует принципам оперативности, объективности, систематичности, комплексности, полифункциональности, универсальности и др. [4].

В сельской малокомплектной школе самостоятельная работа обучающихся занимает более половины учебного времени, и вопрос об эффективности контроля качества учебных достижений обучающихся является одним из самых актуальных. Как организовать систему контроля за деятельностью обучающихся, какие при этом технологии использовать, какие задания дают наиболее объективную оценку уровня учебных достижений обучающихся?

Для оценки уровня учебных достижений обучающихся в образовательной практике используют тесты учебных достижений. Под тестами учебных достижений понимается система заданий специфической формы, определенного содержания возрастающей трудности, позволяющая качественно оценить структуру, измерить уровень ЗУН и представлений по изучаемому предмету. Педагогический тест – это многокомпонентная структура, состоящая из следующих элементов:

- спецификации;
- инструкции по применению;
- инструкции по выполнению для испытуемых;
- инструкции по обработке результатов тестирования;

- системы тестовых заданий, расположенных определенным образом в общей структуре теста;
- эталонов правильных ответов.

С позиций системного подхода тест рассматривается как система, состоящая из подсистем – тестовых заданий, обладающих рядом свойств, характерных для системы в целом: целостность, неаддитивность, эмерджентность, синергизм, обособленность, совместимость и адаптивность.

Целостность теста характеризуется рядом свойств и особенностей. Многогранность целостности отражается наличием у теста общей цели – гарантии оценки качества подготовки обучающихся и характеризуется дифференциацией, интеграцией и асимметрией тестовых заданий, решающих различные задачи для достижения заданной цели.

Неаддитивность означает появление нового качества теста как системы, возникающего в результате интеграции тестовых заданий в единое целое. При этом интеграция тестовых заданий во многом определяется их совместимостью – взаимосвязанностью внутри теста в логике совместимости самого теста с элементами других систем, входящих в систему обучения:

- требованиями образовательного стандарта и образовательной программы к содержанию и результатам подготовки обучающихся;
- программой и кодификатором учебной дисциплины;
- другими контролирующими (оценочными) материалами (контрольными заданиями, опросниками, заданиями на выполнение лабораторных работ и пр.).

При успешной интеграции тестовых заданий тест проявляет свойство синергизма, что означает однонаправленность действий тестовых заданий и теста в целом.

Эмерджентность означает появление у теста эмерджентных свойств, которые не присущи составляющим ее элементам – тестовым заданиям.

Обособленность теста как системы заключается в некоторой его изолированности системы от взаи-

модействия с другими контролирующими (оценочными) материалами в общей иерархии построения системы обучения, контроля и оценки образовательных достижений обучающихся.

Адаптивность – свойство теста как системы, характеризуется возможностью применения теста в изменяющихся внешних и внутренних условиях вследствие коррекций в целях обучения и содержания учебных дисциплин.

В учебном процессе педагогические тесты учебных достижений применяются для следующих видов контроля:

1. Предварительный (входной) контроль служит для получения сведений об исходном (стартовом) уровне учебных достижений. Применяется для оценивания базисных знаний и умений, определяет наличие пробелов и параллельно осуществляет подготовку к изучению учебного предмета.

2. Текущий контроль выполняет функцию обратной связи. Результаты контроля используются для корректировки учебного процесса. Регулярное проведение текущего контроля оказывает мотивационный эффект.

3. Тематический контроль проводится для оценки результатов усвоения определенной темы, служит для повторения пройденной темы и подготовки к изучению следующей темы.

4. Рубежный контроль (промежуточный) позволяет осуществить проверку воспроизведения и понимания материала, выяснить объем накопленных знаний перед началом изучения новой темы.

5. Итоговый контроль – это вид контроля, который призван дать представление о достигнутых результатах. Рубежный контроль рассматривается

как частный случай итогового контроля.

Проведение всех видов педагогического контроля с применением тестовых технологий позволяет проводить сравнительный анализ результатов, отслеживать индивидуальную динамику уровня учебных достижений обучающихся и выражать ее в количественных единицах.

Особенность проведения контроля с использованием тестовых технологий заключается в прямой зависимости надежности и объективности результатов тестирования от качества педагогического теста. Результаты тестирования будут объективными и достоверными только в том случае, если для контроля используются валидные и надежные педагогические тесты [5]. Качество педагогического теста определяется целым рядом критериев и показателей, среди которых содержательная валидность, критериальная валидность, параллельность вариантов тестирования, согласованность заданий в тесте и др. В научной литературе по дидактической тестологии нет единого мнения о том, каким количеством критериев должно оцениваться качество педагогических тестов. Так, ряд авторов выделяют два критерия – надежность и валидность [6]; другие рассматривают в качестве критериев качества теста: надежность, валидность и объективность [7, 8]; третьи – надежность, валидность, объективность, экономичность и полезность. Основной перечень критериев и показателей качества педагогических тестов представлен в табл. 1.

Не менее важным вопросом при разработке педагогических тестов учебных достижений является этап отбора содержания для проведения контроля. Методологическую основу контроля и оценки каче-

Таблица 1

Критерии качества педагогических тестов и инструментарий их оценивания

№	Критерий оценки качества педагогических тестов	Определение	Методы оценивания критериев качества педагогических тестов
1	Содержательная валидность	Степень репрезентативности содержания ТЗ по отношению к измеряемой характеристике	Экспертное оценивание
2	Очевидная валидность	Внешнее представление о тесте с точки зрения испытуемого	Анкетный опрос
3	Функциональная валидность	Степень соответствия содержания ТЗ по отношению к выявляемому уровню усвоения опыта деятельности	Экспертное оценивание
4	Прогностическая валидность	Степень точности и обоснованности суждения о диагностируемой переменной	Методы математической статистики
5	Критериальная валидность	Степень соответствия между результатами тестирования и внешним, не относящимся к тесту критерием	Методы математической статистики
6	Конструктивная валидность	Отражение в тесте теоретической модели курса, его структурных пропорций и основных компонентов на всех уровнях	Экспертное оценивание
7	Надежность	Отражает точность измерения и устойчивость результатов к воздействию посторонних случайных факторов	Методы математической статистики
8	Сбалансированность	Отражает сбалансированность ТЗ по сложности и структуре	Методы математической статистики
9	Оптимальность	Отражает корректный подбор длины теста и оптимального времени его предъявления и времени его выполнения для группы тестируемых	Экспертное оценивание

ства образования с использованием тестовых технологий составляет теория педагогических измерений, согласно которой промежуточные и конечные цели обучения описываются в соответствии с таксономией учебных целей [5]. Под таксономией педагогических целей понимается система, содержащая упорядоченную классификацию целей, используемая при планировании обучения и оценки его результатов, являющаяся надежным инструментом при проверке новых курсов [9]. В отличие от традиционной классификации таксономия представляет собой иерархическое описание целей, согласно которому более простые, частные цели кумулируются (лат. *simulatio* «накопление, увеличение») в более сложные, общие. Таксономия позволяет описать структуру целей, степень их сложности и взаимные отношения. Основу таксономии составляет операционализация учебных целей.

Операционные цели содержат описание конечного поведения обучающегося (действие и содержание описания), условия его проявления и стандарты (критерии) его достижения. Таким образом, таксономия учебных целей является инструментом, позволяющим оценить результаты учебного процесса с квалитетических позиций, т. е. выходить на качественный и количественный анализ результатов обучения.

Конкретная деятельность по каждому предмету и каждому его разделу определяется планируемыми целями обучения. Цели обучения содержат систему задач, которую должны научиться решать обучающиеся при изучении данного предмета. За-

дачи, составляющие цели обучения, определяют не только содержание контроля, но и являются критериями оценки уровня учебных достижений обучающихся. Умение использовать содержание изученного предмета (раздела) при решении этих задач и означает, что цели обучения по данному предмету достигнуты. Соотнесенность содержания контрольных заданий с целями обучения – главное требование к организации контроля с использованием тестовых технологий. Особенность применения педагогических тестов для контроля и оценки качества обучения заключается в том, что каждое отдельное тестовое задание является эмпирическим признаком успешности усвоения учебного материала.

Например, в таксономии Блума в состав требований к планируемым результатам обучения включают систему объектов изучения, описание видов учебной деятельности и качество усвоения учебного материала. Все это входит в систему научных знаний.

Первым компонентом требований системы научных знаний является характеристика объектов изучения с учетом глубины их освещения учителем и планируемого уровня усвоения учениками.

К элементам системы научных знаний исследователи отнесли понятия и факты, законы, теории, идеи, знания о способах деятельности, методологические и оценочные знания.

Решение проблемы обеспечения контроля надежным и валидным инструментарием видится в проведении комплексной экспертизы качества тестовых материалов на этапе их разработки.

Таблица 2

Таксономия Блума когнитивных целей обучения по М. В. Кларину

Категория	Конкретизированные цели	Альтернативные глаголы действия
Знание (простое)	Рассматривается как узнавание, запоминание, воспроизведение информации. Запоминание осуществляется на основе взаимосвязи факторов, терминов, понятий, суждений, операций	Определить, описать, выявить, обозначить, перечислить, подобрать (пару), назвать, кратко обрисовать, отобразить, утвердить
Понимание	Обнаруживается в понимании смысла изученного материала, в умении увидеть главное, определить причину и следствие явления и т. п.	Превратить, защитить, различить, оценить, объяснить, расширить, обобщить, дать примеры, заключить, перефразировать, предсказать, переписать, подытожить
Применение информации	Применение (фактов, правил, теорий, методов) в конкретных и новых ситуациях, применение правил, методов, принципов, законов	Изменить, вычислить, продемонстрировать, узнать, манипулировать, видоизменить, действовать, подготовить, получить (какой-то результат), связать, показать, решить, использовать
Анализ	Качество знания, позволяющее разделять информацию на взаимозависимые части, выявление взаимосвязей между ними, осознание принципов организации целого	Анализировать, изобразить, дифференцировать, распознавать, разъединять, выявлять, иллюстрировать, вывести, наметить, указывать, устанавливать (связь), отобразить, отделить, подразделять
Синтез	Творческое преобразование информации из разных источников и создание новой, объясняющей явление или событие, предсказывающей что-либо и т. п.	Категоризировать, соединять, составлять, собирать, создавать, разрабатывать, изобретать, переписывать, подытоживать, рассказывать, писать
Оценка	Суждение о целостности идеи, метода, теории на основе проникновения в суть явлений и их сравнения, умение оценивать значение объекта для конкретной цели	Оценить, сравнить, сделать вывод, противопоставить, критиковать, описывать, проводить, различать, объяснять, обосновывать, истолковывать, устанавливать связь, подытоживать, поддерживать

Основу методологии экспертизы качества педагогических тестов составляют методы:

- теории измерения изучаемых объектов, заимствованные у психометрики;
- теории квалиметрии и метрологии;
- теории экспертных оценок;
- теории эффективности применяемых процедур оценивания и критериев оценки качества объектов;
- теории информативности и полезности применяемых процедур.

Под комплексной экспертизой понимается процесс системного исследования совокупности внутренних и внешних свойств тестовых материалов методами и приемами комплексного оценивания с целью получения интегрального показателя качества [10]. Термин «экспертиза» происходит от латинского *expertus*, что означает «опытный». Эффективность применения экспертизы во многом зависит от принципов ее организации и проведения, к таковым относятся: принцип системности экспертизы, комплексности экспертизы, объективности, независимости мнений экспертов, профессиональной компетентности экспертов, преемственности этапов в проведении экспертизы, необходимости контроля за соблюдением процедур экспертизы и др.

Экспертиза качества является одним из главных этапов разработки педагогических тестов, которая проводится для оценки:

- оценки качества спецификации тестового материала;
- степени соответствия тестового материала целям обучения и содержанию учебного предмета, требованиям образовательной программы и образовательного стандарта;
- соответствия формы, содержания (значения, смысла) морфологических (слово, символ, знак) и синтаксических (выражения, обороты) единиц тестового задания форме и содержанию аналогичных единиц, использованных (используемых) при изложении данного предмета.

Технологические этапы экспертизы включают в себя следующие процедуры:

1. Формирование цели экспертизы качества тестов.
2. Классификация свойств тестов, определяющих критериальную базу оценивания качества.
3. Декомпозиция комплексных критериев качества тестов.
4. Формирование процедур и алгоритмов оценивания, выбор методов экспертного оценивания, апробационного тестирования и шкалирования с целью получения итоговых комплексных показателей качества тестов.
5. Формирование процедур и алгоритмов оценивания, выбор методов экспертного оценивания,

апробационного тестирования и шкалирования с целью получения итоговых комплексных показателей качества целых сформированных тестов.

6. Формирование базовых (нормативных) значений показателей качества тестов.
7. Нормирование показателей качества тестов, приведение их к стандартному виду.
8. Перевод количественных показателей качества тестов в их качественные аналоги.

Особую значимость в процедуре формирования логики экспертизы качества тестовых материалов приобретает проблема определения нормативных значений показателей качества. Нормативный показатель качества теста – это значение показателя качества, принятое за основу при оценивании качества объекта экспертизы. Базовые показатели качества теста могут быть установлены несколькими способами:

- в результате заимствования из зарубежных и отечественных источников;
- в результате проведения апробационного тестирования на выборочной совокупности тестируемых, соответствующих свойствам и характеристикам репрезентативной выборки;
- в результате формирования «эталонного теста» из совокупности использованных тестов при проведении тестирования.

Особенность, а вместе с ней и методическая сложность экспертизы заключаются в том, что показатели качества тестовых материалов обладают разнородной выраженностью, т. е. часть показателей может быть выражена числом (количественно), а часть – лингвистическими переменными. Поэтому важными процедурами при проведении экспертизы качества являются формализация, нормирование и интеграция показателей качества. Данная проблема может быть решена только в рамках комплексного оценивания качества тестовых материалов.

Комплексная экспертиза качества педагогических тестов включает в себя следующие этапы:

- 1) предварительная экспертиза;
- 2) внутренняя экспертиза;
- 3) тестологическая экспертиза;
- 4) комплексная экспертиза [11].

Главная задача этапа предварительной экспертизы – отбор тестовых материалов, отвечающих формальным требованиям теста. Основными критериями оценивания на данном этапе являются: наличие спецификации, длина и формулировка тестового суждения, наличие инструкции по выполнению задания, правильного ответа и вариантов ответов (дистракторов).

На этапе внутренней экспертизы качества тестовых материалов осуществляется экспертное оценивание содержательной валидности. Учитывая тот

факт, что содержание тестового материала – слабо формализуемая информация, для оценивания качества необходимо использовать методы экспертного оценивания. На этом этапе для оценивания содержательной валидности привлекаются специально подготовленные специалисты в предметной области – эксперты. Существенным недостатком работы экспертов является субъективность их оценок. Это обусловлено тем, что у каждого специалиста свои представления о «правильности» контроля знаний. Поэтому с целью снижения доли субъективизма в оценивании качества тестовых материалов рекомендуется формирование групп экспертов количеством не менее 5 и использование современных методов поддержки принятия решений на основе экспертного оценивания.

Проведение тестологической экспертизы качества тестовых материалов направлено на оценивание функциональной валидности, определение статистических норм показателей качества тестового

материала, выявление необходимого времени и правил участия (пользование справочными материалами и пр.) обучающихся в процедуре тестирования.

На заключительном этапе проводится комплексное оценивание качества тестовых материалов. Данный этап оказывается наиболее сложным в методическом плане: необходимость соотношения атрибутивных показателей – результатов экспертного оценивания с количественными показателями, полученными на этапе тестологической экспертизы с целью определения характеристики интенсивности проявленного признака качества.

Проведение комплексной экспертизы и соблюдение требований к процедурам оценки качества педагогических тестов являются гарантией разработки валидных и надежных тестовых измерителей, что в свою очередь повысит эффективность контрольно-оценочной деятельности образовательного процесса.

Список литературы

1. Червонный М. А. Проблемы малокомплектной школы в современной образовательной ситуации (примеры разных стран) // Вестн. Томского гос. пед. ун-та (Tomsk State Pedagogical University Bulletin). 2009. Вып. 5 (83). С. 95–98.
2. Войтеховская М. П., Обухов В. В., Червонный М. А. Задачи педагогических университетов по решению задач МКШ // Там же. Вып. 12 (90). С. 7–10.
3. Обоянцева О. В., Червонный М. А. Особенности функционирования и развития сельских муниципальных образовательных систем // Там же. 2010. Вып. 11 (101). С. 57–60.
4. Червонный М. А. Проектирование средств и систем оценивания // Педагогическая диагностика. 2005. № 4. С. 80–91.
5. Беспалько В. П. Основы теории педагогических систем // Проблемы и методы психолого–педагогического обучения техн. обуч. систем. Воронеж: Изд-во Воронежского ун-та, 1977. 304 с.
6. Майоров А. И. Тесты школьных достижений: конструирование, проведение, использование. СПб: Образование и культура, 1997. 304 с.
7. Михайлычев Е. А. Дидактическая тестология. М.: Народное образование, 2001. 432 с.
8. Чельшкова М. Б. Теория и практика конструирования педагогических тестов: учеб. пос. М.: Логос, 2002. 432 с.
9. Кларин М. В. Инновационные модели обучения в зарубежных педагогических поисках. М.: Арена, 1994. 222 с.
10. Максимова О. А., Пермяков О. Е. Основы технологии комплексной экспертизы качества педагогических тестов: монография. Томск: ТОИПКРО, 2008. 100 с.
11. Максимова О. А., Пермяков О. Е. Процедуры комплексной экспертизы качества тестовых заданий и тестов при формировании банка данных // Журнал научных публикаций аспирантов и докторантов. Курск, 2008. № 4. 110–114.

Максимова О. А., кандидат педагогических наук, доцент кафедры.

Томский государственный педагогический университет.

Ул. Киевская, 60, Томск, Россия, 634061.

E-mail: Maksimova@tspu.edu.ru

Червонный М. А., кандидат педагогических наук, доцент, зав. кафедрой.

Томский государственный педагогический университет.

Ул. Киевская, 60, Томск, Россия, 634061.

E-mail: mach@tspu.edu.ru

Материал поступил в редакцию 21.11.2011.

O. A. Maksimova, M. A. Chervonny

PECULIARITIES OF DEVELOPMENT AND COMPREHENSIVE EXAMINATION OF THE QUALITY OF EDUCATIONAL TESTS USED FOR MONITORING AND EVALUATION OF THE QUALITY OF EDUCATION OF RURAL STUDENTS

The article deals with scientific and methodological foundations of development and technology complex examination of pedagogical tests for monitoring and evaluating the quality of teaching in rural schools. The authors show features of application testing technology in the learning process. The authors offer approaches to improve the quality of educational tests.

Key words: *educational testing, testing technology, criteria for the quality of teacher testing, complex examination of the quality of the test, taxonomy of educational objectives.*

Maksimova O. A.

Tomsk State Pedagogical University.

Ul. Kievskaya, 60, Tomsk, Russia, 634061.

E-mail: Maksimova@tspu.edu.ru

Chervonny M. A.

Tomsk State Pedagogical University.

Ul. Kievskaya, 60, Tomsk, Russia, 634061.

E-mail: mach@tspu.edu.ru