

СОВРЕМЕННОЕ ЯЗЫКОВОЕ ОБРАЗОВАНИЕ И ТЕСТИРОВАНИЕ LANGUAGE EDUCATION AND TESTING



Трапицын Сергей Юрьевич

доктор педагогических наук, профессор, заведующий кафедрой управления образованием и кадрового менеджмента, Российский государственный педагогический университет имени А.И. Герцена, Российская Федерация, 191186, Санкт-Петербург, набережная реки Мойки, д. 48, trapitsin@gmail.com
ORCID: 0000-0003-3752-8848
SPIN-код: 1582-2439



Куновски Марина Николаевна

кандидат филологических наук, доцент, заведующий кафедрой русского языка № 2 института русского языка, Российский университет дружбы народов, Российская Федерация, 117198, Москва, ул. Миклухо-Маклая, д. 10, стр. 3, kunovski-mn@rudn.ru
ORCID: 0000-0002-5872-2052
SPIN-код: 7029-6172



Поморцева Наталья Владимировна

доктор педагогических наук, доцент, заведующий кафедрой русистики, этноориентированной педагогики и цифровой дидактики института русского языка, Российский университет дружбы народов, Российская Федерация, 117198, Москва, ул. Миклухо-Маклая, д. 10, стр. 3, pomortseva-nv@rudn.ru
ORCID: 0000-0003-4224-8138
SPIN-код: 5744-8264

DOI: 10.22363/3034-2090-2025-11-2-86-96

EDN: ODLWGV

УДК 37.012.3

Научная статья

Педагогический тест как инструмент оценки качества образования

С.Ю. Трапицын¹ , Н.В. Поморцева² , М.Н. Куновски²

¹Российский государственный педагогический университет им. А.И. Герцена, Санкт-Петербург, Российская Федерация

²Российский университет дружбы народов, Москва, Российская Федерация

trapitsin@gmail.com

Аннотация. Приведены результаты анализа диагностических возможностей и областей применения педагогических тестов как инструмента оценки качества образования. Особое внимание уделено требованиям к тестам, обеспечивающим объективное и надежное измерение результатов образования,

© Трапицын С.Ю., Поморцева Н.В., Куновски М.Н., 2025



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
<https://creativecommons.org/licenses/by-nc/4.0/legalcode>

сформулированы критерии оценки качества тестовых систем. Показано, что эффективность применения педагогических тестов достигается четкой постановкой целей тестирования, осознанным и аргументированным определением предмета оценивания, соблюдением правил отбора и конструирования тестовых заданий, а также применением релевантных методов обработки и интерпретации результатов тестирования. Представлена типология тестов, дана характеристика целевого предназначения, специфики построения и областей применения различных типов тестов. Определены перспективы развития педагогической тестологии.

Ключевые слова: педагогические измерения, качество знаний, требования к педагогическому тесту

Вклад авторов: нераздельное соавторство.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

История статьи: поступила в редакцию 20.02.2025; одобрена после рецензирования 20.03.2025; принята к печати 22.03.2025.

Для цитирования: Трапицын С.Ю., Поморцева Н.В., Куновски М.Н. Педагогический тест как инструмент оценки качества образования // Русский тест: теория и практика. 2025. Т. 11. № 2. С. 86–96. <https://doi.org/10.22363/3034-2090-2025-11-2-86-96> EDN: ODLWGV

Введение

Управление качеством образования невозможно без наличия полной, достоверной и актуальной информации об образовательных процессах, их результатах и условиях их осуществления. Значимость и актуальность решения этой задачи возрастает с развитием цифровых технологий и автоматизированных средств контроля. В этой связи одним из перспективных направлений развития педагогической науки и практики является педагогическая текстология. Использование тестов в качестве средств оценки академических достижений, несмотря на уже довольно длительную историю их применения, до сих пор вызывает неоднозначную реакцию в педагогическом сообществе и широкой общественности, порождая множество мифов и толкований. Вместе с тем, представляется, что негативная реакция на все большее распространение тестирования как одного из простых и достаточно эффективных инструментов оценки качества образования вызвана, в нередких случаях, не слишком отчетливыми и компетентными представлениями о методологии, методах и средствах педагогического тестирования, а также низким качеством самих тестов, несоблюдением базовых принципов их создания и применения. Еще один

источник заблуждений и критики кроется в непонимании того факта, что никакой контроль, даже самый совершенный и объективный, не способен изменить (улучшить) качество образования, как не способен термометр изменить температуру. Любой педагогический контроль, и тест не исключение, выступает лишь инструментом для получения информации о степени освоения учащимися определенного содержания, при этом тест — инструмент гораздо более точный, чем излюбленные некоторыми педагогами критерии оценивания типа «дал достаточно полный и развернутый ответ», «глубоко понимает суть излагаемого вопроса» или «допускает незначительные ошибки и неточности» и пр., позволяющие им выступать в роли объективного и беспристрастного (так ли?) арбитра, оценивающего качество знаний учащегося (но, увы, не свою способность их передать).

В этой связи перед исследователями, специалистами, педагогами-практиками возникает вопрос — что представляет собой педагогический тест и каков используемый в нем механизм оценки? Попытка ответа на данный вопрос и представляет собой **цель исследования**.

Педагогическое тестирование привлекает внимание исследователей как с точки

зрения истории его возникновения, становления и развития в России (Ганин, Ганина, 2012) как педагогического явления и научного направления, начиная с введения самого термина «тест», эволюции тестирования как метода контроля обученности, вплоть до появления Единого государственного экзамена как общеобязательного государственного метода контроля знаний, так и с точки зрения исследования тестовой диагностики качества образования, анализа современного состояния педагогического тестирования, его достоинств и недостатков (Вербицкий, Креславская, 2012). Педагогическое тестирование активно адаптируется под новые форматы представления заданий. В.Ю. Переверзев еще в 2005 г. описал подходы к созданию тестовых заданий для компьютерного формата предъявления, методы отбора содержания и технологию разработки спецификации педагогических тестов для различных дисциплин¹.

В работе М.Б. Чельшковой и В.И. Звонникова, посвященной современным средствам оценивания результатов обучения, представлены теоретические и методические подходы к конструированию педагогических тестов, анализ их функций и возможностей применения, использование тестов на различных уровнях: в школьных контрольно-оценочных системах, на выпускных экзаменах, в системах мониторинга качества образования². М.Б. Чельшкова считает тест научно обоснованным методом оценки знаний учащихся. Она подчеркивает важность теоретических основ и научных методов при конструировании педагогических тестов, при отборе содержания, разработке формы заданий в соответствии с требованиями тестовой технологии, при этом важно ориентироваться в первую очередь на уровень знаний учащихся, на сохранение мотивации к обучению. Важным этапом является анализ результатов тестирования: на основе

полученных эмпирических данных проводится статистическое обоснование качества теста и дальнейшее его совершенствование³.

И.Л. Пасховер дает определение термина «педагогический тест», а также описывает ключевые критерии для их классификации и детально рассматривает этапы создания тестов и принципы выбора содержания и формы тестовых заданий (Пасховер, 2011). Классификация тестовых материалов по видам контроля, содержанию тестовых заданий, методологии оценивания, формату предъявления и масштабу применения представлена С.В. Пашниным (Пашнин, 2009).

В.С. Аванесов анализирует феномен вероятности правильного ответа на тестовые задания, подчеркивая, что использование одной и той же системы заданий для различных целей и уровней подготовленности испытуемых недопустимо. Такой подход приводит к увеличению погрешности измерений и, как следствие, к снижению их качества; достижение абсолютной объективности в педагогическом тестировании невозможно, поскольку измерения неизбежно включают субъективные компоненты. Для повышения уровня объективности при разработке тестовых материалов необходимо целенаправленное снижение субъективных факторов (Аванесов, 2008). Исследователь акцентирует внимание на образовательном потенциале тестов как инструмента интеллектуального развития, стимулирующего активизацию учебного процесса и улучшение качества знаний (Аванесов, 2004). Педагогическую диагностику как предпосылку, условие и результат обучения как фактор мотивации в обучении рассматривает и К. Ингенкамп, говоря о роли и функции педагогического контроля в учебном процессе (Ингенкамп, 1991).

Исследователи акцентируют внимание на таких аспектах разработки и применения

¹ Переверзев В.Ю. Технология разработки тестовых заданий : справочное руководство : учеб. пособие для студ. образ. учреждений сред. проф. образования. М. : Е-Медиа, 2005. 265 с. EDN: QUVHKK

² Звонников В.И., Чельшкова М.Б. Современные средства оценивания результатов обучения : учеб. пособие для студ. выс. учеб. заведений, обуч. по пед. спец. М., 2008. 18 с. EDN: QWPIVR

³ Чельшкова М.Б. Теория и практика конструирования педагогических тестов : учеб. пособие. М. : Логос, 2002. 432 с.

тестов в образовательной практике, как стандартизация и нормализация, точность параметризации, статистическое обоснование гипотез тестирования, анализ регрессионных и корреляционных зависимостей, а также шкалирование результатов тестирования. Эти элементы имеют ключевое значение для эффективного использования тестов в процессе аттестации учащихся (Майоров, 1997; Нейман, Хлебников, 2000). Б.П. Битинас выделяет несколько основных функций педагогической диагностики: функцию обратной связи, которая предоставляет преподавателю информацию для корректировки педагогической траектории; функцию оценки результативности педагогической деятельности; функцию прогнозирования последующей работы на основе полученных данных. Педагогическая диагностика направлена на повышение качества образовательного процесса с особым акцентом на развитие личности обучающихся, учитывая образовательный потенциал тестов (Битинас, 1993).

Методология и методы педагогического тестирования

Согласно положениям интеррогативной логики, структура любого тестового вопроса (задания) должна содержать в себе две основные части: содержательную (знаниевую) и собственно вопрошающую. При этом первая часть может быть истинным, неполным (требующим дополнения) или ложным высказыванием, вторая, собственно вопрос (задание), определяет характер действия, которое должен совершить тестируемый. Такая структура педагогического теста задает базовые требования к нему:

- *содержательная валидность* — соответствие содержательной (знаниевой) части теста содержанию изучаемого предмета (раздела, темы);
- *функциональная валидность* — соответствие теста цели контроля (распознавание, демонстрация знания, понимание, объяснение и пр.);

- *определенность* — относится, скорее, к вопрошающей части теста, задавая понятные тестируемому требования к действию, которое он должен выполнить;

- *простота* — содержательная часть теста должна формулироваться в лаконичной форме, ее представление должно в явном виде отражать содержание проверяемого фрагмента предметной области, а определяющий признак — быть необходимым и достаточным;

- *однозначность* — тест не должен содержать неоднозначных трактовок, допускать множественные интерпретации, а также должен включать четкие критерии и правила оценивания;

- *надежность* — точность, воспроизводимость и устойчивость результатов при повторном тестировании одного и того же обучающегося (группы), а также устойчивость к случайным ошибкам и погрешностям измерений;

- *дифференцирующая способность* — тест должен позволять обнаруживать статистически значимые различия как в уровне освоения отдельных фрагментов оцениваемого содержания и/или выполняемых действий, так и в уровне подготовки обучающихся.

Разнообразие целей, пользователей, используемых методов и способов обработки данных определяет специфику систем тестирования при соблюдении общих требований к ним, а именно:

- 1) тест должен соответствовать характеру решений по улучшению качества образования, которые будут приниматься по его результатам;

- 2) результаты тестирования должны оцениваться одновременно с анализом возможностей, которые имели учащиеся при освоении проверяемого содержания. При этом одинаковое внимание должно уделяться как оценке образовательных достижений, так и оценке возможностей их достичь;

- 3) характеристики получаемых данных должны соответствовать предмету оценивания.

Все декларируемые как измеряемые элементы содержания должны действительно

измеряться, задания должны быть валидны по отношению к проверяемым результатам обучения. Задания, используемые при оценке, должны давать устойчивые результаты. Они должны быть аутентичны, отражать ситуации, с которыми учащиеся встретятся в профессиональной и личной жизни. Результаты должны быть надежными. В процессе тестирования учащиеся должны иметь равные возможности продемонстрировать свои достижения;

4) оценки по результатам тестирования должны быть справедливыми для всех учащихся, независимо от их принадлежности к отдельным группам, дифференцируемым по различным характеристикам (полу, месту проживания, национальности и др.);

5) выводы о достижении планируемых результатов и о том, насколько реально этих результатов достигнуть, должны быть обоснованы не только полученными данными, но и анализом и описанием условий, при которых это становится возможным. Например, более высокие результаты определенной группы учащихся могут быть объяснены тем, что учащиеся имели высокий уровень мотивации к обучению, их обучали выдающиеся педагоги, они имели большее количество часов на освоение предмета и т.д.

Результаты исследования

Бурное развитие и широкое распространение тестов в связи с развитием онлайн-курсов и дистанционного обучения привело к ситуации, когда значительное их число не отвечает указанным требованиям и тестами, по сути, не является, что умножает число противников и критиков тестирования. Ситуация усугубляется, когда (в нередких случаях) не только разработчики тестовых программ и средств, но и использующие их педагоги не вполне осознают цели тестирования, условия его применения, ограничения, диагностические возможности.

Анализ актуальной практики оценки качества образования показывает, что ее проведение существенно осложняется отсутствием

ясных формулировок и измерителей целей, планируемых к достижению в процессе обучения. Несомненно, образовательные стандарты задают определенные ориентиры в отношении личностных, метапредметных и предметных результатов, такие, например, как освоение социальных норм, правил поведения, ролей и форм социальной жизни; владение основами самоконтроля, самооценки, принятия решений; понимание, интерпретация и комментирование текстов и т.п., однако столь общие формулировки создают широкое пространство для их интерпретации каждым отдельным педагогом: «показал хорошие знания», «продемонстрировал способность применять знания на практике», «умеет решать нестандартные задачи», «проявляет интерес к изучению» и пр.

При такой постановке целей не вполне понятно, какие именно задания теста позволяют однозначно, полно и точно оценить вышеприведенные характеристики качества знаний; содержание, степень сложности и последовательность вопросов определяются каждым педагогом интуитивно, и часто не лучшим образом. Нет единого мнения и о том, какое количество вопросов требуется для проверки того или иного объема учебного материала. Соответственно, оцениваются разные показатели качества знаний, что снижает диагностическую ценность системы проверочных заданий.

Отметим, что это касается не только тестов, любой из применяемых педагогами методов и форм контроля качества знаний имеет свои преимущества и недостатки, а также свои ограничения. Недостатком устного экзамена является высокая степень неопределенности и случайности, полученная на экзамене оценка часто субъективна. В то же время такая форма контроля обладает достаточной гибкостью и адаптивностью, педагог имеет возможность оперативно реагировать на ответ, задавать уточняющие и/или дополнительные вопросы, осуществлять непосредственный контакт с обучающимся, создавая мотивирующий эмоциональный фон. Письменный

контроль, хотя и более объективен, чем устный экзамен, не дает педагогу возможности непосредственно следить за процессом выполнения задания. Кроме того, письменная форма контроля, помимо знания предмета, требует от обучающегося владения письменной речью, грамотного выражения своих мыслей, наличие способности к поиску, отбору, структурированию информации, ее критическому анализу и пр.

Оценка качества знаний непосредственно педагогом связана с еще одной проблемой — субъективным восприятием справедливости оценки обучающимися. Достаточно часто они склонны объяснять низкую оценку предвзятостью педагога. Результаты исследований свидетельствуют, что у хорошо успевающих обучающихся совпадение между собственной оценкой и оценкой, выставленной педагогом, бывает в 46 % случаев, а у слабо успевающих — лишь в 11 %. Решение данной проблемы требует развития у обучающихся рефлексивных умений, способности к адекватной самооценке.

Говоря о целях использования педагогических тестов как средства оценки качества образования, следует четко определить объект и предмет оценивания. Согласно стандарту ИСО 9000:2015, качество есть степень соответствия совокупности характеристик объекта требованиям. Очевидно, что объектом педагогического тестирования является обучающийся (группа обучающихся), характеризующийся большим количеством признаков (свойств). Очевидно также, что предметом оценивания выступает, в общем случае, уровень его обученности (качество результатов обучения). При этом набор диагностируемых свойств (признаков) может быть различным, в частности, по содержательным характеристикам знаний: полнота, глубина, прочность, системность и т.п., по предметной области: математические, естественно-научные, гуманитарные и т.д., по учебным действиям: поиск информации, анализ, сравнение, классификация, аргументация и пр.

Другой целью педагогического контроля может быть определение проблемных областей образовательного процесса, например, разделов и тем, которые вызывают наибольшие трудности в освоении, различия в качестве преподавания, используемых технологиях, условиях обучения.

Многие исследователи утверждают, что сам факт педагогического контроля влияет на мотивацию обучающихся. Вопреки утверждениям критиков тестирования, исследования показали, что это справедливо и для педагогических тестов. В этих исследованиях были получены интересные результаты, в частности, зафиксировано возрастание у учащихся самоуважения и чувства собственного достоинства, интереса к обучению, стремление к более тщательной подготовке и т.п. По данным П. Бродфут, М. Кроссли и др., в процессе конструирования тестовых заданий школьные учителя значительно повысили свое профессиональное мастерство (Broadfoot, 1979; Crossley, Broadfoot, Schweisfurth, 2007). П. Бродфут приводит также данные, что, например, тесты на оценку уровня мышления обучающихся в процессе решения проблем и нахождения нестандартных решений позитивно отразились на их академических достижениях и привели к росту познавательного интереса непосредственно в учебном процессе.

Наконец, целью тестирования может быть оценка качества самих тестов, их диагностирующей и дифференцирующей способности, валидности, надежности и пр.

Таким образом, целями педагогического тестирования могут выступать:

- определение степени соответствия результатов обучения требованиям ФГОС;
- сопоставление индивидуальной академической успешности с аналогичными показателями группы;
- оценка динамики изменения индивидуальной и групповой академической успешности;
- создание тиражируемой базы качественных тестовых заданий, облегчающих педагогам проведение оценочных процедур;

- проведение сопоставительных исследований качества образования на межстрановом, страновом, региональном уровне, уровне образовательной организации.

В зависимости от этих целей можно выделить базовые подходы к тестированию: нормативно-ориентированный, критериально-ориентированный и личностно-ориентированный.

Нормативно-ориентированное тестирование применяется, как правило, в селективных системах обучения, когда требуется осуществить отбор какой-либо группы (обычно демонстрирующей лучшие результаты) обучающихся, например, вступительный экзамен в вуз, проведение предметных олимпиад и т.п., или, наоборот, выделить так называемую «группу риска» — обучающихся, демонстрирующих стабильно низкие результаты обучения. Данный подход используется также для сравнения в определяемых заданиями теста области содержания академических достижений одного обучающегося с другими, групп обучающихся между собой (например, изучающих одинаковые предметы в разных вузах, контрольной и экспериментальной групп в педагогическом эксперименте и пр.). Кроме того, с помощью нормативно-ориентированного теста можно сравнивать сложность тестовых заданий, степень освоения обучающимися различных фрагментов учебного материала и т.д. Результаты нормативно-ориентированного теста базируются, как правило, на статистических данных, полученных на достаточно представительной выборке испытуемых.

Критериально-ориентированное тестирование значительно чаще применяется в педагогическом контроле и предназначено, преимущественно, для диагностики уровня освоения обучающимися учебного материала, определения соответствия стандарту, аттестации обучающихся, выставления академической отметки. Каждому значению оценочной шкалы (уровню академической успешности) соответствует определенный результат прохождения теста, фиксируемый обычно в проценте правильно выполненных заданий. Для

интерпретации тестовых баллов требуется детализированная спецификация области проверяемого содержания. Возможна также модель построения критериально-ориентированного теста, когда каждой позиции оценочной шкалы соответствуют тестовые задания различного уровня сложности, позволяющие оценить степень соответствия обучающегося данному уровню (критериальному требованию).

Конструирование заданий для критериально-ориентированного теста требует более детальных и точных правил и спецификации проверяемой области содержания. В зарубежной литературе их нередко называют *domain-referenced test*. Существуют и так называемые *mastery tests* критериально-ориентированные тесты. Они применяются для дифференциации обучающихся на освоивших и не освоивших какую-либо область содержания предмета. Таким образом, способность теста предоставлять информацию об усвоении обучающимися определенной области содержания и делает его критериально-ориентированным, а не факт использования порогового значения или других критериев дифференциации.

Личностно-ориентированное тестирование имеет целью оценку прогресса обучающегося и находит широкое применение в системах мониторинга качества образования. Это может быть повторное прохождение одного и того же теста для оценки прочности усвоения учебного материала (так называемых «остаточных знаний») или тест с усложняющимися заданиями, позволяющими диагностировать переход обучающегося с одного качественного уровня освоения учебного материала на более высокий, например, с репродуктивного (запоминание и воспроизведение) к продуктивному (понимание и применение) или творческому (создание нового, решение нестандартных задач).

В целом педагогическое тестирование является эффективным средством, информирующим всех заинтересованных лиц о качестве образования. Оценка образовательных достижений, проводимая с помощью тестов,

являясь систематически повторяемым процессом, включает 4 составляющие: цель тестирования, предмет и объект оценки, методы обработки данных и пользователей результатов тестирования.

При общем признании эффективности применения тестовых систем для оценки качества образования исследователи и специалисты отмечают ряд проблем, которые снижают качество оценивания на основе тестов:

1) чрезмерное увлечение тестированием приводит к значительному росту нагрузки на педагогов, что нередко становится причиной их негативного отношения к самой идее тестирования и вызывает скрытое или явное сопротивление введению тестовых систем оценивания;

2) существуют сложности с апробацией тестов, их экспертизы, выравнивания диагностирующих возможностей. Трудности апробации заданий и теста в целом, проявляется также в низкой адаптивности тестов к быстро меняющимся условиям;

3) тесты часто критикуют за то, что они отходят от критериально-ориентированного подхода в определении качества знаний и реализуют, скорее, нормативно-ориентированную модель дизайна тестов, т.е. выполняют, скорее, надзорную, нежели развивающую и поддерживающую функцию;

4) сложность интерпретации результатов тестирования — что использовать для оценки качества образования: сырые баллы за выполнение теста или сложный индикатор добавленной стоимости, в основе которого лежит метод многоуровневого анализа результатов обучающегося (группы, образовательной организации) одновременно с их контекстуальными характеристиками;

5) существуют проблемы сравнения и учета данных, полученных в результате текущего контроля, и данных итогового тестирования.

Одна из основных проблем педагогического тестирования и главный аргумент критиков широкого использования тестов в практике оценивания качества образования — вопрос доверия к результатам, способности теста релевантно оценивать качество знаний. Эти

характеристики теста напрямую зависят от его воспроизводимости и могут быть статистически оценены, т.е. требования к качеству тестов определяют необходимость проведения статистического анализа тестовых заданий. Основаниями для проведения такого анализа служат ситуации, когда хорошо подготовленный обучающийся может, в силу различных причин, не справиться с тестовым заданием, или, наоборот, когда слабый обучающийся чисто случайно дает правильные ответы на все вопросы теста. Кроме того, каждое тестовое задание по мере его использования приобретает некоторую статистику, позволяющую делать выводы о его содержательной и функциональной валидности и надежности в зависимости от доли успешно или неуспешно выполнивших его тестируемых. В общем случае, эта статистика должна соответствовать нормальному закону распределения. Невыполнение этого условия может (должно) служить поводом для серьезного и глубокого анализа качества тестового задания. Отметим, что при таком подходе вероятность традиционных ошибок педагогического оценивания («ошибки великодушия», «гибкого эталона», «ошибки ореола», «центральной тенденции» и пр.) достаточно мала, хотя и не исключена (например, преднамеренное или непреднамеренное упрощение (усложнение) педагогом тестовых заданий, использование заданий с «квазивыбором», субъективное определение правила перевода результатов выполнения теста в академическую шкалу оценивания и т.п.).

Технология разработки тестов

Для исследователей и специалистов особый интерес представляет технология разработки тестов. Поскольку в соответствии с законом «Об образовании» каждый выпускник должен подтвердить достижение образовательных стандартов для получения аттестата или диплома, требуется создание простой и надежной системы оценки степени достижения этих стандартов. При этом необходимо найти баланс между аутентичностью

и технологичностью оценивания. Для обеспечения такого баланса могут использоваться различные подходы в рамках классической и современной теорий тестов, такие, например, как технология «якорных» тестов (common anchor technics), дискриминантный анализ (DIF — Differential Item Functioning) и пр.

Общими требованиями к разработке тестов выступают:

- 1) соответствие программе. Содержание тестов должно максимально охватывать содержание программы;
- 2) ориентация на лучший опыт преподавания и обучения;
- 3) проверка (оценка) всех возможных уровней достижения;
- 4) учет реальных условий.

Одним из основных требований к тесту как измерителю учебных достижений является его способность к дифференциации учащихся различного уровня подготовки на основе результатов выполнения теста. Для надежного установления различий в уровнях академической успешности распределение баллов тестовых заданий должно быть как можно более равномерным по всей шкале. Для этого в тест включают как трудные, так и легкие задания. В общем случае сложность заданий теста может варьировать от 10 до 90 %, целесообразно включать в тест также очень простые (< 10 %) и очень трудные (> 90 %) задания. Легкие задания поддерживают мотивацию учащихся при выполнении теста, трудные задания, при условии, что они соответствуют требованиям программы, выявляют неосвоенные элементы образовательных стандартов. Вместе с тем их число не должно быть большим, так как их включение снижает эффективность теста как измерительного инструмента.

Для поддержания одинаковых стандартов (обеспечения равных условий) при проектировании тестов необходимо соблюдать следующие правила:

- 1) отбирать в тест задания в строгом соответствии с установленными критериями оценивания;
- 2) осуществлять выравнивание тестов по сложности;

- 3) тщательно анализировать содержание теста, его соответствие содержанию пройденного материала;

- 4) устанавливать проходной балл (если это необходимо) на основе статистики результатов тестирования по каждому уровню.

Описание уровней достижений — отправная точка отсчета как для отбора проверяемого содержания, так и для разработки отдельных заданий и схемы их проверки, а также для установления стандартов достижений. В спецификации к тесту следует давать описание уровней достижений исходя из того, что знает и может продемонстрировать «типичный» учащийся данного возраста по данному предмету.

Одним из способов повышения надежности теста является модель, в которой при достаточно большом числе тестовых заданий их можно разбить на две части. В этом случае устойчивая статистика снижения результатов тестирования во второй части теста может свидетельствовать об утомляемости студентов, снижении способности к концентрации внимания и т.п. Поэтому более продуктивной будет методика оценки статистики результатов выполнения четных и нечетных заданий. У одинаковых по качеству знаний обучающихся результаты выполнения этих двух групп заданий не должны иметь статистически значимых различий, и в то же время они должны отличаться для обучающихся разного уровня подготовки.

Заключение

Стратегические приоритеты современного образования и тенденции его развития требуют изменений в системе оценки качества образования. К перспективным направлениям совершенствования системы оценивания учебных достижений, безусловно, относится развитие педагогической тестологии как научной дисциплины и сферы педагогической практики. При этом парадигмальными основаниями для разработки эффективных систем педагогического тестирования должен стать переход

- от оценки отдельных знаний и умений к комплексной и междисциплинарной оценке, от оценки предметных знаний к измерению метапредметных и надпредметных результатов и универсальных учебных действий;

- от одномоментной оценки к оценке динамики учебных достижений (добавленной ценности, прогресса);

- от оценки индивидуальных академических результатов к оценке способности работать в команде, приходиться к коллегиальному решению;

- от оценки исключительно результатов обучения к анализу всей совокупности влияющих на них факторов,

а также переход к аутентичной оценке, использование контекстуализированных заданий, заданий с несколькими правильными ответами, предоставление обучающемуся возможности активного конструирования содержания ответа.

В заключение приведем ряд концептуальных идей, которые следует учитывать при реализации перспективных направлений педагогической тестологии.

1. Ни один из методов оценки качества знаний не является универсальным, перспективный подход в этом отношении — принцип комплиментарности.

2. В процессе педагогических измерений сложно реализовать оценивание всей совокупности качественных признаков, поэтому важно определить, что конкретно, какие именно характеристики, определяющие образовательные достижения, выступают предметом оценки. Результаты тестирования дают лишь некоторую (неполную) информацию о качестве знаний, и потому на основе теста, хотя и обеспечивающего относительную объективность оценива-

ния, нельзя давать категоричные заключения об уровне академической успешности.

3. Качество знаний — это результат обучения, которое во многом определяется личностью учителя, условиями образовательного процесса и другими факторами. Интерпретация результатов тестирования без учета этих факторов не может быть действительно объективной.

4. Любое измерение всегда включает ошибку измерения, что требует анализа статистики результатов тестирования, как в отношении теста, так и в отношении обучающегося.

5. Отсутствие однозначно определенной единицы измерения качества знаний может приводить к неверной интерпретации результатов тестирования (например, получение одинакового процента правильных ответов (баллов) за выполнение заданий разного уровня сложности).

6. Измеряемые тестом показатели качества знаний должны коррелировать с другими наблюдаемыми характеристиками обучающегося (например, с его текущими отметками, активностью на занятиях и пр.).

7. На результаты тестирования оказывает значимое влияние мотивация и психологическое состояние обучающегося. Это требует создания возможности оценивать готовность обучающегося к тестированию до начала его проведения.

8. Педагогический контроль, в основе которого лежат психометрика и теория тестов, в перспективе должен смениться новым типом измерений, которые в большей степени направлены на оценку прогресса обучающегося, помощь ему и академическое консультирование, чем на проверку степени его соответствия некоему эталону.

Список литературы

Аванесов В.С. Основы педагогической теории измерений // Педагогические измерения. 2004. № 1. С. 15–21.

Аванесов В.С. Проблема объективности педагогических измерений // Педагогические измерения. 2008. № 3. С. 3–40.

Битинас Б.П. Педагогическая диагностика: сущность, функции, перспективы // Советская педагогика. 1993. № 2. С. 10–15.

Вербницкий А.А., Креславская Е.Е. Тестирование в образовании: проблемы и перспективы // Педагогика. 2012. № 8. С. 3–13. EDN: PGWLYB

- Ганин В.В., Ганина Н.В. Становление педагогического тестирования в России (конец XIX — начало XXI века) // Знание. Понимание. Умение. 2012. №4. С. 127–132. EDN: PXADFD
- Ингенкамп К. Педагогическая диагностика. М. : Педагогика, 1991. 240 с.
- Майоров А.Н. Тесты школьных достижений: конструирование, проведение, использование. СПб. : Образование и культура, 1997. 304 с. EDN: FBDTPM
- Нейман Ю.М., Хлебников В.А. Введение в теорию моделирования и параметризации педагогических тестов. М. : Прометей, 2000. 169 с.
- Пасховер И.Л. Педагогический тест как инструмент системы оценки и контроля качества образования // Magister Dixit. 2011. №4. С. 175–185. EDN: OXAPON
- Пашиин С.В. Педагогические тесты в современном образовательном процессе, их классификация // Современная высшая школа: инновационный аспект. 2009. №4. С. 34–41. EDN: LPBQNX
- Broadfoot P. Assessment, schools and society (contemporary sociology of the school). Methuen Publishing Ltd, 1979. 156 p. <https://doi.org/10.4324/9780203808702> EDN: SIZRWX
- Changing educational contexts, issues and identities: 40 years of comparative education. 1st ed. / M. Crossley, P. Broadfoot, M. Schweisfurth (Eds). Routledge, 2007. 432 p. <https://doi.org/10.4324/9780203961995>

DOI: 10.22363/3034-2090-2025-11-2-86-96

EDN: ODLWGV

UDC 37.012.3

Research article

Pedagogical test as a tool for assessing the quality of education

Sergey Y. Trapitsyn¹  , Natalia V. Pomortseva² , Marina N. Kunovski² 

¹Herzen State Pedagogical University of Russia, *Saint Petersburg, Russian Federation*

²RUDN University, *Moscow, Russian Federation*

 trapitsin@gmail.com

Abstract. The results of the analysis of diagnostic capabilities and areas of application of pedagogical tests as a tool for assessing the quality of education are presented. Special attention is paid to the requirements for tests that ensure objective and reliable measurement of educational outcomes, and criteria for evaluating the quality of test systems are formulated. It is shown that the effectiveness of pedagogical tests is achieved through a clear definition of testing goals, a conscious and reasoned definition of the subject of assessment, compliance with the rules of selection and design of test tasks, as well as the use of relevant methods of processing and interpreting test results. The typology of texts is presented, the characteristics of the target purpose, the specifics of construction and the areas of application of various types of tests are given. The prospects for the development of pedagogical testology are determined.

Keywords: pedagogical measurements, quality of knowledge, requirements for the pedagogical test

Authors' contribution: undivided co-authorship.

Conflicts of interest. The authors declare no conflict of interest.

Article history: received 20.08.2025; approved after reviewing 20.09.2025; accepted 22.09.2025.

For citation: Trapitsin, S.Y., Pomortseva, N.V., & Kunovski, M.N. (2025). Pedagogical test as a tool for assessing the quality of education. *Russian Language: Research, Testing and Practice*, 11(2), 86–96. <https://doi.org/10.22363/3034-2090-2025-11-2-86-96> EDN: ODLWGV