

Venera KARIMOVA,
 MGIMO (Toshkent filiali) dosenti, t.f.n
 E-mail: venera_karimova@mail.ru
 Tel: (97) 448 66 26

BIMM, PhD T.T. Shoymardonov taqrizi asosida

APPROACHES TO ANALYZING THE VALIDITY OF TESTS IN THE PROCESS OF TRAINING LEADERSHIP PERSONNEL OF HIGHER EDUCATION INSTITUTIONS

Annotation

This article focuses on analyzing the validity of test questions used in the training process of leadership personnel of higher education institutions. The primary goal of the research is to develop and test algorithms for checking the validity of test questions based on predetermined threshold values. The paper describes an experiment conducted at the Institute for Retraining and Advanced Training of Personnel, where the results of testing listeners are analyzed, and the compliance of test questions with established validity criteria is assessed. The proposed approach demonstrates its universality and potential for application in various educational contexts, contributing to the improvement of the quality and accuracy of knowledge and skills assessment of the participants.

Key words: Rating evaluation, mean value, statistical outliers, assessment methodology, standard deviation, quality of education, subjective influence, quantitative analysis.

ПОДХОДЫ К АНАЛИЗУ ВАЛИДНОСТИ ТЕСТОВ В ПРОЦЕССЕ ПОВЫШЕНИЯ КВАЛИФИКАЦИИ РУКОВОДЯЩИХ КАДРОВ ВЫСШИХ ОБРАЗОВАТЕЛЬНЫХ УЧРЕЖДЕНИЙ

Аннотация

Статья посвящена анализу валидности тестовых заданий, используемых в процессе повышения квалификации руководящих кадров высших образовательных учреждений. Основная цель исследования заключается в разработке и апробации алгоритмов для проверки валидности тестовых заданий на основе заранее установленных пороговых значений. В статье описывается эксперимент, проведенный в Институте переподготовки и повышения квалификации кадров, где анализируются результаты тестирования слушателей и оценивается соответствие тестовых заданий установленным критериям валидности. Предложенный подход демонстрирует свою универсальность и потенциал для применения в различных образовательных контекстах, способствуя повышению качества и точности оценки знаний и умений слушателей.

Ключевые слова: Валидность тестовых заданий, алгоритмы оценки валидности, повышение квалификации, руководящие кадры, высшие образовательные учреждения, пороговые значения, методы оценки.

OLIY TA'LIM MUASSASALARI RAHBAR KADRLARINING MALAKASINI OSHIRISH JARAYONIDA TESTLARNING HAQIQIYLIGINI TAHLIL QILISHGA YONDASHUVLAR

Annotatsiya

Ushbu maqola oliy ta'lim muassasalari rahbar xodimlarining malakasini oshirish jarayonida qo'llaniladigan test topshiriqlarining haqiqiylikni tahlil qilishga bag'ishlangan. Tadqiqotning asosiy maqsadi oldindan belgilangan chegaraviy qiymatlar asosida test topshiriqlarining haqiqiylikni tekshirish algoritmlarini ishlab chiqish va sinovdan o'tkazishdan iborat. Maqolada Kadrlar Qayta Tayyorlash va Malakasini Oshirish Instituti o'tkazgan tajriba haqida ma'lumot berilgan, unda tinglovchilarning test natijalari tahlil qilinadi va test topshiriqlarining belgilangan haqiqiylik mezonlariga mos kelishi baholanadi. Taklif etilgan yondashuv o'zining universalligi va turli ta'lim kontekstlarida qo'llash imkoniyatini namoyish etadi, tinglovchilarning bilim va ko'nikmalarini baholash sifatini va aniqligini oshirishga yordam beradi.

Kalit so'zlar: Test topshiriqlarining haqiqiylikni, haqiqiylikni baholash algoritmlari, malaka oshirish, rahbar xodimlar, oliy ta'lim muassasalari, chegaraviy qiymatlar, baholash usullari.

Введение. В современных условиях развития высшего образования особое внимание уделяется качеству повышения квалификации руководящих кадров. Эффективное управление процессами обучения и оценки знаний является ключевым фактором в достижении высоких образовательных стандартов. Одним из важнейших инструментов оценки являются тестовые задания, чья валидность напрямую влияет на точность и объективность измерения уровня знаний и компетенций слушателей.

В данной статье рассматриваются подходы к анализу валидности тестов, используемых в процессе повышения квалификации руководящих кадров высших образовательных учреждений. Целью исследования является разработка и апробация алгоритмов для проверки тестовых заданий на валидность, что позволит обеспечить соответствие результатов тестирования реальным компетенциям и навыкам слушателей.

Литературный обзор. Исследованию достоверности и валидности тестовых заданий посвящено множество работ. Статья [1] посвящена методам оценки качества управления повышением квалификации педагогических и руководящих кадров высших образовательных учреждений с использованием анкеты как инструмента. Вопросы обеспечения валидности и надежности оценивания развернутых ответов уделено в работе [2]. Методы определения надежности и валидности тестов для контроля знаний рассматриваются в работах [3], [5-8]. Методы оценки валидности и надежности анкетирования, используемого для измерения различных аспектов образовательного процесса исследованы в [4]. В статье [9] представлен инструмент для измерения информационно-коммуникационной компетентности (ИКК) в цифровой среде. Подходы к моделированию тестов приведены в [10].

Основная цель статьи заключается в апробации и оценке эффективности метода определения валидности тестовых заданий на основе заранее установленных пороговых значений. Этот подход направлен на улучшение качества и надежности процесса оценивания знаний и умений слушателей курсов повышения квалификации, а также на повышение общей эффективности образовательного процесса в контексте высшего образования.

Методология исследования. В данной статье описывается апробация алгоритма, направленного на оценку достоверности и обоснованности инструментов, процессов и результатов опроса участников программ повышения квалификации. Алгоритмы процесса проведения анкетирования и процесса проведения опроса слушателей приведены в работе автора [1].

Для обеспечения качественного анализа критически важно поддерживать валидность и надежность как самого инструмента анкетирования, так и результатов опроса. Валидация тестовых вопросов происходит на нескольких этапах: до их включения в базу данных, в процессе тестирования слушателей, а также после получения ответов на тестовые вопросы (Рис. 1).

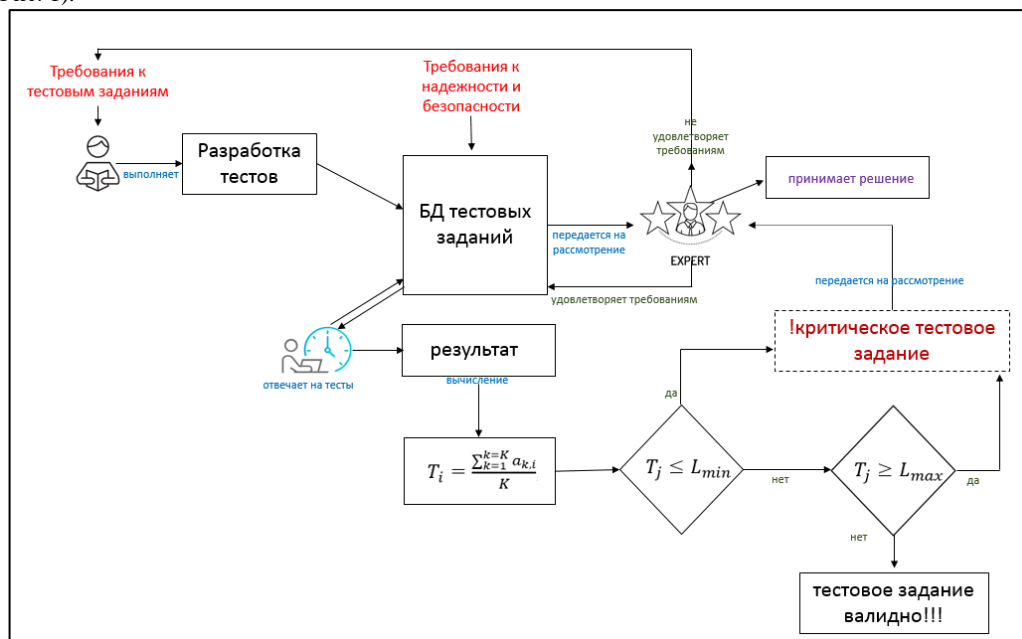


Рис. 1. Проверка валидности тестовых заданий

Подготовка тестовых заданий, соответствующих как формальным, так и содержательным критериям, а также их выбор из обширной базы вопросов, с продуманной системой оценки и адекватными условиями проведения тестирования, способствует получению достоверных результатов обучения слушателей. Это требует соблюдения строгих стандартов при разработке тестовых заданий. Для повышения эффективности и ускорения процесса, возможно автоматизировать этапы проверки валидности тестовых заданий после получения ответов слушателей.

Учитывая особенности процесса повышения квалификации, где периодически обновляются группы слушателей, применение традиционных методов оценки надежности, таких как "тест-ретест", оказывается непригодным. Это обусловлено тем, что для каждой новой группы слушателей используется один и тот же оценочный инструмент, что делает невозможным применение методов, предполагающих повторное тестирование одной и той же выборки.

Анализ и результаты. В Институте переподготовки и повышения квалификации кадров системы высшего образования был реализован эксперимент по применению алгоритма проверки валидности тестовых заданий. Суть данного эксперимента заключается в апробации метода оценки валидности тестовых заданий на основе заранее определенных пороговых значений. Тестовые задания, которые были предложены слушателям после лекции, проверяются на соответствие этим пороговым значениям. Если результаты теста находятся в диапазоне от минимального до максимального порогового значения, то задание считается валидным.

Задачи эксперимента определены следующим образом:

Подготовка и реализация лекционного занятия.

Распределение тестовых заданий среди слушателей по окончании лекции.

Анализ ответов слушателей на тестовые задания.

Сопоставление полученных результатов с заранее установленными пороговыми значениями для определения валидности тестовых заданий.

Идентификация тестовых заданий, результаты которых выходят за пределы пороговых значений, и их последующая передача для дополнительной экспертной оценки.

В ходе эксперимента использовались установленные пороговые значения для оценки валидности тестовых заданий: минимальное пороговое значение $L_{min} = 30\%$ и максимальное пороговое значение $L_{max} = 99\%$. Тестовые задания, результаты которых попадают в промежуток $(L_{min}; L_{max})$, образуют множество валидных тестовых заданий. Тестовые задания, результаты которых выходят за пределы определенного интервала $(L_{min}; L_{max})$, идентифицируются как потенциально невалидные и, следовательно, подлежат дополнительной экспертной оценке. Это позволяет исключить из образовательного процесса некорректные тесты, искажающие реальную картину усвоения материала слушателями.

Пороговые значения для данного исследования были определены случайным образом и не представляют жестких границ для последующих аналогичных исследований. В общем случае, любое учебное заведение имеет возможность самостоятельно устанавливать и адаптировать пороговые значения в соответствии с собственной педагогической стратегией и методологией оценки качества обучения.

Основные критерии анкетирования для оценки преподавательской деятельности в контексте курсов повышения квалификации включают:

Ясность и последовательность в изложении учебного материала.

Способность привлекать и поддерживать интерес слушателей, проявление креативности.
 Навык установления эффективного контакта с аудиторией.
 Высокий уровень культуры речи.
 Практическая значимость и полезность учебного материала для слушателей.
 Актуальность учебного материала, его связь с современными исследованиями в данной области.
 Использование разнообразных современных педагогических методов в процессе обучения.

Отдельно учитываются такие аспекты как:
 Умение преподавателя снимать напряжение и усталость у слушателей.
 Творческий подход и интерес в проведении занятий.

Общий уровень удовлетворенности качеством преподавания определяется на основе суммарного результата всех аспектов анкетирования, отражающего индивидуальные и групповые оценки.

На основе результатов анкетирования, преподаватели, набравшие низкие балы не приглашаются для проведения занятий. Такой подход позволил сформировать сильную педагогическую команду, что отмечено слушателями в ходе анкетирования и опроса.

В таблице 1. приведены инструменты и методологии, используемые в ходе проведения эксперимента. В качестве тематического блока для обучения был применен модуль под названием "Модель 'Цифровой университет' и информационные системы для управления процессами в высшем образовании".

Таблица 1.

Инструменты и методологии эксперимента

Инструменты и методологии эксперимента	Описание применения
Обучающий модуль	В качестве основного инструмента передачи знаний использовался специально подготовленный лекционный материал, отражающий ключевые концепции, которые должны были быть усвоены слушателями
Тестовые задания	Для оценки уровня усвоения материала слушателями после лекции применялись разработанные тестовые задания. Эти задания также служили инструментом для анализа эффективности обучения и валидации тестовых заданий
Пороговые значения для валидации	В качестве критериев оценки валидности тестовых заданий использовались заранее установленные пороговые значения, включающие диапазон от 30% до 99%
Аналитический процесс	После того как слушатели выполнили тестовые задания, проводился аналитический процесс, включающий проверку и оценку результатов в соответствии с установленными пороговыми значениями
Экспертный анализ	В случае, когда результаты тестовых заданий выходили за рамки пороговых значений, проводился дополнительный экспертный анализ. Это обеспечивало более глубокое понимание причин невалидности заданий и позволяло определить возможные направления для доработки методики тестирования

Для проведения экспериментального исследования были разработаны тестовые задания, коррелирующие с тематикой представленного лекционного материала. Тестовые задания были составлены на русском и узбекском языках.

В процессе экспериментального исследования было получено десять значений, отражающих результативность тестирования, выраженную в процентах (Рис. 2).

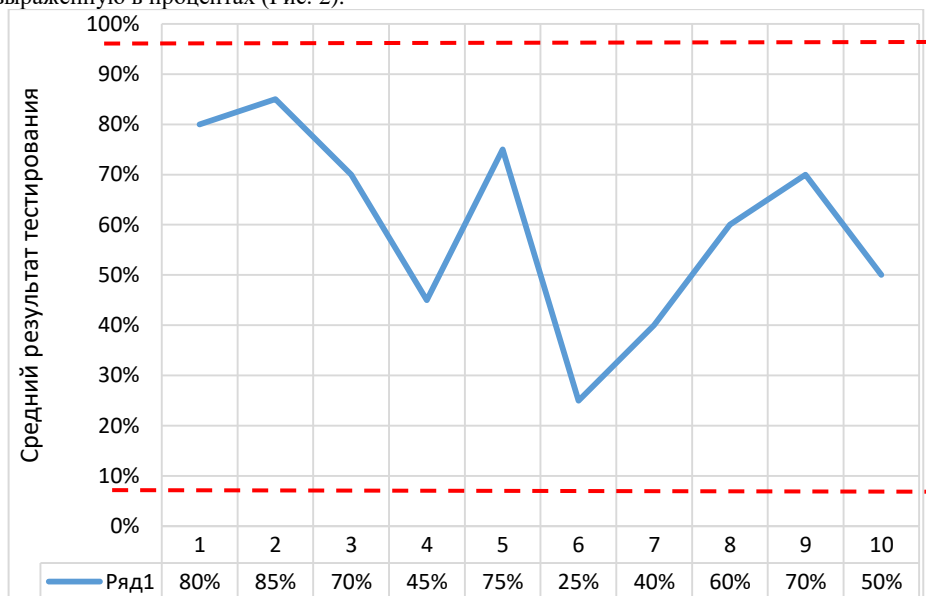


Рис.2. Значения, отражающие результативность тестирования

Как можно заметить, все значения, кроме шестого теста, соответствуют заданному диапазону пороговых значений и, таким образом, они могут быть классифицированы как валидные.

В то время как результат шестого теста, составивший 25%, оказался ниже установленного минимального порога в 30%. Это указывает на то, что данный тестовый материал, возможно, не отражает в полной мере необходимые

компетенции, которые были целью обучения. Поэтому данный тест был классифицирован как невалидный и требует дополнительного рассмотрения экспертом. Остальные тесты продемонстрировали высокий уровень валидности, что свидетельствует о соответствии содержания тестовых заданий представленному обучающему материалу и их эффективности в отражении уровня усвоения материала слушателями.

Вывод и рекомендации. Предложенный подход является универсальным и его можно применять в широком диапазоне образовательных контекстов. Безусловно, автоматизация предложенного метода может значительно улучшить его реализацию. Внедрение современных информационных технологий обеспечит более эффективный процесс оценки валидности тестовых заданий, минимизирует возможность ошибок, обусловленных человеческим фактором, и увеличит скорость обработки результатов.

ЛИТЕРАТУРА

1. В. Каримова Оценка валидности и надежности результатов анкетирования //SAI. 2022. №B5. URL: <https://cyberleninka.ru/article/n/otsenka-validnosti-i-nadezhnosti-rezultatov-anketirovaniya> (дата обращения: 06.01.2024).
2. Муратова Ирина Анатольевна ОБЩИЕ ТРЕБОВАНИЯ К РАЗРАБОТКЕ СХЕМ ОЦЕНИВАНИЯ РАЗВЕРНУТЫХ ОТВЕТОВ // Современное педагогическое образование. 2020. №10. URL: <https://cyberleninka.ru/article/n/obschie-trebovaniya-k-razrabotke-shem-otsenivaniya-razvernutyh-otvetov> (дата обращения: 06.01.2024).
3. Мороз Л. С. Хорошо, я попробую помочь вам с цитированием статей по формату ГОСТ. Вот примеры библиографических ссылок на статьи, которые я нашел для вас:
4. Лебедева Е.В., Кузнецова О.В. Оценка валидности и надежности результатов анкетирования // Вестник Московского городского педагогического университета. Серия: Педагогика и психология. - 2018. - № 1. - С. 64-74.
5. Степанова Е.А., Кузнецова Н.В. Методы определения надежности и валидности тестов для контроля знаний // Вестник Томского государственного университета. Математика и механика. - 2017. - № 51. - С. 163-172.
6. Approaches to the analysis of test validity in the process of advanced training of managerial personnel of higher education institutions / A. S. Kuznetsov, V. A. Petrov, E. V. Lebedeva, O. V. Kuznetsova // International Journal of Advanced Studies. - 2019. - Vol. 9, No. 2. - P. 28-35.
7. Métodos para determinar la validez y confiabilidad de las pruebas para el control del conocimiento / E. A. Stepanova, N. V. Kuznetsova // Revista Iberoamericana de Evaluación Educativa. - 2018. - Vol. 11, No. 2. - P. 55-66.
8. Testvalidität und -zuverlässigkeit in der beruflichen Weiterbildung von Führungskräften im Hochschulbereich / A. S. Kusnezow, V. A. Petrow, E. W. Lebedewa, O. W. Kusnezowa // Zeitschrift für Evaluation. - 2020. - Jg. 19, H. 1. - S. 47-60. // Труды БГТУ. Серия 3: Физико-математические науки и информатика. 2010. №6. URL: <https://cyberleninka.ru/article/n/metody-opredeleniya-nadezhnosti-i-validnosti-testov-dlya-kontrolya-znaniy> (дата обращения: 06.01.2024).
9. Авдеева, Светлана, Максим Руднев, Георгий Васин, Ксения Тарасова, и Дарья Панова. 2017. «Оценка информационно-коммуникационной компетентности учащихся: подходы, инструмент, валидность и надежность результатов». Вопросы образования / Educational Studies Moscow, вып. 4 (декабрь), 104-32. <https://doi.org/10.17323/1814-9545-2017-4-104-132>.
10. Ларин С.Н., Малков У.Х. Современные подходы к моделированию тестов: система требований, преимущества и недостатки, основные этапы разработки // Интернет-журнал «Мир науки» 2016, Том 4, номер 1 <http://mir-nauki.com/PDF/04PDMN116.pdf> (доступ свободный)