

ОЦЕНКА ВАЛИДНОСТИ И НАДЕЖНОСТИ РЕЗУЛЬТАТОВ АНКЕТИРОВАНИЯ**Венера Каримова**

к.т.н., доцент Московского Государственного института международных отношений
(Ташкентский филиал)

<https://doi.org/10.5281/zenodo.7005916>

***Аннотация.** Данная статья посвящена методам, используемым для оценки качества управления повышения квалификации педагогических и руководящих кадров высших образовательных учреждений. Для оценки достоверности проводимого анализа данных и обоснованности полученных выводов в статье оценивается точность и надежность используемого инструмента — анкеты. В статье предложены алгоритмы процесса проведения опроса слушателей курсов повышения квалификации и проверки валидности тестовых заданий. Построена математическая модель отбора допустимых тестовых заданий с использованием пороговых значений. Обеспечение валидности тестовых вопросов в работе происходит до поступления в базу данных тестовых вопросов, до тестирования слушателей и после ответов слушателей. Надежность анкеты в статье определяется путем измерения распределения ответов слушателей по заданным вопросам. Приведены критерии анкетирования слушателей. Уровень личностной и групповой удовлетворенности качеством преподавания определяется на основе суммарного итогового показателя по всем вопросам анкеты.*

***Ключевые слова:** надежность, валидность, анкета, повышение квалификации, оценка, опрос, алгоритм.*

ASSESSMENT OF THE VALIDITY AND RELIABILITY OF THE RESULTS OF THE QUESTIONNAIRE

***Abstract.** This article is devoted to the methods used to assess the quality of management of advanced training of pedagogical and managerial personnel of higher educational institutions. To assess the reliability of the data analysis and the validity of the conclusions obtained, the article evaluates the accuracy and reliability of the questionnaire tool used. The article proposes algorithms for the process of conducting a survey of students of advanced training courses and verifying the validity of test tasks. A mathematical model of the selection of acceptable test tasks using threshold values is constructed. The validity of test questions in the work is ensured before the test questions are entered into the database, before testing the listeners and after the answers of the listeners. The reliability of the questionnaire in the article is determined by measuring the distribution of listeners' responses to the questions asked. The criteria for the survey of listeners are given. The level of personal and group satisfaction with the quality of teaching is determined on the basis of the total final indicator for all questions of the questionnaire.*

***Keywords:** reliability, validity, questionnaire, professional development, assessment, survey, algorithm.*

ВВЕДЕНИЕ

Большинство методов, используемых для оценки качества управления повышения квалификации педагогических и руководящих кадров высших образовательных учреждений в качестве инструмента предполагают анкетирование. Но анкетирование, как

один из инструментов мониторинга, является достаточно сложным процессом, так как большое влияние на результат оказывает человеческий фактор. При проведении анкетирования слушателей повышения квалификации, необходимо критически подходить к составу опрашиваемых групп. Поэтому для оценки достоверности проводимого анализа данных и обоснованности полученных результатов необходимо оценить точность и надежность используемого инструмента — анкеты.

МАТЕРИАЛЫ И МЕТОДЫ

Обоснованность, или валидность (validity), – это эквивалентность измерений характеристикам измеряемого объекта. Другими словами, это мера соответствия оценок, получаемых в процессе измерения, представлениям о сущности свойств исследуемых объектов и их роли в исследуемых процессах. Надежность (reliability) – это характеристика, отражающая устойчивость и согласованность получаемых результатов измерения. Надежность является составляющей обоснованности (валидности) инструмента, это необходимый, но недостаточный элемент валидности [1, стр. 81].

Оценке валидности и надежности результатов анкетирования посвящено множество работ, в частности работы [1-9]. В данной статье предлагаются алгоритмы для оценки достоверности и обоснованности инструментов, процессов и результатов опроса слушателей центров повышения квалификации.

На рисунке 1 приведен алгоритм процесса проведения опроса слушателей курсов повышения квалификации, который включает следующие этапы:

- Постановка цели анкетирования;
- Разработка анкет для опроса слушателей;
- Проверка валидности разработанных анкет;
- Рецензирование анкет
- Проведение опроса;
- Обработка результатов опроса; слушателей
- Анализ результатов опроса слушателей;
- Анализ проверки выполнения поставленной цели;
- Принятие и утверждение управленческого решения.

Цель анкетирования зависит от поставленной задачи, например, оценка модулей обучения, оценка проведения занятий, оценка восприятия степени организации повышения квалификации и т.д.

Разработка анкет для опроса осуществляется преподавателями при оценке знаний слушателей по проводимым модулям или группой экспертов при оценке образовательных процессов и качества повышения квалификации. Это должны быть квалифицированные специалисты, четко понимающие цели проводимого анкетирования.

Особые требования предъявляются к аналитической группе, проводящей экспертизу и проверку валидности анкет. Аналитическая группа обосновывает содержательную валидность анкеты, и в случае отклонения, отправляет анкету на редактирование или обновление. Оценка и аргументация содержательной валидности анкет преимущественно имеет описательный характер. В случае положительного решения, начинается проведение опроса слушателей.

До проведения опроса необходимо провести инструктаж и объяснить слушателям цель анкетирования. Инструктаж осуществляется лицами, понимающими цель анкетирования и умеющих грамотно объяснить технологию проведения опроса.

После проведения опроса слушателей производится автоматизированная обработка результатов опроса, на основе которого проводится анализ результатов опроса и измеряется достижение цели. В случае достижения поставленной цели, принимается управленческое решение в зависимости от поставленной цели. В случае невыполнения поставленной цели, проводится анализ, рассчитываются отклонения и выявляются причины и проводится повторный опрос.



Рис.1. Алгоритм процесса проведения опроса слушателей

Для более качественного анализа необходимо обеспечить валидность и надежность, как самого измерительного инструмента анкеты, так и результатов опроса. Обеспечение валидности тестовых вопросов происходит до поступления в базу данных тестовых вопросов, до тестирования слушателей и после ответов слушателей (Рис.2.).

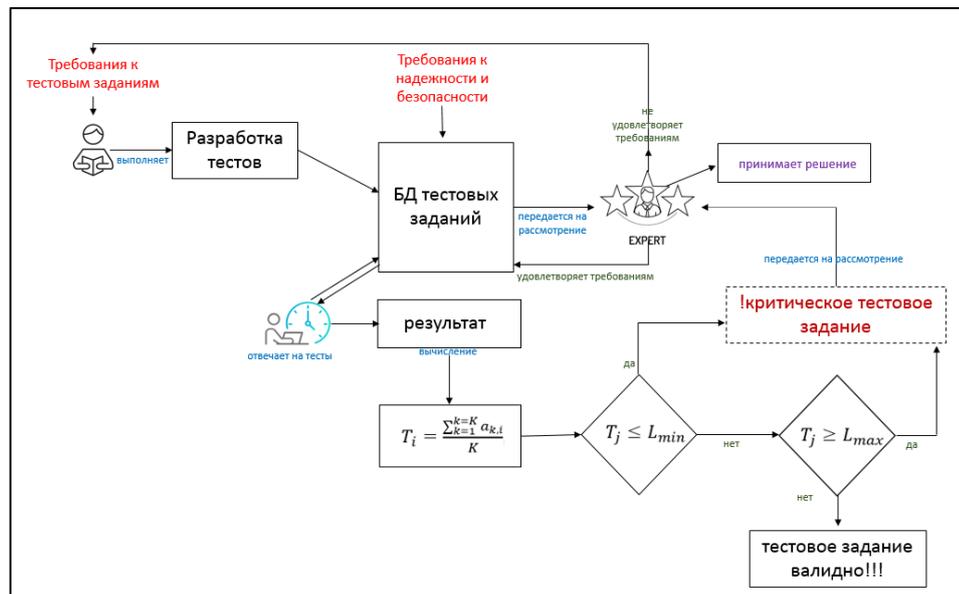


Рис. 2. Проверка валидности тестовых заданий

РЕЗУЛЬТАТЫ

Правильно методически подготовленные тестовые задания как по форме, так и по содержанию, использование выборки из большой базы тестовых вопросов, продуманная шкала оценивания теста и условия проведения тестирования, позволяют получить достаточно объективную картину результатов обучения слушателей по модулям. Существует большое количество работ, включающее требования к разработке тестовых заданий. Важно, чтобы эти требования соблюдались разработчиками тестовых заданий. Поэтому необходима проверка тестовых заданий независимыми экспертами до тестирования слушателей и после ответов слушателей на тестовые задания. Но для того, чтобы этот процесс не занимал много времени, можно автоматизировать процесс проверки тестовых заданий на валидность после ответов слушателей.

Пусть вектор $(a_{k,i})$ – задание на которое отвечает слушатель, где $k = \overline{1, M}$ и переменная i изменяется в пределах от 1 до N . Здесь изменяется в пределах от 1 до – количество слушателей, N – количество тестовых заданий. Тогда имеем множество векторов по каждому тестовому заданию:

$$\begin{pmatrix} a_{1,1}, a_{2,1}, \dots, a_{k,1} \\ a_{1,2}, a_{2,2}, \dots, a_{k,2} \\ \vdots \\ a_{1,i}, a_{2,i}, \dots, a_{k,i} \end{pmatrix} \quad (1)$$

Вычисляем среднее по каждому вектору $T_i = \frac{\sum_{k=1}^{k=K} a_{k,i}}{K}$, где K изменяется в пределах от 1 до M и i изменяется в пределах от 1 до N .

Вводим пороговое минимальное значение L_{min} и пороговое максимальное значение L_{max} .

Таким образом мы получаем множество дискретных величин в интервале от 0 до 100 ограниченных минимальным и максимальным пороговыми значениями (Рис.3).

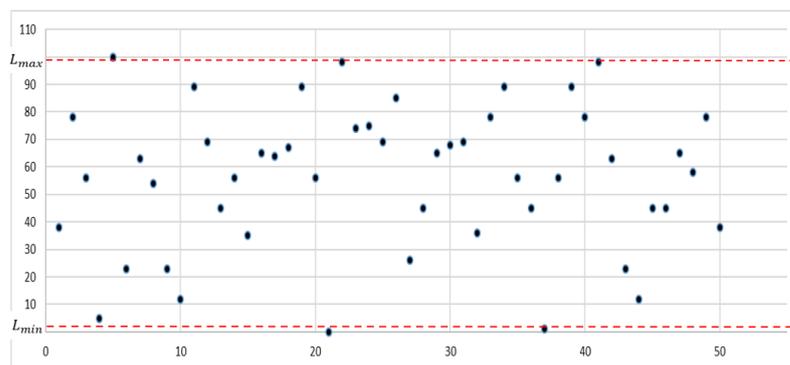


Рис.3. Множество допустимых тестовых заданий

Все значения, которые не входят в ограниченное множество допустимых значений, то есть удовлетворяют условиям $T_j \leq L_{min}$ и $T_j \geq L_{max}$ признаются критическими и передаются на рассмотрение независимому эксперту, который принимает решение об исключении тестового задания из базы тестов или его корректировке.

Учитывая специфику повышения квалификации, через определенный промежуток времени каждый раз имеется новая выборка для одного и того же инструмента. Поэтому для оценки надежности не могут использоваться методы, типа «тест-ретест».

Надежность анкеты определяется путем измерения распределения ответов слушателей по заданным вопросам. Для оценки согласованности ответов слушателей по каждому пункту анкеты при однократном тестировании достаточно использовать коэффициент α -Кронбаха [10]:

$$\alpha = \frac{N}{N-1} * \left(1 - \frac{\sum_{i=1}^N \sigma_i^2}{\sigma_t^2} \right) \quad (2)$$

где N – количество слушателей;

σ_i^2 - дисперсия откликов (оценок слушателей) по каждому пункту анкеты;

σ_t^2 - дисперсия суммарной шкалы (дисперсия суммы откликов каждого слушателя на пункты анкеты).

Расчет данного коэффициента очень просто автоматизировать, также возможно использовать специальные программы для вычисления дисперсий, такие как Matlab или MS Excel. В этом случае результаты вычисления дисперсий можно представить в виде таблицы 1.

Таблица 1

Результаты вычисления дисперсий

Респонденты: Пункты анкеты	1	2	...	k-1	k	
1	A1	A2	...	A(k-1)	Ak	$\sigma_1^2 (A1, A2, \dots, A(k-1), Ak)$
2	B1	B2	...	B(k-1)	Bk	$\sigma_2^2 (B1, B2, \dots, B(k-1), Bk)$
...
N-1	C1	C2	...	C(k-1)	Ck	$\sigma_{N-1}^2 (C1, C2, \dots, C(k-1), Ck)$
N	D1	D2	...	D(k-1)	Dk	$\sigma_N^2 (D1, D2, \dots, D(k-1), Dk)$
	\sum_1	\sum_2		\sum_{k-1}	\sum_k	$\sigma_t^2 \left(\sum_1, \sum_2, \dots, \sum_{k-1}, \sum_k \right)$

В данной таблице N – количество опрашиваемых слушателей, k – количество вопросов в анкете.

Если $\alpha \geq 0,71$, то пункты опрашиваемой анкеты признаются согласованными.

На надежность анкеты, определяемой по формуле Кронбаха влияет количество и качество пунктов, входящих в эту анкету. При исключении любого пункта из анкеты коэффициент α Кронбаха будет изменяться (уменьшаться или увеличиваться). При исключении пунктов, которые не противоречат другим пунктам анкеты, коэффициент α Кронбаха уменьшается. И, напротив, при исключении пунктов, которые не согласуются с другими пунктами анкеты, значение коэффициента α Кронбаха будет увеличиваться.

Поэтому при $\alpha < 0,71$ рекомендуется рассчитать этот показатель также при попеременном исключении пунктов из анкеты (Таблица 2).

Таблица 2

Расчет коэффициентов α Кронбаха при попеременном исключении пунктов анкеты

Исключаемый пункт анкеты	Значение коэффициента α Кронбаха без данного пункта анкеты
1	

2	
...	
M-1	
M	

Здесь M – число исключаемых пунктов анкеты.

ОБСУЖДЕНИЕ

Если исключение одного из пунктов анкеты приводит к повышению значения коэффициента α Кронбаха до требуемого уровня, значит именно данный пункт не согласуется с остальными и требует исключения из анкеты.

Одной из важных задач в проблеме повышения качества обучения слушателей повышения квалификации является оценка качества преподавания, так как в отличие от студента, слушатели, являясь коллегами преподавателей оценивают не только знания, но и степень мастерства и методику преподавания.

В Головном центре повышения квалификации руководящих кадров высшего образования ежемесячно проводят анкетирование слушателей повышения квалификации по окончании курса. Основные критерии, включают следующие пункты:

- 1) ясное и последовательное изложение материала;
- 2) умение заинтересовать слушателей и креативность;
- 3) умение наладить контакт с аудиторией;
- 4) обладание высокой культурой речи;
- 5) полезность материала для слушателей;
- 6) связь учебного материала с современным состоянием исследований в рассматриваемой области
- 7) разнообразие современных педагогических методов при проведении занятий.

Отдельно слушатели отмечают: умение снять напряжение и усталость аудитории и творческий подход, и интерес.

Уровень личностной и групповой удовлетворенности качеством преподавания определяется на основе суммарного итогового показателя по всем вопросам анкеты.

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

ВЫВОДЫ

Таким образом, можно сделать вывод, что разрешение проблемы валидности измерения качества повышения квалификации в большей степени зависит от методики проведения измерения качества и мотивационной работой, апробированной в реальных условиях.

REFERENCES

1. Дубина И.Н. Математические основы эмпирических социально-экономических исследований: учебное пособие. - Барнаул : Изд-во Алт. ун-та, 2006. 263 с.
2. Вакуленко А.А., Сытник Д.А. Методика экспертизы электронного анкетирования в системах контроля и оценки качества объектов. - Международный научный журнал. - Номер: 5 Год: 2013 Страницы: 45-49
3. Гуляева Т.И., Коломейченко А.С., Польшакова Н.В., Шуметов В.Г., Яковлев А.С. Опыт количественного анализа результатов анкетирования студентов по вопросам качества обучения: методика, модели, перспективы // ИТС. 2019. №2 (95). URL: <https://cyberleninka.ru/article/n/opyt-kolichestvennogo-analiza-rezultatov-anketirovaniya-studentov-po-voprosam-kachestva-obucheniya-metodika-modeli-perspektivy> (дата обращения: 16.08.2022).
4. Рахманова Ю. В. Методы социологического исследования: учебное пособие. — СПб.: Изд-во РГПУ им. А. И. Герцена, 2021. — 124 с.
5. Соколова Л.В. Тестирование как эффективный метод оценивания качества учебных достижений студентов / Л.В. Соколова, А.В. Молчанова // Развитие образования. – 2021. – Т. 4, №4. – С. 27-33. DOI:10.31483/r-100655
6. Кузнецов В.А. Пушкарев Г.С. Ярославская Е.И. Надежность и валидность русскоязычной версии многомерной шкалы восприятия социальной поддержки (MSPSS). Психологические исследования, 2015, 8(41), 10.
7. Reliability and Validity of a New Survey to Assess Global Health Competencies of Health Professionals// Mirella Veras, Kevin Pottie, Vivian Welch, Ron Labonte, Javier Eslava-Schmalbach, Cornelia M. Borkhoff, Elizabeth A. Kristjansson, Peter Tugwell Glob J Health Sci. 2013 Jan; 5(1): 13–27. Published online 2012 Oct 22. doi: 10.5539/gjhs.v5n1p13 PMID: PMC4776957
8. Scherpenzeel, Annette & Saris, Willem. (1997). The Validity and Reliability of Survey Questions: A Meta-Analysis of MTMM Studies. Sociological Methods & Research - SOCIOL METHOD RES. 25. 341-383. 10.1177/0049124197025003004.
9. SCHERPENZEEL AC, SARIS WE. The Validity and Reliability of Survey Questions: A Meta-Analysis of MTMM Studies. Sociological Methods & Research. 1997;25(3):341-383. doi:10.1177/0049124197025003004
10. https://ru.wikipedia.org/wiki/Альфа_Кронбаха (дата обращения:15.08.2022)