

УДК 37.013.77

СТАНДАРТЫ ПСИХОЛОГО-ПЕДАГОГИЧЕСКОГО ТЕСТИРОВАНИЯ

Хусейнова А.А., Ковалева С.В.

*Донской государственный технический университет, Ростов-на-Дону,
e-mail: slastena_96@bk.ru*

В статье рассматривается ряд малоизученных вопросов, касающихся стандартов психолого-педагогического тестирования. Целью исследования является обоснование необходимости создания, анализа и применения стандартов тестирования для организации и проведения успешного контрольно-оценочного мероприятия. Особое внимание уделено оценке существующих стандартов в области тестирования, в том числе зарубежных, и сравнению их с российскими стандартами тестирования в образовании и оценке персонала. В работе выделены основные структурные составляющие любого стандарта в области тестирования, которые принимаются для обеспечения точности, сопоставимости и интерпретации результатов, а также определена роль стандартов в связи с появлением новых сфер их применения. Обращается внимание на то, что в стандартах особое внимание уделяется аспектам справедливости, которые обычно являются обязанностью тех, кто разрабатывает, использует и интерпретирует результаты. Справедливость обозначена в четырех принципах: справедливость как отсутствие смещения; справедливость как равное отношение в процессе тестирования; справедливость как равенство в результатах тестирования; справедливость как возможность учиться. Проведенные исследования подтвердили, что использование стандартов тестирования в образовании и оценке персонала улучшает взаимную согласованность между разработчиками и пользователями тестовых методик, а также приводит к более четкому пониманию критериев объективной оценки.

Ключевые слова: стандарт, тест, психолого-педагогическое тестирование, контрольно-оценочная деятельность

STANDARDS OF PSYCHO-PEDAGOGICAL TESTING

Khuseynova A.A., Kovaleva S.V.

Don State Technical University, Rostov-on-Don, e-mail: slastena_96@bk.ru

The article deals with a number of poorly studied issues related to the standards of psychological and pedagogical testing. The purpose of the study is to substantiate the need for the creation, analysis and application of testing standards for the organization and conduct of successful control and evaluation activities. Special attention is paid to the evaluation of existing standards in the field of testing, including foreign ones and their comparison with the Russian standards of testing in education and personnel assessment. The paper highlights the main structural components of any standard in the field of testing, which are undertaken to ensure accuracy, comparability and interpretation of the results, as well as the role of standards in connection with the emergence of new areas of their application. Attention is drawn to the fact that the standards pay special attention to the aspects of equity, which are usually the responsibility of those who develop, use and interpret the results. Justice is defined in four principles: justice as the absence of bias; fairness as equitable treatment in the testing process; fairness as equality in outcomes of testing; and fairness as opportunity to learn. Studies have confirmed that the use of testing standards in education and personnel assessment improves mutual consistency between developers and users of test methods, as well as leads to a clearer understanding of the objective evaluation criteria.

Keywords: standard, test, psychological and pedagogical testing, control and evaluation activity

В настоящее время для оценки качества обучения в образовании и квалификации персонала в производственной сфере широко применяется тестирование, специальная форма стандартизированных, ограниченных по времени, кратких тестов, предназначенных для установления количественных и качественных индивидуальных различий в подготовке [1]. Для снижения рисков оценочного процесса и гарантии того, что в полной

мере используются все преимущества данного инструмента психолого-педагогического измерения, разрабатываются стандарты на разработку, организацию и проведение тестирования.

Цель исследования: обоснование необходимости создания и анализа стандартов тестирования для организации и проведения успешного контрольно-оценочного мероприятия в области образования и оценки персонала.

Материалы и методы исследования

Исследование базируется на трудах отечественных и зарубежных исследователей, а также на опубликованных стандартах тестирования. В процессе исследования были использованы следующие методы: теоретические – анализ литературных источников и педагогического опыта в аспекте изучаемой проблемы, синтез, абстрагирование; эмпирические – наблюдение, изучение документации, продуктов деятельности обучающихся и работников, педагогический эксперимент.

Результаты исследования и их обсуждение

Как отмечал исполнительный директор Американской ассоциации образовательных исследований – American Educational Research Association (AERA) – Фелис Дж. Левин: «Стандарты – это продукт исчерпывающего рассмотрения наилучшего руководства, возможного в образовательном и психологическом тестировании. Сотрудничество ассоциаций и ученых представляет собой привлечение лучшего опыта, а полученный продукт предлагает важные рекомендации для разработчиков тестов, политиков и преподавателей в области тестирования и измерения» [2].

В последнее время практика тестирования быстро развивалась, в дополнение к положительным явлениям наблюдался и рост недостатков – низкая квалификация значительной части пользователей, проблемы с лицензированной чистотой, а также разрозненные и неоднозначные рекомендации для практиков. Деятельность в области стандартизации тестирования как раз направлена на выполнение следующих функций: упорядочение объектов; установление в нормативных документах по стандартизации организационных, технических и других норм и требований; право на использование и соблюдение общих правил и требований, изложенных в нормах стандартизации.

Лаура Гамильтон, старший ученый Комитета управления по пересмотру стандартов психолого-педагогического тестирования, отмечала, что многие учебные заведения в штате Вашингтон

используют тесты для оценки прогресса учащихся. Применение стандартов в данной области обяжет учителей использовать только качественный и надежный материал, помогающий по настоящему оценивать успеваемость учащихся и исключит субъективный аспект оценивания даже при низких ставках. Комитет по стандартам до сих пор обсуждает оценку с использованием технических средств как эффективное оценочное средство и запрашивает информацию от внешних экспертов для развития.

Роль стандартов в связи с появлением новых сфер их применения все более возрастает. Стандартизация является ключевым фактором государственной политики, способствующей развитию образовательной и производственной сфер, навыков работы с информационными технологиями; потребности в непрерывном образовании; способностей к интегральному использованию знаний и др. Стандартизация тестирования в области образования и оценке персонала – это возможность повышения качества оценочного процесса (рис. 1).

В области тестирования уже разработано достаточно большое число стандартов и руководств на международном, национальном и региональном уровне.

К таковым относятся созданные: Международной тестовой комиссией – International Test Commission (ITC) руководства пользователей тестов; Британским психологическим обществом – British Psychological Society (BPS) системы квалификаций; Европейской федерацией психологов – European Federation of Psychologists Association (EFPA) тестовые стандарты. Существуют также стандарты и руководства, касающиеся разработки тестов и оценки их качества. К ним относятся голландский критерий рецензирования, норвежская система сертификация тестов и т.д. [3]. Российский стандарт тестирования персонала, созданный в 2014–2015 гг. по инициативе Национальной конфедерации «Развитие человеческого капитала» (НК РЧК), включает требования национальных и международных руководящих принципов по разработке и применению тестовых методик [4].



Рис. 1. Роль стандартизации тестирования в образовании и оценке персонала

Стандарт образовательного и психологического тестирования – Standards for educational and psychological testing (2014) – это так называемый золотой стандарт в руководстве по тестированию в Соединенных Штатах Америки, признанный во всем мире. Он написан для профессионалов в области психолого-педагогического тестирования и людей, интересующихся данной тематикой, и касается профессиональных и технических вопросов разработки и использования тестов в образовании, психологии и занятости. На сегодняшний день практически все государственные программы оценивания ссылаются на стандарты в разработке и использовании своих образовательных программ и практик оценивания. В этом стандарте особое внимание уделено следующим областям:

- изучению вопросов подотчетности, связанных с использованием тестов в образовательной сфере;
- разработке главы, связанной со справедливостью тестирования;
- расширению концепции доступности тестов для всех экзаменуемых;
- всестороннему представлению роли и важности тестов, необходимости обеспечения валидности и надежности;
- принятию во внимание растущей роли технологии тестирования в образовании и профессиональной сфере;
- совершенствованию структуры стандарта для лучшего информирования и др. [5].

Проанализируем основные структурные составляющие любого стандарта в области тестирования, которые предпринимаются для обеспечения точ-

ности и сопоставимости, а также обоснованной интерпретации результатов.

1. Валидность позволяет оценить пригодность методики для изучения конкретного явления. На это могут влиять четыре основных критерия (рис. 2).

Критерием валидности теста является независимый источник данных о том или ином явлении (психологическое свойство), изучение которого проводится посредством тестирования. Пока полученные результаты не проверены на соответствие критерию и цели тестирования, валидность не может быть оценена. Стандарты учитывают следующие составляющие валидности: содержание тестового задания; внутренняя структура теста; взаимоотношения между оцениваемыми факторами; корреляция результатов тестирования с данными, полученными от внешних экспертов, и т.д.

Валидность измерения необходима для поддержки интерпретации результатов тестов как предсказания будущего успеха испытуемого. Свидетельства валидности должны быть собраны до начала тестирования и проанализированы дополнительные данные, поскольку испытание продолжается при оперативном использовании. Более высокие уровни валидности необходимы, когда результаты теста имеют высокие ставки.

2. Надежность и ошибка измерения. Надежность в стандартах причисляется к постоянству предоставленных измерений, когда процедура тестирования воспроизводится несколько раз на одной и той же выборке испытуемых. При этом ни один испытуемый не может быть целиком последовательным, а ино-

гда из-за субъективности в ходе подсчитывания баллов оценка учащегося и средний балл группы постоянно отражают погрешность измерения. Сообщение о погрешности измерения необходимо для верной оценки и применения тестового инструмента. Главными свойствами надежности теста могут быть: воспроизводимость итогов, полученных в результате проведения исследования; уровень верности методики измерений; постоянство итогов в течение установленного периода времени.

3. Нормы и тестовые баллы в качестве количественного или качественного критерия для оценки результатов теста определяют степень достижения или выраженности определенного качества личности. Нормы необходимы при интерпретации результатов теста (первичных показателей) как эталона, с которым сравниваются результаты теста. Достоверность интерпретаций, связанных с нормой, частично зависит от контрольной группы, с которой сравниваются результаты теста. Важно, чтобы нормы основывались на разумной научной выборке достаточных размеров. Что касается сравнения результатов, в стандартах четко указано, что не должно быть чрезмерного искажения норм для разных выборок или взаимосвязи между ними. В тех случаях, когда результаты нормотворческого процесса значительны, и особенно когда речь идет о большом числе субъектов, эксперты должны быть обеспокоены тем, что процесс установления норм, в кото-

рых баллы определяют будущее субъекта, должен быть четко документирован и обоснован.

4. Вспомогательная документация для тестов. Стандарты содержат информацию о том, что дополнительные документы обязаны включать данные, предоставляющие возможность испытуемым и оценщикам установить полезность теста. Документация, как правило, определяет: вид проверки; ее предполагаемое применение; процедуру, взаимосвязанную с исследованием проверки; технические сведения, сопряженные с оценкой, интерпретацией и подтверждением валидности; выбор масштаба и стандартизацию, в случае если данное следует; управляющие основы с целью администрирования оценивающего события и др.

5. Справедливость в тестировании и использовании теста. Стандарты сосредоточены на аспектах справедливости, за которые обычно отвечают те, кто разрабатывает, использует и интерпретирует тесты. Они описывают справедливость в следующих четырех принципах, в которых должен использоваться данный термин (рис. 3).

Стандарты описывают понятие «смещение» как компонент, который приводит к систематически более низким или более высоким баллам для идентифицируемых групп испытуемых. Аналогичным образом определяются два основных источника смещения: запрещенные источники информации и ответы на задания.

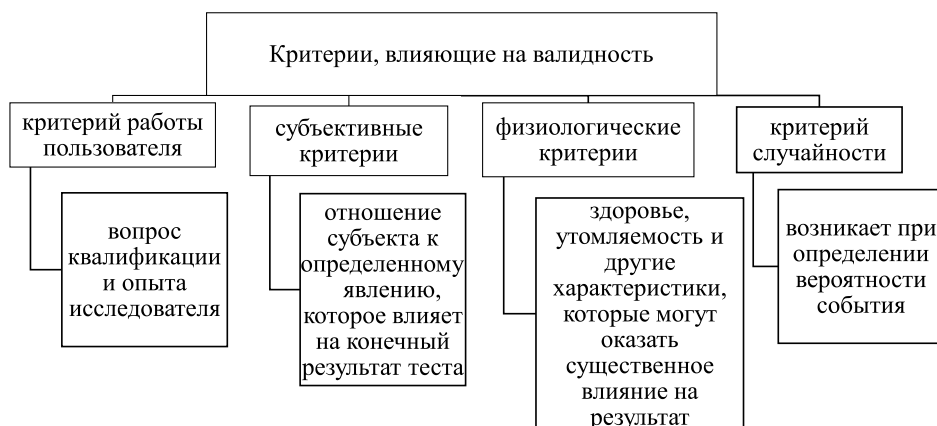


Рис. 2. Критерии, влияющие на валидность теста



Рис. 3. Использование термина «справедливость» в стандартах тестирования в образовании и оценке персонала

6. Обязанности пользователей тестовыми материалами. Тестовые пользователи называются группой специалистов, которые деятельно участвуют в толковании и применении итогов тестирования. При выборе теста и интерпретации предполагается, собственно, то, что испытуемый будет иметь точное представление о целях испытания и его вероятных последствиях. Если испытание проводится с целью, для которой имеется мало или вообще не имеется документации, пользователь несет ответственность за получение доказательств валидности и надежности испытания для этой цели. Иногда планируется утверждённая модификация формата теста, способа администрирования, руководства или языка, применяемого при проведении оценочного мероприятия, пользователь обязан владеть веским основанием для вывода о том, что достоверность, безопасность и своевременность общепризнанных мерок никак не будут установлены под угрозой. Обеспечение доступа результатов для всех испытуемых, то есть минимизация барьеров для доступа, достигается путем:

- использования процедур проектирования и разработки испытаний для снижения барьеров, создание «универсальных конструкций» оценивания;

- устранения проблем, которые могут ограничить способность испытуемого продемонстрировать то, что они знают и могут сделать из-за структурных проблем с тестом;

- необходимость обеспечения каждого обучающегося соответствующим набором средств для демонстрации знаний и компетенций [6].

В стандартах всех стран сегодня должны рассматриваться некоторые конкретные проблемы, связанные с безопасностью тестирования; автоматизированной проверкой и автоматическим результатом.

Говоря об отечественном опыте, отметим, что тестирование было широко принято в практике служебной аттестации российских компаний, однако долгое время не существовало единого стандартизированного документа, содержащего список требований и критериев оценки качества тестирования [4, 7]. Разрабатывая российский стандарт, принимались во внимание рекомендации национальных и международных организаций по развитию и применению тестов. Первый выпуск российского стандарта тестирования персонала – общепринятый практико-методический грант. Это закладывает основу для дальнейшего развития более подробных и узкоспециализированных стандартов и также формирует основание «профессионального стандарта эксперта в служебной аттестации» наряду со стандартами других методов и процедур оценки. Обширная публикация содержит общую терминологию, классификацию тестов и тестовых ситуаций, требования к валидности, описание принципов объективности, расчетные формулы, критерии оценки качества тестов и рекомендации по обработке результатов.

Стандарт содержит минимальный набор требований для всех этапов разработки и тестирования. Есть информация для HR-менеджеров, которые организуют крупные оценочные проекты, и подробные инструкции по конкрет-

ным аспектам тестирования (например, как организовать рабочее место – температура, уровень шума, освещение). Он также дает практические советы – как отличить «чистые» методы, как использовать IT-технологии, как объединить результаты тестирования с другими методами оценки для принятия окончательных кадровых решений и т.д. Стандарт базируется на одном из основных принципов распределения ответственности: разработчик несет ответственность за качество теста, а пользователь методики – за качество тестирования.

Измерительный механизм, сформированный согласно стандарту, отличается диагностичностью, самостоятельностью, четкостью, легкостью и воссоздаваемостью алгоритмизируемой технологии измерения степени достижений обучающихся либо оценки качества подготовленности персонала, установления его соответствия критериям. Опыт показывает, что стандартизированные педагогические измерения повышают мотивацию обучения и помогают достигать лучших результатов [8].

Заключение

В заключении отметим, что стандарты тестирования – это свод профессиональных, этических и культурных норм, которые используются на практике. Использование стандартов тестирования в образовании и оценке персонала улучшает взаимную согласованность между разработчиками и пользователями тестовых методик, а также приводит

к более четкому пониманию критериев объективной оценки.

Список литературы

1. Векслер В.А., Коноваленко О.Л. Психолого-педагогические аспекты тестирования // NovaInfo. 2015. № 35–1 [Электронный ресурс]. URL: novainfo.ru/article/3773 (дата обращения: 01.11.2018).
2. Батурин Н.А., Мельникова Н.Н. Технология разработки тестов: часть V // Вестник Южно-Уральского государственного университета. 2011. № 5 [Электронный ресурс]. URL: https://cyberleninka.ru/article/n/tehnologiya-razrabotki-testov-chast-v (дата обращения: 10.11.2018).
3. Бертрам Д. Стандарты тестов, квалификация и сертификация пользователей тестов // Вестник ЮУрГУ. 2011. № 5 [Электронный ресурс]. URL: http://dspace.susu.ru/xmlui/bitstream/handle/0001.74/2740/2.pdf?sequence=1&isAllowed=y (дата обращения: 03.12.2018).
4. Батурин Н.А., Вучетич Е.В., Костромина С.Н., Кукаркин Б.А., Куприянов Е.А., Лурье Е.В., Митина О.В., Науменко А.С., Орел Е.А., Полетаева Ю.С., Попов А.Ю., Потапкин А.А., Симоненко С.И., Синицына Ю.Д., Шмелев А.Г. Российский стандарт тестирования персонала // Организационная психология. 2015. Т. 5. № 2. С. 67–138.
5. The Standards for Educational and Psychological Testing / American Psychological Association [Электронный ресурс]. URL: https://www.apa.org/index.aspx (дата обращения: 11.10.2018).
6. Appendix, F. Summary of the standards for educational and psychological testing // California state personnel board merit selection manual: policy and practices. 2003 [Электронный ресурс]. URL: http://home.ubalt.edu/tmitch/645/articles (дата обращения: 12.10.2018).
7. Ефремова Н.Ф. Тестирование: теория, разработка и использование в практике учителя: методическое пособие. М.: Национальное образование, 2012. 223 с.
8. Ефремова Н.Ф. Учебные достижения как объект тестирования и показатель качества в образовании // Вопросы тестирования в образовании. 2004. № 9. С. 53–61.