

# СТАНДАРТЫ ТЕСТОВ, КВАЛИФИКАЦИЯ И СЕРТИФИКАЦИЯ ПОЛЬЗОВАТЕЛЕЙ ТЕСТОВ<sup>1</sup>

*Дэйв Бертрам*

Описаны проблемы обеспечения компетентности при использовании тестов и обеспечения эффективности и безопасности тестирования. Акцент делается на применении тестирования в сфере персонала. Представлен обзор изменений взаимоотношений между поставщиками и пользователями тестов вследствие развития технических разработок. Сделан обзор некоторых действующих стандартов и руководств по применению тестов, их качества и проведения оценки. Описана деятельность EFPA (Европейской Федерации психологических ассоциаций) совместно с другими международными организациями.

*Ключевые слова:* психодиагностика, психологическое тестирование, стандарты тестов, квалификация пользователей тестов.

## Кто является пользователями тестов?

Начнем с обсуждения отношений между поставщиками и пользователями тестов. Международные исследования показали, что большинство тестов в сфере персонала и образования используется не-психологами. В тех областях, где психологи используют тесты, психодиагносты довольно часто имеют небольшой практический опыт, не обладают необходимыми для сертификации компетенциями.

Необходимо гарантировать соответствующие уровни компетентности для всех, кто применяет тесты [14, 20]. Такие компетенции должны быть соотнесены с задачами оценки. Другими словами, необходимо определить, что человек должен знать и уметь и чего не должен, чтобы быть компетентным в особых ситуациях тестирования. Например, менеджеры по персоналу нуждаются в знаниях, умениях, навыках для применения тестов соответствующим образом для рекрутинга, отличных от знаний, умений, навыков медсестры психиатрии при оценке людей с умственными проблемами.

В 1998 году Международная тестовая комиссия провела опрос, включающий следующие вопросы: кто применяет тесты; доступность тестов различным категориям пользователей; качество стандартов тестов, кодексов, механизмов контроля, используемых при по-

ставке тестов; экспертиза и профессионализм в исследованиях и разработках тестов, адаптации тестов и психометрика внутри стран; мнения и убеждения об использовании тестов. Этот опрос явился результатом «корпоративных» ответов, полученных от респондентов национальных ассоциаций психологов. Так, данные представляют мнения и убеждения ключевых членов профессионального психологического сообщества о проблемах тестирования в своих странах. Данные были получены от 36 стран из 48, которых попросили ответить. Здесь вкратце будут представлены лишь некоторые результаты.

По ответам респондентов, опрошенных в 36 странах, число пользователей тестов составляет 3 234 994 человек. При рассмотрении результатов этого опроса необходимо помнить, что ответы, полученные от психологов, включают много оценок. Таким образом, результаты должны рассматриваться как ориентировочные, а не описательные. Согласно полученным результатам, 86,3 % пользователей тестов – не-психологи. Наибольшая группа – 78,8 % пользователей – в образовании. Наименьшая группа – 9 % всех пользователей специализируются в тестировании персонала. Клиника и право (12,2 %) – единственные области, где число не-психологов не превышает число психологов.

<sup>1</sup> Статья подготовлена на основе пленарного доклада на Всероссийской научно-практической конференции с международным участием по проблемам психодиагностики, тестам и тестированию, проведенной 8–10 сентября 2010 г. в Южно-Уральском государственном университете (г. Челябинск).

Перевод статьи Ю.Л. Юсуповой.

В исследовании также спрашивалось об обучении пользователей тестов. Респонденты указывали, что по их оценкам только 41 % пользователей получили специальное обучение. В образовании – наименьшее количество специально обученных пользователей, в то время как в клинике – самые высокие показатели. Психологи имеют тот же или более высокий уровень показателя обучения, чем не-психологи во всех областях. Тем не менее, различия очень малы в области тестирования персонала.

Ключевыми моментами этого исследования являются: роль психологов и не-психологов в тестировании; потребность в большем обучении для всех групп (психологи видят потребность быть более вовлеченными в использование тестов, но не соответствуют требованиям); различия между сферами использования (в здравоохранении, работе, образовании): особенно выражено различие между тестированием в сфере персонала и всеми другими сферами; предоставление тестов людям, не компетентным для их использования; различия в результатах между разными группами стран. Последние вопросы были подробно рассмотрены Бертрамом и Койн [4, 5], где можно найти подробные детали этого исследования.

Если вкратце, результаты этого опроса подтверждают точку зрения о том, что существует много различий в использовании тестов как между сферами применения внутри стран, так и между странами. Различия далее подтверждаются с помощью результатов последующих опросов EFPA относительно практики отдельных психологов, отношении к тестированию и использованию некоторых диагностических инструментов [21–23]. Такое различие в практике очень затрудняет создание единого решения проблемы гарантирования того, что все пользователи тестов обеспечивают компетентность и этичность использования инструментов оценки.

### **Влияние Интернета на применение тестов**

Интернет начал оказывать влияние на ежедневную жизнь, когда всемирная паутина стала развиваться в середине 1990-х гг. и когда повсеместно начали появляться компьютеры. Однако только к концу десятилетия Интернет начал оказывать значительное влияние на применение тестов. С 2000 г. влияние Интернета стало возрастать в очень быстром темпе, особенно в профессиональной сфере

применения тестов. В моей собственной компании, SHL, мы сейчас проводим около 90 % всех наших тестов через Интернет, а не в бланковом варианте. Когда я стал членом SHL в 1998 г., только малая часть наших тестов была в компьютерном варианте и ни одного теста не проводилось через Интернет.

Интернет оказывает ряд эффектов на паттерны использования тестов [10]. Во-первых, предоставляет доступ к тестам из любой точки мира, где есть доступ к Интернет-браузеру. При расширении это означает использование не только компьютера, но и других технологий (iPhone, Blackberry, iPad, digitalTV). В прошлом у издателей были большие товарные склады, заполненные бумажными материалами, которые высылались по почте пользователям тестов. Каждый раз материалы необходимо было пересматривать или исправлять, что было комплексной и дорогостоящей работой, особенно, если издатель имел много непроданных материалов – и гарантировать, что все пользователи тестов работают, ознакомившись с материалами. Через Интернет это можно сделать автоматически.

Однако более важно то, что Интернет изменил природу отношений между поставщиками и пользователями тестов. Сдвиг произошел в смещении локуса контроля от клиента к дистрибьютору тестов. Дистрибьюторам тестов больше не нужно отправлять пользователям тестов большие технические руководства, ключи, таблицы норм, тестовые буклеты, бланки ответов и другие подобные материалы через электронную почту. Сейчас вся эта нагрузка перешла на Интернет-сервер, управляемый дистрибьюторами тестов. Когда пользователь тестов хочет применить тест, он обычно выставляет порядок указанных адресов электронной почты или адресов людей, которые будут тестироваться. Затем тестируемый получает на электронную почту заготовленные для него инструкции, пароль, гиперссылки и другую соответствующую информацию, дающую возможность доступа и прохождения теста. После того, как это сделано, пользователь тестов получает результаты в форме отчета, описывающего результаты. Дистрибьютор тестов имеет доступ к данным всех тестов. Раньше эта информация хранилась у пользователей тестов и была труднодоступна для дистрибьютора. Сейчас это создает значительные возможности для разработчиков тестов и дистрибьюторам тестов в обновлении и повышении качества материалов.

Однако это возлагает больше ответственности на дистрибьюторов тестов, поэтому они становятся более включенными в процесс администрирования, подсчета баллов и интерпретации, каждый из которых раньше были зонами ответственности пользователя тестов.

Эти изменения оказали влияние на ролевые отношения между людьми и их зонами ответственности. Поскольку речь идет о стандартах, потребность в объективности, справедливости и этическом поведении остается неизменной. Изменилось лишь то, что во многих случаях традиционная ответственность в аспектах хорошего применения сдвигается от пользователей тестов к дистрибьюторам тестов.

Учитывая влияние Интернета на пользователей тестов, необходимо выделять различные способы оценки. Четыре таких способа описаны в «Руководстве по компьютерному и интернет-тестированию», разработанном ИТС (2006 г.) [18]:

- 1) открытый доступ к тестам (или отсутствие контроля разработчика);
- 2) контролируемый доступ (или отдаленная супервизия);
- 3) локальная супервизия;
- 4) полностью управляемое администрирование.

Основное влияние Интернет оказывает на развитие второго из этих четырех способов. В рамках этой статьи не может быть дано описание всех аспектов развития, которые имели отношение к компьютерному тестированию, но в следующем разделе будет описано то, что получилось.

Открытый способ – ненадежен. Он широко используется для незначимых оценок, когда нет необходимости контролировать тех, кто имеет доступ к тесту или действительно способен подтвердить подлинность пользователя тестов. Аналогично, если использовать открытый способ, невозможно гарантировать конфиденциальность содержания теста. Типичное использование открытого способа могло быть на сайте поиска работы, где люди использовали один или несколько тестов и получали результаты, информацию и рекомендации по выбору специальности.

Контролируемый способ имеет существенное отличие. Здесь дистрибьютор может контролировать один или большее количество переменных. Минимальный контроль осуществляется, когда используется уникальный логин-пароль и нормальный недоступный URL (т. е. пользователю нужно только найти

сайт, предоставляемый по электронной почте). К тому же есть возможность контролировать IP-адрес пользователя, для того чтобы знать, где используется тест. Также можно контролировать интервал времени, в течение которого пользователи имеют доступ к тесту, частоту использования теста, степень дистанционного контроля функционирования компьютера, и в некоторых обстоятельствах возможно добавление элементов дистанционной супервизии. Отдаленная супервизия включает использование аудио- и видеонаблюдения за испытуемыми, а также анализ их ответов здесь и сейчас. Последнее может быть использовано как часть данных судебного анализа для выявления отклоняющегося поведения (например, если кто-либо отвечает правильно очень быстро, его можно заподозрить в том, что у него есть правильные ответы; если кто-либо отвечает медленно, но случайно, они могут тратить время на «копирование» пунктов, чтобы продать их потом другим).

В отношении контролирующего способа при администрировании возникают две главные проблемы. Первая касается установления подлинности личности кандидата, вторая – безопасности содержания теста. С этими же проблемами мы сталкиваемся и при традиционном контролируемом тестировании, когда мы знаем об обмане и попытках украсть содержание [11]. Был разработан ряд методов специально для решения этих проблем. Например, при тестировании в сфере оценки персонала контроль осуществляется на ранних стадиях оценки для отбора (т. е. прежде чем будет составлен окончательный список кандидатов). Проверка подлинности личности кандидата, прошедшего тест на этой ранней стадии, является реальным выходом, если этот кандидат включен в окончательный список. Таким образом, методы были созданы для переоценки кандидатов из окончательного списка, чтобы гарантировать, что их результаты соответствуют тем, которые они получили до попадания в окончательный список. Более сложные методы, нацеленные на установление подлинности, вовлекают автоматизированное распознавание лиц кандидата, использование анализа образца нажатия клавиши или распознавание отпечатка пальца.

Проблема безопасности содержания теста касается, главным образом, тестов с максимальным выполнением (т. е. тех, в которых есть правильные ответы). Если вопросы таких тестов становятся широко известны, то ва-

лидность теста ставится под угрозу. Для Интернет-тестирования осуществлять контроль предполагается при помощи создания инструментов, основанных на теории IRT, с привлечением больших банков заданий. Поэтому обмануть такие тесты практически невозможно. SHL разработал набор тестов способностей, названных Верифай, на основе IRT для гарантии того, что каждый кандидат получает сопоставимый, но уникальный набор заданий. Кроме того, Верифай включает в себя второй короткий тест, который может применяться на кандидатах из окончательного списка, чтобы проверить, действительно ли кандидат сам получил баллы до попадания в окончательный список, а не кто-то другой.

Произошло очень быстрое техническое развитие, связанное с Интернет-тестированием. В области тестирования в целях оценки персонала и тестирования в образовании сейчас мы достаточно далеко продвинулись по сравнению с тем, что было пять лет назад. Вероятно, эти изменения продолжатся и применение технологий будет все больше и больше влиять на разработку, продвижение и использование тестов во всех областях. Стандарты тестов и уровни квалификации пользователей должны следовать этим изменениям. В следующем разделе мы рассмотрим несколько главных линий развития стандартов тестов и увидим, как они меняются, чтобы идти в ногу с техническими изменениями.

## Стандарты и руководства

Почему нужны стандарты? Стандарты выполняют две функции. Во-первых, они снижают риск нанесения вреда испытуемому тестами, а во-вторых, они гарантируют, что все преимущества тестов используются. Что касается риска навредить, то риски неодинаковы, они зависят от цели тестирования. Снижение таких рисков обеспечивается компетентностью пользователей в той области тестирования, в которой они практикуют. Из-за того, что риски неодинаковы, невозможно создать отдельный специальный уровень компетентности, который относился бы ко всем во всех ситуациях. Поэтому необходимы системы, разрабатывающие уровни компетентности пользователей и области применения. Для гарантии того, что преимущества тестов используются, необходимо удостовериться в их соответствии требованиям. Как было отмечено ранее, это привело к открытию доступа к тестам профессионалам не-психологам. Особенно это касается тестирования в сфере пер-

сонала и тестирования в образовании. Открытый доступ к тестам возможен только, если все пользователи компетентны оптимально использовать информацию, получаемую при помощи тестов. В целом, стандарты помогают сбалансировать риски и преимущества.

Стандарты объединяют людей, продукты и процессы. В области тестирования для каждого из них было разработано большое количество стандартов и руководств. Например, ИТС разработала руководства пользователей тестов [17–19], BPS разработала систему квалификации пользователей тестов, EFPA разработала стандарты пользователей тестов. ИТС разработала руководства по адаптации тестов и руководства для компьютерного и Интернет-тестирования. Есть также стандарты и руководства, касающиеся тестов и их качества [1, 2]. К ним относятся Голландский критерий рецензирования COTAN, процедуры рецензирования Института Бьюроса, процедуры рецензирования и регистрации BPS, норвежская сертификация тестов DNV, процесс рецензирования в Германии, основанный на DIN33430. Многие из них касаются также практики добросовестного тестирования и оценки. Как позже будет отмечено, сейчас разрабатывается новый стандарт ISO (ISO 10667) специально для урегулирования процесса оценки в сфере персонала. Еще одна работа, касающаяся процесса оценки, – это Руководства для процесса оценки, разработанные Fernandez-Ballesteros с соавторами [12, 13].

## Руководства ИТС

Международная комиссия по тестам (ИТС: [www.intestcom.org](http://www.intestcom.org)) [19] – это «Ассоциация национальных психологических ассоциаций, комиссий по тестам, издателей и других организаций, призванных продвигать эффективное тестирование и принципы оценки, а также для надлежащего развития, оценки и использования образовательных и психологических инструментов». В ИТС входит 20 полноправных членов (национальные профессиональные психологические ассоциации), 46 аффилированных членов (различные комиссии по тестам, издатели и исследовательские организации, вовлеченные в тестирование) и более 300 индивидуальных представителей (лица, работающие с тестами или интересующиеся тестированием). Членами ИТС являются большинство стран Западной и Восточной Европы, в том числе и Россия, страны Северной Америки и некоторые страны Среднего и Дальнего Востока, Южной Америки и Афри-

ки. ИТС зарегистрирована как некоммерческая организация и связана с IAAP и IUPsyS.

Международная комиссия по тестам (ИТС) обеспечивает обмен информацией между участниками и стимулирует сотрудничество по проблемам создания, распространения и применения психологических тестов и других психодиагностических инструментов. Чтобы достигнуть этих целей ИТС организует международные встречи, продвигает публикацию релевантной информации посредством собственных и других научных выходов и стимулирует международное сотрудничество в научно-исследовательских работах о научном и этичном применении тестов.

Все руководства ИТС разработаны международными группами экспертов и открыты международной критике и пересмотру через циклы консультаций. К настоящему моменту ИТС опубликовала четыре руководства:

1. Руководства для перевода и адаптации тестов [15, 16, 24]. Это первое руководство, разработанное ИТС, широко принятое в качестве стандартов по адаптации тестов. Разрабатывается новая версия стандартов, которая будет опубликована в 2011 году.

2. Международные руководства по применению тестов [3, 10]. Они касаются добросовестной практики и этичного использования тестов. Руководства переведены на 14 языков и используются в качестве основы для Европейских стандартов пользователей тестов и других национальных стандартов.

3. Руководства по компьютеризированному и Интернет-тестированию [18]. Они касаются аспектов технологий, которые затрагивают тесты и тестирование и полагаются на руководства по применению тестов. Это набор принципов, касающихся защищенности теста, конфиденциальности данных и добросовестности разработчика, издателя и пользователя теста.

4. Справочник тестируемому по технологии компьютерного тестирования [19]. Это новое руководство, разработанное на основе руководства по компьютерному тестированию, но адресованное тестируемому. Справочник довольно небольшой и простой, он представляет собой список того, что тестируемые могут ожидать в процессе тестирования и перечень их ответственности.

**Европейская федерация психологических ассоциаций (EFPA)**

В состав EFPA входит 35 Европейских психологических ассоциаций, представляю-

щих интересы около 250 000 психологов. Новым членом Федерации, вступившим в ее состав в 2009 году, является Россия. Главный офис EFPA находится в Брюсселе, EFPA формально признан Европейским союзом как негосударственная организация.

Постоянный Комитет EFPA по тестам и тестированию (SCTT) отвечает за руководство Исполнительным органом EFPA по вопросам касательно тестов и тестирования. В него входят национальные представители, и в настоящее время это один из самых больших и наиболее активных комитетов EFPA. Комитет издал два главных набора стандартов (размещенных на сайте EFPA):

- Модель рецензирования EFPA по описанию и оценке психологических тестов. Версия 3.41, 2005. (На сайте есть русский перевод, выполненный под руководством Н.А. Батурина).

- Европейские стандарты пользователей EFPA по применению тестов в сфере оценки персонала. Версия 1.91, Сент. 2005.

SCTT работал совместно с Европейской ассоциацией труда и Организационными психологами (EAWOP) для разработки квалификации пользователей тестов, основанных на Стандартах пользователей тестов. Они совместно разработали и профессиональные стандарты по оценке персонала, основанные на стандартах EFPA, которые больше ориентированы на оценку, а не на тестирование. Эта работа оказала большое влияние на содержание проекта ISO 10667, о котором будет сказано позже.

SCTT в настоящее время разрабатывает процедуры аккредитации национальных квалификаций пользователей тестов на основе оценки того:

- насколько их содержание соответствует Стандартам Пользователей тестов EFPA;
- насколько методы, используемые для оценки пользователей, соответствуют главным принципам оценки компетентности EFPA;
- насколько обеспечено качество процедуры присвоения квалификаций.

Стандарты пользователей разработаны на основе Стандартов квалификации пользователей тестов и имеют похожую структуру. Они делятся на три раздела, первый из которых касается вопросов профессионализма и этики, а два других – процессуальных вопросов.

Раздел 1. Принятие ответственности за этичное использование теста

1) профессиональная и этичная манера поведения;

2) гарантия компетентного использования тестов;

3) принятие ответственности за использование тестов;

4) гарантия того, что материалы теста надежно хранятся;

5) гарантия конфиденциального хранения результатов.

Раздел 2. Следование принципам добросовестной практики в использовании тестов

1) оценка потенциальной полезности тестирования в ситуации оценивания;

2) выбор отвечающих требованиям тестов, соответствующих ситуации;

3) должное внимание проблемам честности в тестировании;

4) соответствующие анализ и интерпретация результатов;

5) ясное и точное сообщение результатов;

6) рассмотрение уместности тестов и их использования.

Раздел 3. Следование принципам добросовестной практики в администрировании тестирования

1) подготовка к тестированию;

2) использование тестов должным образом;

3) точный подсчет результатов.

Модель квалификации EFPA различает три уровня компетентности:

Уровень 1 может практиковать под супервизией. Те, кто администрируют тесты от имени квалифицированного пользователя тестов, могут работать на этом уровне. Уровень 1 касается администрирования тестов и любой другой квалификации, связанной с каким-либо аспектом использования теста. К основным положениям квалификации уровня 1, например, относится применение линейными руководителями отчетов по тестам.

Уровень 2 может практиковать независимо, но в ограниченном круге ситуаций. Например, специалист по персоналу, который использует определенные типы тестов для оценки, отбора или развития персонала, должен соответствовать требованиям уровня 2.

Уровень 3 может практиковать независимо в широком круге ситуаций с экспертной оценкой в одной или более областях. Психологи, специализирующиеся в тестировании, должны соответствовать требованиям уровня 3. Некоторые другие профессионалы-эксперты в тестировании также могут работать на этом уровне.

(Недавнее исследование содержания этих квалификаций выявило, что они соответству-

ют уровням 4, 5 и 7 Европейской системы квалификации.)

Как отмечено выше, для уровня 2 EFPA рассматривает процедуры аккредитации схем национальной сертификации. Однако EFPA рассматривает необходимость развития уровня 3 квалификации для психологов, чтобы он соответствовал требованиям специального реестра EuroPsy.

В то время как третий уровень квалификации по классификации EFPA доступен только психологам, первый и второй уровни квалификации доступны широкому спектру специалистов, который также зависит от особенностей законов и практики в каждой отдельной стране. Важно, что EFPA устанавливает единые для всех стандарты использования тестов: психологов и не психологов. Таким образом, осуществляется влияние на большую часть европейского сообщества, которое имеет отношение к тестам и тестированию, и контроль над всеми сферами, в которых тестирование может нанести какой-либо ущерб, а не только над сферой профессиональной психологической практики.

## **Британское Психологическое Сообщество (BPS)**

BPS занимается сертифицированием пользователей тестовых методик с 1991 года. Во второй половине 1980-х BPS разработала компетентностный подход к сертификации [6, 7]. Изначально в британской системе различались квалификационный уровень А (который включал использование групповых тестов способностей и подобных методик) и квалификационный уровень В (который включал использование личностных опросников и остальных методов, основанных на самоотчете). Уровень В был разделен на две квалификации: промежуточную и полную. В начале все эти квалификации были разработаны для отбора персонала и психодиагностики в организациях. Уровень А и средний уровень В соотносятся с уровнем 2 по классификации EFPA, а полный уровень В с уровнем 3.

Несколько позднее в BPS ввели отдельную квалификацию в администрировании тестов (такую же, как в уровне 1 EFPA) и квалификацию для использования тестов в образовательной сфере: Сертификат компетентности в тестировании в образовании (ССТЕТ). Он основывался на уровне 2 EFPA.

Система поддержания стандартов и гарантий качества по отношению к квалификациям BPS основывается на процессе ратифи-

кации. Сертифицированный психолог, работающий на Сообщество, контролирует качества оценки и компетентность пользователей тестов. Те, кто оценивают квалификацию пользователей по системе Сообщества, должны быть сертифицированными психологами: они должны получить соответствующую квалификацию пользования тестами и должны быть признаны Сообществом компетентными в оценке других пользователей.

К концу 2009 года, Сообществом было выдано следующее число сертификатов:

26 798 – уровень А в сфере персонала (1991 – н. в.);

9599 – промежуточный уровень В в сфере персонала (1995 – н. в.);

480 – полный уровень В в сфере персонала (1995 – н. в.);

726 – администрирование тестов в сфере персонала (2003 – н. в.);

1564 – образовательный ССЕТ уровень А (2004 – н. в.).

У Сообщества есть Реестр компетенций в использовании тестов. Любой, кто имеет сертификат в тестировании, может быть записан в данный Реестр. В декабре 2009 в Реестре было записано 9150 человек. Большая часть сертифицированных пользователей тестов (около 90 %) не являются психологами. Единственным исключением является полный уровень В, где большинство психологи.

Уровни А и В квалификации пользователей тестов стали серьезным успехом в установлении независимых критериев для стандартов использования тестов. Они также способствовали тому, что BPS вне психологического сообщества стала расцениваться как организация, ответственная за определение и поддержание стандартов в данной области. Более детальное описание работы BPS в отношении тестов и тестирования можно найти на сайте Психологического тестового центра Сообщества: [www.psychtesting.org.uk](http://www.psychtesting.org.uk)

BPS в настоящее время адаптировала стандарты использования тестов EFPA и находится в процессе глобальной ревизии своей системы квалификации пользователей тестов. Мы сейчас осуществляем процесс приведения наших квалификационных процедур в такую форму, которая соответствовала бы модели, принятой EFPA. С начала 2011 года система BPS будет приведена к трехуровневой модели EFPA. Это будет заключаться не только в смене названий квалификаций, но в смещении содержания квалификационных категорий.

Также будет изменен статус Реестра. Сообщество больше не будет выдавать сертификаты компетентности, а будет оценивать людей, определяя их право войти в Реестр. Вхождение в Реестр станет обязательным. Те, кто уже входит в Реестр, должны будут подтвердить свою компетентность, а в последующем все, входящие в Реестр, должны будут подтверждать свою квалификацию каждые 6 лет.

Кроме работы по определению квалификации пользователей тестов, BPS разработало процедуры рецензирования самих тестов, основанные на модели Bartram, Lindley&Foster (1990) [3]. За последние несколько лет мы перешли в рецензировании тестов к критериям EFPA, а также полностью перевели систему рецензирования в онлайн-формат. К тому же процедуры регистрации тестов были разработаны на основании стандартов EFPA по минимальным требованиям к тестовой продукции.

#### **Новые разработки международных стандартов**

ISO 10667 – это первый ISO стандарт, который посвящен различным аспектам качества психологической оценки. Это стандарт качества в области предоставления услуг психологической оценке. Как таковой этот стандарт описывает требования к процессу, а не к человеку или продукту. В нем проработаны все области оценки в рабочем и организационном контексте.

Этот стандарт разделяется на две части: одна касается обязанностей клиентов, а другая – поставщиков услуг. ISO 10667 обеспечивает основу, на которой могут быть выстроены более конкретные государственные и международные стандарты и квалификации, такие как описанные выше. Начальной точкой проекта ISO был DIN 33430. Этот стандарт был разработан в Германии и использовался там как основа для квалификации тестов и их пользователей. DIN (Германский институт стандартов) внес предложение о разработке стандарта ISO, основанного на этой системе. Страны-члены ISO проголосовали в мае 2006 г. за это предложение, и Проектный комитет ISO (PC230) впервые собрался в Берлине в марте 2007. Проектный комитет избрал небольшую группу по разработке чернового варианта, и в последующие три года проектный комитет и группа попеременно разрабатывали стандарт. Между встречами проектного комитета происходили встречи различных групп технической поддержки или «теневых коми-



тетов». В процессе разработки стандарта было направлено много усилий на то, чтобы объединить работу ИТС, EFPA, и EAWOP, а также учесть национальные стандарты и руководства. В ходе работы проектного комитета был разработан и далее принят на обсуждение черновой вариант международных стандартов. Результаты обсуждения в настоящий момент анализируются, и на основе данного анализа будет сформирована программа работы окончательной встречи проектного комитета в декабре 2010 г. Планируется опубликовать финальный вариант стандартов в начале 2011 года.

Стандарты касаются проведения оценки на индивидуальном, групповом и организационном уровне. Их целью является продвижение добросовестной практики и содействие ясности в отношениях клиента и поставщика услуг. Они призваны быть практическим руководством для клиентов и поставщиков услуг, в них описаны их роли и зоны ответственности перед, во время и после процесса оценки. В них также описываются права и обязанности участников оценки и других людей, вовлеченных в процесс оценки.

Точнее, стандарты относятся к инструментам и процедурам оценки людей:

- на индивидуальном уровне (например, интервью, профориентация, отбор персонала, планирование карьеры);
- групповом уровне (например, психологический климат, оценка командной работы);
- организационном уровне (например, удовлетворенность работников, корпоративная культура, вовлеченность).

Стандарты содержат принципы для оценки на протяжении всего карьерного пути: рекрутирование и отбор, карьерная ориентация и сопровождение, персональное развитие, планирование карьеры, трудоустройство уволенных и т. д.

Стандарты представлены в двух документах: в одном устанавливаются требования к поставщикам услуг, в другом – к клиентам. Каждый документ делится на 6 секций, или «статей»:

1. Сфера рассмотрения.
2. Термины и определения.
3. Принципы.
4. Процедуры перед оценкой.
5. Проведение оценки.
6. Обзор после оценки.

Основные положения стандартов сосредоточены в статьях 4, 5 и 6:

#### 4. Процедуры перед оценкой:

- определение задач и целей оценки;
- заключение договора на оценку.

#### 5. Проведение оценки:

- планирование оценки;
- информирование участников оценки;
- проведение оценки;
- интерпретация результатов;
- подготовка и предоставление отчетов;
- формулирование и предоставление обратной связи;
- определение качества оценочного процесса.

#### 6. Обзор после оценки.

Важно отметить, что ISO стандарты доступны для всех. В отличие от предшествующих стандартов, которые были разработаны для внутреннего пользования психологами, данные стандарты разработаны профессионалами из различных сфер деятельности, которые соприкасаются с тестированием и оценкой. Более того, любая организация может создать и использовать процедуры для сертификации по стандартам ISO. Их «статус» как сертифицирующей организации будет зависеть от того, кем они являются и имеют ли они независимую аккредитацию (например, в Великобритании существует национальная аккредитационная служба, которая предоставляет подтверждение того, что та или иная организация может предоставлять сертификационные услуги по ISO стандартам).

Тем не менее, психологи играли важную роль в разработке ISO 10667. Все члены комитета по разработке чернового варианта являются психологами с широким спектром опыта в оценке. Национальные психологические ассоциации должны принять во внимание их потенциальную роль в качестве органов сертификации для этого стандарта.

Скорее всего, стандарт ISO во многом определит содержание будущих стандартов, касающихся психологической оценки.

#### Выводы

Этичность оценки зависит от компетентности оценщика. Профессиональные регалии не гарантируют компетентность, и следовательно, мы должны соотносить действия, производимые в процессе оценки, с компетентностью, необходимой действующему лицу для их выполнения. Компетентность пользователя наилучшим образом обеспечивается системой квалификации, основанной на стандартах, в которой наибольшее внимание уделяется знаниям и практическим навыкам, не-



обходимым для использования данных тестов в данном наборе ситуаций. В Великобритании подход BPS был широко принят другими профессионалами, в связи с необходимостью в компетентности и осознанием наличия границ компетентности. Также этот подход привел к признанию BPS в качестве организации, устанавливающей стандарты для компетенций.

В рамках данной статьи нет возможности описать другие инициативы (см. ранний обзор [4, 5] и более свежий [21]), но необходимо упомянуть о сотрудничестве между Великобританией, Норвегией и Швецией как в сфере квалификации пользователей тестов, так и в рецензировании тестов. Эти три страны осуществляют очень схожий подход к данной проблематике, в котором небольшие различия между ними вызваны различной историей и практикой в отношении тестирования. Германский DIN 33430 являлся главным шагом в установлении обеспечения гарантий качества в оценке на предприятиях и основой для создания стандарта ISO 10667. Скорее всего, подобные разработки обретут еще большее значение в будущем.

Основным выводом является то, что психологическая профессия должна перестать «смотреть внутрь» и начать «смотреть вовне», с точки зрения своей роли в широком сообществе. Психология как профессия имеет потенциал для установления плана работы и развития в сфере тестов и тестирования. Если этого не произойдет или этот план будет касаться только того, что происходит в сфере тестирования внутри профессионального сообщества, другие возьмут на себя лидирующую роль. Это не удовлетворяет ни чьим интересам.

Международные стандарты определяют добросовестность практики, но не могут быть принудительными из-за широкого спектра различных национальных профессиональных политик, политических аспектов, правил и регулятивных аппаратов. Стандарты также значительно варьируются по функциям и областям применения:

- качество используемых инструментов и процедур – стандарты продукта;
- компетентность пользователей – кадровые стандарты;
- качество процесса, в рамках которого действуют пользователи – стандарты предоставления услуг.

Все три типа стандартов важны, но, возможно, основанная на компетентностях ква-

лификация пользователя теста является наиболее важной, поскольку обеспечивает практичный и эффективный способ внедрения продуктов и предоставления услуг, а также обеспечения добросовестной практики.

### Литература<sup>2</sup>

1. AERA (1999). *American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. Standards for Educational and Psychological Testing. Washington DC: American Educational Research Association.*
2. APA (1985). *American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1985). Standards for Educational and Psychological Testing. Washington DC: American Psychological Association.*
3. Bartram, D., Lindley, P.A. and Foster, J.M. (1990). *A Review of Psychometric Tests for Assessment in Vocational Training. The Training Agency: Sheffield.*
4. Bartram, D., & Coyne, I. (1998a). *The ITC/EFPPA survey of testing and test use within Europe. In Proceedings of the British Psychological Society's Occupational Psychology Conference (pp. 197–201). Leicester, UK: British Psychological Society.*
5. Bartram, D., & Coyne, I. (1998b). *Variations in national patterns of testing and test use. European Journal of Psychological Assessment, 14, 249–260.*
6. Bartram, D. (1995). *The development of standards for the use of psychological tests in occupational settings: The competence approach. The Psychologist, May, 219–223.*
7. Bartram, D. (1996). *Test qualifications and test use in the UK: The competence approach. European Journal of Psychological Assessment, 12, 62–71.*
8. Bartram, D. (1998). *The need for international guidelines on standards for test use: A review of European and International initiatives. European Psychologist, 3, 155–163.*
9. Bartram, D. (2001). *The development of international guidelines on test use: the International Test Commission Project. International Journal of Testing, 1, 33–53.*
10. Bartram, D. (2008). *The advantages and disadvantages of on-line testing. In Cartwright, S,*

<sup>2</sup> Перечень процитированных в статье литературных источников представлен с сохранением оформления автора статьи.

& Cooper C. (Eds), *The Oxford Handbook of Personnel Psychology*. Oxford University Press.

11. Cizek, G.J. (1999). *Cheating on tests: How to do it, detect it, and prevent it*. Mahwah, NJ: Lawrence Erlbaum Associates.

12. Fernandez-Ballesteros, R. (1997). Task force for the development of guidelines for the assessment process (GAP). *The International Test Commission Newsletter*, 7, 16–20.

13. Fernandez-Ballesteros, R., De Bruyn, E.E.J., Godoy, A., Hornke, L.F., Ter Laak, J. Vizcarro, C., Westhoff, K., Westmeyer, H., & Zaccagnini, J.L. (2001). *Guidelines for the assessment process (GAP)*. *European Journal of Psychological Assessment*, 17, 187–200.

14. DeMers, S.Y., Turner, S.M. (Cochairs), Andberg, M. Foote, W. Hough, L. Ivnik, R. Meier, S. Moreland, K. & Rey-Casserly, C.M. (2000). *Report of the Task Force on Test User Qualifications*. Washington, D.C.: Practice and Science Directorates, American Psychological Association.

15. Hambleton, R. (1994). *Guidelines for adapting educational and psychological tests: A progress report*. *European Journal of Psychological Assessment*, 10, 229–244.

16. Hambleton, R., (2005). *Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures*. In, R.K. Hambleton, P. Merenda, & C. Spielberger (Eds.).

*Adapting Educational and Psychological Tests for Cross-Cultural Assessment*. Lawrence Erlbaum.

17. ITC (2001). *International Guidelines on Test Use*. *International Journal of Testing*, 1, 95–114.

18. ITC (2006). *International Guidelines on computer-based and Internet-delivered testing*. *International Journal of Testing*, 6, 143–172.

19. ITC (2010). *A Test-Taker's Guide to Technology-Based Testing*. [<http://www.intestcom.org>]

20. Joint Committee on Testing Practices. (2000). *Rights and Responsibilities of Test Takers: Guidelines and Expectations*. Washington DC: Joint Committee on Testing Practices.

21. Muniz, J., & Bartram, D. (2007). *Improving International Tests and Testing*. *European Psychologist*, 12(3), 206–219.

22. Muñiz, J., Prieto, G., Almeida, L., & Bartram, D. (1999). *Test use in Spain, Portugal and Latin American countries*. *European Journal of Psychological Assessment*, 15, 151–157.

23. Muñiz, J., Bartram, D., Evers, A., Boben, D., Matesic, K., Glabeke, K., Fernández-Hermida, J.R., & Zaal, J.N. (2001). *Testing Practices in European Countries*. *European Journal of Psychological Assessment*, 17, 201–211.

24. Van de Vijver, F. & Hambleton, R. (1996). *Translating tests: some practical guidelines*. *European Psychologist*, 1, 89–99.

*Поступила в редакцию 14 января 2011 г.*

Дэйв Бертрам. Профессор, директор по научно-исследовательским работам, SHL Group Ltd (Великобритания).

Dave Bartram. Prof, Research director, SHL Group Ltd.