



## Российский стандарт тестирования персонала (временная версия, созданная для широкого обсуждения в 2015 году)

**БАТУРИН Николай Алексеевич** (д. психол. н., проф. ЮУрГУ, Пситест, Челябинск, Россия)

**ВУЧЕТИЧ Евгений Викторович** (EHECT Partners Group, Москва, Россия)

**КОСТРОМИНА Светлана Николаевна** (д. психол. н., проф. СПбГУ, Санкт-Петербург, Россия)

**КУКАРКИН Борис Александрович** (EHECT Partners Group, Москва, Россия)

**КУПРИЯНОВ Евгений Андреевич** (к. психол. н., НИУ ВШЭ, Москва, Россия)

**ЛУРЬЕ Евгений Валентинович** (ЭКОПСИ Консалтинг, Москва, Россия)

**МИТИНА Ольга Валентиновна** (к. психол. н., вед. науч. сотр., МГУ, Москва, Россия)

**НАУМЕНКО Анна Сергеевна** (к. психол. н., независимый консультант, Лондон, Великобритания)

**ОРЕЛ Екатерина Алексеевна** (к. психол. н., доцент НИУ ВШЭ, Москва, Россия)

**ПОЛЕТАЕВА Юлия Сергеевна** (SHL, Москва, Россия)

**ПОПОВ Алексей Юрьевич** (к. психол. н., ЭКОПСИ Консалтинг, Москва, Россия)

**ПОТАПКИН Алексей Анатольевич** (к. психол. н., независимый консультант, Москва, Россия)

**СИМОНЕНКО Светлана Игоревна** (к. психол. н., Detech, Москва, Россия)

**СИНИЦЫНА Юлия Дмитриевна** (TalentQ, Москва, Россия)

**ШМЕЛЕВ Александр Георгиевич** (д. психол. н., проф. МГУ, Группа компаний «Гуманитарные технологии», Москва, Россия)

В последние десятилетия использование тестовых методик получило широкое распространение в практике оценки персонала российских компаний, однако до настоящего момента не существовало единого и общепринятого документа, содержащего перечень требований и критериев, позволяющих оценить качество тестирования. Российский стандарт тестирования персонала был создан в 2014—2015 гг. по инициативе Национальной конфедерации «Развитие человеческого капитала» (НК РЧК) и поддержан экспертным сообществом. Данный стандарт призван разъяснить и популяризировать в русскоязычном пространстве один из базовых принципов распределения ответственности: за качество теста отвечает разработчик, а за качество тестиро-

**Благодарности.** Участвовали в отдельных заседаниях, дистанционных онлайн-опросах, предоставили комментарии и рецензии следующие эксперты: Маргарита Акимова, Тахир Базаров, Надежда Волкова, Александр Ерофеев, Владимир Жильцов, Александр Комков, Виктория Кононова, Олеся Кувшинова, Ирина Лисица, Юрий Михеев, Алексей Серебряков, Юрий Синягин, Юрий Тукачев, Людмила Собчик, Татьяна Хватинина, Александр Чумаков.

Рабочая группа и НК РЧК выражают благодарность профессору А.Г. Шмелеву за научно-методическое и организационное обеспечение разработки стандарта, Алексею Попову, Анне Науменко и Дарье Портновой за редактирование и оформление текста, а также Михаилу Морозову, Ивану Осадчему и Надежде Волковой за техническую помощь, помощь в протоколировании собраний, а также в организации сбора экспертных суждений и замечаний.

вания — пользователь методики. Данный стандарт разрабатывается в эпоху информационно-компьютерной революции, которая уже привела к распространению новых видов тестирования — прежде всего, дистанционного онлайн-тестирования, ставшего особенно актуальным для крупных территориально распределенных компаний. Стандарт призван решить данную задачу на уровне общих для всего профессионального сообщества методических, технологических и этических требований к процедурам тестирования и правилам использования их результатов. При разработке российского стандарта принимались во внимание рекомендации национальных и международных руководств по разработке и применению тестовых методик. Настоящее первое издание российского стандарта тестирования персонала — это общепринимое практико-ориентированное методическое руководство. Оно закладывает фундамент для дальнейшей разработки более детальных и узкоспециализированных стандартов, а также ложится в основу «Профессионального стандарта специалиста по оценке персонала» наряду со стандартами иных оценочных методов и процедур.

**Ключевые слова:** российский стандарт тестирования персонала; тестирование; онлайн-тестирование; тестовые методики; психодиагностика; оценка персонала; психометрические тесты; надежность; валидность; профессиональный отбор; тесты профессиональных знаний.

## Оглавление

### Глава 1. Характеристика Российского стандарта тестирования персонала

- 1.1. Обоснование необходимости разработки стандарта
- 1.2. Используемые прототипы
- 1.3. Статус данного стандарта
- 1.4. Назначение и цели стандарта
- 1.5. Целевые группы стандарта
- 1.6. Выполнение предписаний стандарта: обязательность и желательность

### Глава 2. Понятие теста, классификация тестов и ситуаций тестирования

- 2.1. Определение теста
- 2.2. Признаки теста
- 2.3. Психометрические требования к качеству тестов
- 2.4. Классификация тестов
- 2.5. Нетестовые оценочные процедуры и методики
- 2.6. Социально-психологическая ситуация тестирования

### Глава 3. Подготовка тестирования в организации

- 3.1. Типы контроля
- 3.2. Подготовка инструкции
- 3.3. Подготовка бланков или буклетов
- 3.4. Подготовка администратора тестирования
- 3.5. Подготовка помещения для тестирования
- 3.6. Проектирование сеансов тестирования
- 3.7. Подготовка участника — информирование
- 3.8. Дополнительное информирование

### Глава 4. Требования к методикам тестирования

- 4.1. Принципы объективности и стандартизации
- 4.2. Оценка качества тестовой методики
- 4.3. Вероятностный смысл результатов и погрешность измерения
- 4.4. Требования к надежности
- 4.5. Требования к валидности
- 4.6. Требования к тестовым нормам и репрезентативности
- 4.7. Требования к достоверности
- 4.8. Требования к лицензионной чистоте методик тестирования

### Глава 5. Требования к квалификации исполнителей

- 5.1. Пользователи тестовых методик
- 5.2. Администраторы сеансов тестирования
- 5.3. Консультанты-наставники или обучающие работе с тестовой методикой
- 5.4. Разработчики тестовых методик

5.5. Эксперты в предметной области

5.6. Программисты-разработчики компьютерных программ тестирования

## **Глава 6. Проведение тестирования**

6.1. Регламент тестирования

6.2. Предварительная запись на сеансы тестирования

6.3. Предварительное информирование и консультация

6.4. Условия проведения тестирования

6.5. Правила поведения для участника тестирования

6.6. Информированное согласие

6.7. Действия и обязанности администратора тестирования

6.8. Права администратора (наблюдателя)

6.9. Проведение дистанционного онлайн-тестирования

6.10. Завершение сеанса тестирования

6.11. Облегченные требования к консультативному тестированию

6.12. Развивающее испытание, но не прямое обучение

## **Глава 7. Обработка, интерпретация и анализ результатов**

7.1. «Сырые баллы»

7.2. Стандартизация тестовой шкалы

7.3. Стандартные шкалы

7.4. Процентили

7.5. Использование локальных нормативных групп

7.6. Погрешность измерения

7.7. Интерпретация результатов тестирования

7.8. Различные подходы к анализу результатов тестирования

7.9. Расчёт интегрального рейтинг-балла

7.10. Разделение тестируемых на группы

7.11. Сравнение с «идеальным профилем»

## **Глава 8. Информационные технологии в тестировании**

8.1. Условия для выбора компьютерной формы

8.2. Виды компьютерного тестирования и типы ситуаций

8.3. Особенности подготовки участников к компьютерному тестированию

8.4. Подготовка оборудования

8.5. Режим секретности при компьютерном тестировании

8.6. Аутентификация и авторизация

8.7. Технические возможности в профессиональной АСТ

8.8. Требование лицензионной чистоты

8.9. Соглашение о конфиденциальности с соисполнителями

8.10. Действия организатора в случае нарушения информационной безопасности

8.11. АОС и АСТ

8.12. Технический регламент

## **Глава 9. Сопряжение с другими оценочными процедурами и контроль качества тестирования**

9.1. Схемы сопряжения

9.1.1. Различные комбинации тестовых и нетестовых технологий

9.2. Согласование результатов при сопряжении

9.3. Выборочный контроль

9.4. Проверка валидности тестирования

9.5. Принцип превентивного (упреждающего) тестирования

9.6. Количественная и ранговая модели сопряжения с нетестовыми методиками

## **Глава 10. Обратная связь и принятие решений по результатам тестирования**

10.1. Необходимость обратной связи

10.2. Функции и цели обратной связи для тестируемых

10.3. Формы обратной связи

10.4. Содержание обратной связи

10.5. Требования к исполнителям и технической реализации

- 10.6. Информирование заказчиков тестирования и лиц, принимающих решения
- 10.7. Принятие решений на основе результатов тестирования
- 10.8. Использование и распространение результатов тестирования

### **Глава 11. Этические нормы в тестировании персонала**

- 11.1. Этика, право и этико-правовые дилеммы
- 11.2. Компетентность как этическая норма
- 11.3. Этические нормы, связанные с преимуществами особой информированности
- 11.4. Этика деловой коммуникации (взаимоотношений) в ситуации оценки
- 11.5. Национально-культурные особенности
- 11.6. Информированное согласие и использование результатов тестирования
- 11.7. Конфиденциальность и работа с полученными данными

#### **Заключение**

ПРИЛОЖЕНИЕ 1. Перечень рекомендуемой литературы

ПРИЛОЖЕНИЕ 2. Глоссарий основных терминов и сокращений

ПРИЛОЖЕНИЕ 3. Некоторые важнейшие расчетные формулы

## **Глава 1. Характеристика Российского стандарта тестирования персонала**

### **1.1. Обоснование необходимости разработки стандарта**

В последние десятилетия использование тестовых методик получило широкое распространение в практике оценки персонала российских компаний, однако до настоящего момента не существовало единого и общепринятого документа, содержащего перечень требований и критериев, позволяющих оценить качество тестирования. Очевидно, что некорректное использование тестовых методик приводит к серьезным негативным последствиям как для людей, выполняющих данные методики, так и для организаций, использующих их результаты в процессе выработки и принятия различных по масштабу и значимости управленческих решений. Реализация потенциала конкретных тестовых методик зависит в том числе от того, насколько грамотно и уместно они применяются, а ошибки в применении обесценивают результаты даже в случае использования изначально качественных методик.

Данный стандарт призван разъяснить и популяризировать в русскоязычном пространстве один из базовых принципов распределения ответственности: за качество теста отвечает разработчик, а за качество тестирования — пользователь методики. Данный стандарт разрабатывается в эпоху информационно-компьютерной революции, которая уже привела к распространению новых видов тестирования, прежде всего, дистанционного онлайн-тестирования, ставшего особенно актуальным для крупных территориально распределенных компаний. Стихийное возникновение терминов, описывающих новые технологические процессы, процедуры и виды тестирования (например, дистанционное онлайн-тестирование), обозначило острую потребность в унификации используемых понятий для улучшения взаимопонимания между разработчиками и пользователями тестовых методик.

Стандарт призван решить данную задачу на уровне общих для всего профессионального сообщества методических, технологических и этических требований к процедурам тестирования и правилам использования их результатов.

### **1.2. Используемые прототипы**

При разработке российского стандарта принимались во внимание рекомендации национальных и международных руководств по разработке и применению тестовых методик, в число которых входят следующие документы:

- Временный стандарт требований к методикам, принятый в 2011 году секцией психодиагностики Российского психологического общества (Батурин, Мельникова, 2012);
- International Test Commission. ITC Guidelines on Test Use (2013);

— British Psychological Society. Guidelines for the Development and Use of Computer-Based Assessments (2002);

— European Federation of Psychologists' Associations. EFPA Review model for the description and evaluation of psychological and educational tests — Test review form and notes for reviewers (2013);

— New Zealand Psychologists Board. Guidelines on the use of Psychometric Tests (2013);

— American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for educational and psychological testing (1999).

Рассматриваемые стандарты свидетельствуют об опыте применения тестовых методик, накопленном в Европе, США, России и других странах.

Настоящее первое издание российского стандарта тестирования персонала — это общеприменимое практико-ориентированное методическое руководство. Оно закладывает фундамент для дальнейшей разработки более детальных и узкоспециализированных стандартов, а также ложится в основу «Профессионального стандарта специалиста по оценке персонала» наряду со стандартами иных оценочных методов и процедур.

### 1.3. Статус данного стандарта

Стандарт является научно-обоснованным документом. Стандарт является не юридической, но профессиональной, культурной и этической нормой, то есть профессиональному сообществу предлагается принять правила и требования стандарта добровольно. Все положения стандарта являются руководством для практических работников, проводящих оценку кандидатов и работников внутри организации или оказывающих услуги по оценке персонала другим организациям. Документ содержит рекомендации, несоблюдение которых может приводить к искажению или снижению качества результатов тестирования персонала в организациях.

Первая версия стандарта разработана в 2014—2015 гг. специальной комиссией, созданной в рамках Федерации оценки персонала, входящей в состав Национальной конфедерации Развития Человеческого Капитала (НК РЧК). В состав комиссии вошли как ведущие академические специалисты, так и представители коммерческих организаций, разрабатывающие и распространяющие тесты, а также оказывающие услуги по оценке персонала с использованием тестовых технологий. В таком статусе данный документ принят как межкорпоративный стандарт для всех организаций — провайдеров услуг по оценке персонала, входящих в состав Федерации Оценки Персонала. Предполагаемый статус документа в будущем — общероссийский.

### 1.4. Назначение и цели стандарта

Формирование и широкое распространение современных научно обоснованных представлений о методе тестирования<sup>1</sup> и корректном использовании результатов тестовых методик.

Регламентация деятельности по применению тестовых методик.

Информирование о правилах, повышающих эффективность применения тестовых методик.

Нормативная поддержка деятельности по применению тестов и создание условий для профессионального арбитража<sup>2</sup>.

1 Здесь и далее по умолчанию речь ведется о методе тестирования применительно к кандидатам и сотрудникам организаций. В иных случаях делаются специальные примечания и оговорки.

2 «Профессиональный арбитраж» в данном контексте — это урегулирование спорных вопросов, возникающих между пользователями (заказчиками) и разработчиками тестовых технологий, а также между сотрудниками, проходящими тестирование, и организациями, принимающими решение на основе результатов тестов.

Информационное обеспечение оценки качества тестовых методик.

Информационная поддержка лиц, принимающих решения об использовании тестовых методик.

Стимулирование прикладных исследований, подтверждающих эффективность применения тестовых методик.

### 1.5. Целевые группы стандарта

#### 1.5.1. Стандарт адресован тем, кто:

- является пользователем тестовых методик;
- принимает решения об использовании тестовых методик;
- консультирует и проводит экспертизу по вопросам разработки и применения тестовых методик;
- продает тестовые технологии как услугу, то есть проводит проекты, основанные на тестовых технологиях;
- разрабатывает и продает тестовые методики как продукт;
- проводит исследования эффективности применения тестовых методик;
- обеспечивает организационные и технические условия для проведения тестирования;
- проходит тестовые методики.

1.5.2. Главы документа преднамеренно содержат ряд повторов, так как адресованы кроме узкой группы специалистов по тестовой оценке также и другим подгруппам читателей, которым не обязательно читать весь текст документа. Кроме того, повторы дают возможность обращаться за какой-либо справкой к отдельным главам без необходимости чтения всего Стандарта<sup>3</sup>. Например, глава 8 адресована не только специалистам по тестированию, но и специалистам по информационным технологиям, и по-другому освещает для них вопросы, изложенные в главе 6. Взаимосвязанными и частично повторяющимися являются также некоторые параграфы в главах 3 и 6.

В тех случаях, когда организаторы разработки (или подготовки) системы тестирования и администраторы отдельных сеансов тестирования непосредственно на местах — это разные группы исполнителей, первым рекомендуется читать, прежде всего, главу 3, а вторым — главу 6. Глава 11 адресована участникам конфликтных комиссий, которые создаются для рассмотрения возможных жалоб участников тестирования, и эта глава частично пересекается с содержанием глав 3, 5 и 6. Наличие повторов обусловлено также и тем, что они облегчают выполнение требований стандарта на разных этапах тестирования: на этапе подготовки достаточно перечитать главу 3, на этапе проведения — главу 6, на этапе обратной связи — главу 10 и т.п. Большинство рекомендаций по чтению глав различными специалистами суммируются в таблице 1.1.

Таблица 1.1 Рекомендации по чтению глав Стандарта тестирования персонала

Специалисты	Перечень рекомендуемых глав
Пользователи тестов	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11
Администраторы сеансов тестирования	3, 5, 6, 8
Консультанты-наставники	1, 2, 3, 4, 7, 9, 10, 11
Разработчики тестовых методик	1, 2, 4, 5, 6, 7
Эксперты в предметной области	2, 5, 10
Программисты-разработчики	3, 6, 8, 9
IT-специалисты (системные администраторы компьютерных сетей)	8
Лица, принимающие кадровые решения	9, 10, 11

<sup>3</sup> Повторы могут быть отменены при наличии мотивированной необходимости, однако это планируется сделать не ранее 2016 года.

### 1.6. Выполнение предписаний стандарта: обязательность и желательность

Большинство предписаний стандарта следует рассматривать как методические рекомендации, то есть как желательные, но не обязательные для исполнения профессиональные нормативы. Исключения составляют те из них, которые напрямую продиктованы действующим законодательством РФ. Следование данным рекомендациям повышает качество процедуры тестирования и его результатов, их игнорирование — снижает вплоть до критически низкого уровня.

Данный стандарт не подразумевает разделение всех тестовых методик и всех мероприятий по тестированию только на две категории — «пригодные» и «непригодные». Данный стандарт призван помочь разделить тестовые методики и мероприятия по тестированию на большее число качественных категорий: чем больше предписаний стандарта выполняется, тем выше качество теста и соответствующей программы его применения. Таким образом, данный стандарт *не запрещает* использовать тесты невысокого качества, однако помогает ранжировать их по качеству. Для оценки качества применяемой технологии тестирования рекомендуется хотя бы грубо оценить процент признаков этой технологии, соответствующих данному стандарту. Если более половины признаков противоречат нормативам стандарта, то это является серьезным аргументом для пользователя в пользу прекращения отношений с исполнителями или поставщиками данной технологии.

При наличии выбора между несколькими потенциальными организаторами тестирования, каждый из которых соответствует большинству предписаний стандарта (более 50%), с точки зрения качества следует отдавать предпочтение тому из них, кто выполняет большее количество нормативов, демонстрируя тем самым большую степень соответствия своей работы стандарту.

## Глава 2. Понятие теста, классификация тестов и ситуаций тестирования

### 2.1. Определение теста

В данном документе тестом (тестовой методикой, методикой тестирования) называется стандартизированная методика оценки, направленная на измерение индивидуальных свойств и качеств работника (кандидата), прямо или косвенно связанных с возможностью эффективного осуществления им профессиональной деятельности.

### 2.2. Признаки теста

Методика оценки может называться тестом при *одновременном* наличии у нее следующих признаков:

- стандартизированный набор вопросов или заданий иного типа;
- одна или несколько измерительных шкал, позволяющих выразить результаты количественно;
- связь каждого ответа на каждое задание с одной или несколькими измерительными шкалами (наличие «ключей к тесту»)<sup>4</sup>;
- стандартизированная процедура проведения, включающая однозначную (стандартную) инструкцию для тестируемого, правила использования вспомогательной информации, правила завершения или приостановки тестирования и т.п.;

<sup>4</sup> В отличие от социологических опросников, а также от методик прямого субъективного шкалирования (ранжирования), вопросы в тестовых методиках порождают ответы, связь которых с измеряемыми свойствами для респондентов не очевидна, тем самым «ключи к тесту» всегда представляют собой определенный «методический секрет».

- возможность автоматической (без участия человека) обработки результатов, то есть формализованная процедура подсчета баллов по шкалам с помощью весовых коэффициентов (ключей);
- тестовые нормы — фиксированные границы перевода тестовых баллов в оценочные категории;
- формализованная модель интерпретации результатов и/или рекомендации по принятию тех или иных решений, связанные с определенными интервалами значений на шкале (шкалах) и сочетаниями значений шкал (при наличии двух шкал и более);
- направленность на индивидуальную количественную оценку какой-либо характеристики одного человека (а не группы, коллектива и т.п.).

### 2.3. Психометрические требования к качеству тестов

2.3.1. Тест считается качественным измерительным инструментом, если:

- имеются данные, подтверждающие **надежность** тестовых шкал на количественном уровне;
- имеются данные, подтверждающие **валидность** тестовых шкал как на качественном, так и на количественном уровне;
- имеются данные, подтверждающие **репрезентативность** тестовых норм (правил присвоения определенным тестовым результатам заданных оценочных категорий для той выборки, на которой применяется тест);
- имеются формализованные процедуры, позволяющие устанавливать **достоверность** отдельного протокола тестирования.

Измерительные тесты, направленные на диагностику психических свойств работников (интеллектуальные способности, личностные черты, поведенческие установки, мотивационные предпочтения и др.) и отвечающие перечисленным требованиям к их качеству, называются психометрическими<sup>5</sup>.

Один из наиболее важных признаков измерительных (психометрических<sup>6</sup>) тестов — это наличие калибровки (разметки) шкалы тестовых баллов, которая обосновывается статистически (при нормативно-ориентированном тестировании<sup>7</sup>) или экспертно (при критериально-ориентированном тестировании<sup>8</sup>).

2.3.2. Методики, имеющие признаки теста, но не отвечающие психометрическим требованиям к качеству, могут быть названы **квazitестами**. К этой категории могут быть отнесены методики следующего вида:

- шуточные «тесты» развлекательной направленности (например, тест «Какой Вы родитель?» или тесты на корпоративных сайтах, размещаемые только с целью рекламы и привлечения новых кандидатов);

5 Тесты на профессиональные знания, прошедшие проверку психометрических свойств, называются в данном стандарте психометрическими, так как в деятельности тестируемых усматриваются общие психологические закономерности при выполнении различных по содержанию и направленности тестов.

6 В данном стандарте термины «измерительные тесты» и «психометрические тесты» применяются как синонимы, несмотря на то, что ряд авторов тестов знаний не считают тесты профессиональных знаний психометрическими.

7 Нормативно-ориентированное тестирование (НОРТ) предполагает формирование тестовых норм на основе статистического распределения реальных тестовых баллов на представительной выборке стандартизации. Инструментами калибровки являются точки на нормированных шкалах (на Z-шкале и ее линейных производных, например, T баллах, баллах IQ и т.п.).

8 Критериально-ориентированное тестирование (КОРТ) предполагает формирование тестовых норм на основе экспертно-предметных соображений. Инструментами калибровки являются экспертно заданные абсолютные пороговые значения отбора по проценту правильных ответов (например, в квалификационных тестах: к работе допускаются операторы транспортных или энергетических систем, способные выполнять действия с точностью, обеспечивающей безопасность производства, близкую к 100%).



- «авторские» тесты (например, тесты, составленные преподавателями для приближенного текущего контроля степени усвоения знаний узкой группой учащихся по итогам проведенного производственного обучения);
- исследовательские (предварительные) версии будущих психометрических тестов с непроверенными к текущему моменту основными психометрическими свойствами;
- ситуационные тест-опросники, созданные внутренними экспертами организации для ее локальных нужд;
- эвристические процедуры, направленные на развитие самопознания (на осознание человеком наличия у себя определенных психических процессов и свойств) без предполагаемого сравнения себя с другими.

Квазитест включает в себя, как правило, все те же компоненты, что и измерительный (психометрический) тест — фиксированный набор заданий, ключи, нормы, шкалы — но эти компоненты не обоснованы эмпирико-статистической проверкой и сопутствующей ей калибровкой шкалы тестовых баллов. При использовании квазитестов в ходе оценки кандидатов и работников организации к этим методикам необходимо применять те же самые методические и профессионально-этические требования по подготовке, проведению и доступу к результатам, что и к психометрическим тестам.

Если применение какого-либо квазитеста является регулярной практикой, то необходимо организовать работу по проверке его психометрических свойств (произвести последовательный анализ распределения баллов, надежности, валидности и других параметров — см. главу 4). Ни один квазитест не дает возможности прогнозировать успешность профессиональной деятельности с заданной точностью (погрешностью, или вероятностью ошибки), что необходимо понимать при использовании методик подобного типа в работе.

## 2.4. Классификация тестов

Методическое руководство к тесту (руководство пользователя) должно содержать достаточно точное отнесение тестовой методики к определенному классу.

2.4.1. *По содержательной направленности* (что измеряется) тесты делятся на:

- методики оценки когнитивных (познавательных, интеллектуальных) способностей;
- методики оценки психомоторных способностей и навыков (в частности, на точность восприятия, скорость, силу, выносливость и координацию движений и ручных операций);
- методики психологической и психофизиологической диагностики функциональных и эмоциональных состояний работника;
- методики оценки личностных черт, мотивов и социальных установок (личностные, мотивационные и иные);
- методики оценки профессиональных знаний и уровня усвоения должностных обязанностей (в частности, для контроля степени усвоения знаний по результатам производственного обучения);
- методики решения ситуационных задач (в частности, диагностика практических профессиональных умений, оценка сформированности управленческих и иных навыков, уровня развития производственных компетенций, кейс-тесты);
- методики измерения управленческого потенциала, индивидуального стиля делового общения и взаимодействия, межличностных отношений (в частности, диагностика стиля управления, командных ролей);
- социально-психологические и социологические методики (в частности, методики измерения уровня удовлетворенности работой, тест-опросники профессиональных ценностей).

#### 2.4.2. *По типу*<sup>9</sup> (как измеряется) тесты (тестовые задания) делятся на:

- задания с выбором одного из готовых вариантов ответа;
- задания с выбором нескольких из готовых вариантов ответа;
- задания на ранжирование или установление соответствия;
- методики шкалирования, предполагающие вынесение тестируемым балльных оценок для фиксированного набора объектов;
- хронометрические тесты на измерение времени ответа;
- методики, состоящие из заданий со свободным кратким ответом, который может быть автоматически обработан (например, числовой ответ или набор словесных ответов, в котором можно формально установить конечное множество всех правильных ответов);
- методики, предполагающие заполнение пропусков в связном тексте, в которых каждый ответ может быть формально однозначно расценен как правильный или неправильный (метод контекстной вставки, или cloze/close test);
- тесты на сложность (или мощность), состоящие из одного задания (или небольшого числа сложных заданий), в которых тестовый балл определяется комбинацией времени решения этих заданий в сочетании с избранным способом решения;
- тесты, основанные на измерении скорости психомоторной реакции;
- компьютерные имитационные методики (симуляторы), выдающие общий количественный показатель успешности игрока;
- контент-анализ кратких свободных текстов (например, сочинений на тему «Я сам», «Моя компания» и т.п.);
- поведенческие тесты, предполагающие выполнение тех или иных стандартизированных практических заданий и последующее измерение качества полученного результата (например, сварка деталей между собой и измерение качества сварного шва, подготовка документов по заданным критериям).

Следует заметить, что в рамках одного теста или даже одного задания возможна комбинация нескольких подходов (например, измеряется не только правильность ответа, но и скорость решения).

#### 2.4.3. *По форме проведения* тесты подразделяются на:

- компьютерные;
- бланковые;
- настольные, или мануальные (например, раскладывание карточек, сборка из кубиков, деталей и т.п.);
- поведенческие;
- имитационно-игровые.

Родственным этой классификации является деление тестов на вербальные (словесные), невербальные (рисуночные), устные, или аудиальные (предъявляемые на слух), письменные, или визуальные (предъявляемые для зрительного восприятия), мануальные (предъявляемые для выполнения ручных операций), локомоторные (требующие перемещения в пространстве) и т.п. Более подробная классификация различных типов компьютерного тестирования предлагается в главах 3, 6 и 8.

#### 2.4.4. Выбор теста должен определяться наличием ответа на следующие вопросы:

- «что мы хотим измерить?» (какие компетенции, профессионально-важные качества и т.п.);
- «как связаны значения шкалы теста и сочетания значений шкал теста с тем, что мы хотим измерить (предметом оценки)?»;

<sup>9</sup> Классификация 4.2 относится к форме теста (формату тестовых заданий). Формат заданий сам по себе не предопределяет статус теста как психометрического инструмента или квазитеста.

— «для чего мы измеряем ту или иную характеристику?» (какая стоит управленческая задача, как и кем будут рассмотрены результаты тестирования и соответствующие рекомендации, какие возможные управленческие решения принимаются);

— «проводится ли тестирование в ситуации консультирования или в ситуации экспертизы?» (какие решения и кем будут приниматься на основе результата).

Помимо этого выбор теста обосновывается:

- необходимой степенью защиты от сознательных или неосознанных искажений со стороны тестируемых,
- количеством времени, которое тестируемые могут затратить на прохождение теста;
- имеющимися организационно-технологическими возможностями по проведению и обработке результатов тестирования (помещение, оборудование и т.п.);
- характеристиками контингента тестируемых (возраст, количество, образовательный уровень, уровень должности, география расположения рабочих мест и т.п.);
- возможностью сопоставить результаты тестирования с другой доступной информацией про тестируемых (например, результатами ранее проведённого тестирования).

При условии соблюдения требований настоящего Стандарта как к тестам, так и к процедуре организации и проведения тестирования, важным признаком профессионализма является ориентация на предпочтение тестов, наиболее экономически выгодных для заказчика, а также наиболее выгодных вариантов организации и проведения тестирования.

## 2.5. Нетестовые оценочные процедуры и методики

Далеко не все полезные на практике оценочные процедуры и диагностические методики являются тестами. Если хотя бы один из основных компонентов теста (задания, ключи, нормы, шкалы) зависит не от разработчика, а от пользователя методики, ее нецелесообразно называть тестом. Тестами (тестовыми методиками) *не являются*:

— проективные методики, имеющие стандартные наборы заданий (стимулов), но не имеющие формализованных ключей и/или количественных шкал, а также допускающие различия в трактовке свободных ответов у разных экспертов (например, методики ТАТ, Пятна Роршаха, рисуночные проективные методики);

— ситуационные социологические и социально-психологические опросники, предназначенные для выявления качественных особенностей отношения групп сотрудников к своей компании, своей продукции, условиям труда, руководству или корпоративной культуре;

— психофизиологические методики, основанные на регистрации комплекса показателей функционального и эмоционального состояния (полиграф), дающие многомерный паттерн результатов — в этих методиках чаще всего отсутствует стандартизированный набор заданий (вопросов, стимулов), так как набор вопросов подбирается исходя из ситуационного контекста;

— деловые игры, предполагающие применение экспертных шкал для оценки комплекса показателей (компетенций) по признакам наблюдаемого поведения;

— методики шкалирования (рейтингования), основанные на сборе оценок сотрудника со стороны его руководителей, коллег, подчиненных, клиентов (например, «360 градусов», социометрия, групповая оценка личности) — в этом случае не стандартизован круг оценщиков (он является разным для разных оцениваемых), тестовые задания могут отсутствовать и замещаться наблюдением на рабочем месте;

— обучающие методики, предполагающие выдачу обратной связи о правильном или неправильном ответе после каждого задания — в этом случае, несмотря на тестообразный

формат заданий, процедура не является в полной мере стандартизированной<sup>10</sup> и направлена в первую очередь на обучение, а не диагностику;

— интервью любого типа, в том числе с последующей формализованной или полужформализованной оценкой со стороны интервьюера деловых и личностных качеств кандидата (сотрудника) по заданным критериям (компетенциям);

— оценка ключевых показателей эффективности (КПЭ) и любые другие методики, в которых оценивается процесс и/или результаты труда сотрудника на рабочем месте;

— все остальные оценочные методики, не обладающие признаками теста.

## 2.6. Социально-психологическая ситуация тестирования

Условия, в которых проходит тестирование, можно разделить на две основные группы — *ситуация консультирования* и *ситуация экспертизы*.

**2.6.1. Ситуация консультирования** (синонимы: клиентское тестирование, консультативное тестирование, конфиденциальное тестирование, конфиденциальная оценка, помогающее тестирование, самотестирование, самооценка). Ключевые признаки ситуации консультирования:

— решение о прохождении тестового обследования принимает сам тестируемый;

— основной (а иногда и единственный) пользователь результатов — сам тестируемый;

— решения по итогам тестирования также принимает сам тестируемый.

Наиболее распространенные ситуации консультирования: а) тестирование в целях добровольной профориентации (консультирование по проектированию карьеры); б) тестирование с целью составления индивидуального плана развития (проектирования карьеры), когда его инициатором является сам сотрудник; в) тестирование в ситуации добровольного обращения за психологической или медико-психологической помощью; г) тестирование для самооценки уровня собственных профессиональных знаний и/или умений (например, с целью выявления потребности в обучении, промежуточное самотестирование при прохождении образовательной программы).

**2.6.2. Ситуация экспертизы** (синонимы: ситуация внешней оценки, тестовый экзамен, административное тестирование, ассесмент-тестирование, аттестационное тестирование, high stakes assessment, профессиональное испытание). Ключевые признаки ситуации экспертизы:

— проводится, как правило, по инициативе администрации организации;

— основным пользователем результатов также является администрация организации;

— решения по итогам (или с учетом результатов) тестирования принимаются не самим тестируемым, а уполномоченными представителями организации (руководителями, службой персонала и т.п.).

Наиболее распространенные ситуации экспертизы или оценки: а) отбор при приеме на работу; б) формирование кадрового резерва; в) ротация кадров; г) аттестация<sup>11</sup>.

В рамках решения одной кадровой задачи на разных этапах возможно сочетание ситуаций консультирования и экспертизы. Разграничение консультативного и административного тестирования требуется для того, чтобы не смешивать разные стандарты требований для существенно разных ситуаций. Например, для проведения тестового компьютеризированного экзамена (ситуация экспертизы) требуется подготовка особого организационного документа под названием «технический регламент». Этот документ разграничивает права

<sup>10</sup> Например, случайный правильный ответ может воодушевлять слабого тестируемого, а случайный неправильный при наличии «субъективно трудного» задания может угнетать сильного тестируемого.

<sup>11</sup> Аттестация — процесс, жёстко регламентированный ТК РФ, и к нему предъявляются более высокие требования, чем к остальным ситуациям.

парольного доступа для разных категорий пользователей компьютерной системы к тестовым материалам и результатам тестирования. А для консультативного тестирования многие подобные требования стандарта не являются необходимыми (см. подробнее главу 8).

2.6.3. Стратегии поведения тестируемого в ситуации консультирования и в ситуации экспертизы, как правило, различны: если в первой доминирующей мотивацией тестируемого является самопознание, то во второй — стремление показать наилучший результат. Стремление показать наилучший результат может провоцировать попытки осознанной или неосознанной фальсификации со стороны тестируемого. В этом случае тесты, хорошо работающие в ситуации консультирования, могут оказаться непригодными для ситуации экспертизы<sup>12</sup>.

Факт проверки основных психометрических свойств теста в ситуации консультирования не дает возможности гарантировать достижение ими аналогичных показателей в ситуации экспертизы. Выбор теста для использования в ситуации экспертизы должен основываться на данных о его психометрических свойствах, ранее установленных именно в рамках ситуации экспертизы.

### Глава 3. Подготовка тестирования в организации

Тщательная подготовка процедуры тестирования является залогом лояльности тестируемых и позитивного восприятия ими результатов как достоверных и справедливых. В данной главе рассматриваются ее технологические и организационные аспекты.

#### 3.1. Типы контроля<sup>13</sup>

Процесс подготовки тестирования может различаться в зависимости от типа контроля:

3.1.1. Тестирование в присутствии администратора: очное бланковое или компьютерное тестирование, при котором в помещении физически присутствует администратор тестирования (супервизор). Возможна многоуровневая сетевая организация тестирования, предполагающая участие главного администратора в центре и нескольких удаленных администраторов на местах. Проведение такого тестирования требует подготовки администраторов, помещения (включая подготовку аппаратного и программного обеспечения), проектирования сеансов тестирования, информирования участников.

3.1.2. Контролируемое дистанционное тестирование без физического присутствия администратора: ответственность администратора состоит в обеспечении аутентификации (установления личности) тестируемых и дистанционном контроле выполнения ими тестовых процедур. Пример технической реализации дистанционного контроля — видеонаблюдение с помощью веб-камер. Проведение такого тестирования требует подготовки администратора, проектирования сеансов тестирования, информирования участников, тщательной проверки и перепроверки достоверности результатов.

3.1.3. Адресное дистанционное онлайн-тестирование: тестирование проходит без непосредственного или удалённого присутствия администратора, доступ к тестированию предоставляется заранее определённым участникам в виде индивидуальных логина и пароля к тесту или серии тестов. Проведение такого тестирования требует проектирования схемы информационных потоков, обеспечения автоматизированной аутентификации и четкого заочного информирования участников, при этом полная гарантия достоверности резуль-

12 В особенности это касается тест-опросников, не снабженных специальными инструментами по защите от фальсификаций — например, шкалами лжи, ипсативным форматом предъявления заданий и т.д.

13 Классификации ситуаций тестирования по типу контроля и по типу принятия решения не являются тождественными, хотя отсутствие физического наблюдателя в помещении часто сочетается с самостоятельным принятием решения по результатам тестирования самим участником, а присутствие администратора — с принятием решения со стороны заказчика (руководства организации и иных уполномоченных лиц).

татов отсутствует. Совершенствование возможностей контроля в ситуациях 3.1.2 и 3.1.3 является передовым фронтом технологического развития дистанционного тестирования. Подробнее о дистанционном контроле см. главу 8.

3.1.4. Открытое дистанционное онлайн-тестирование: доступ к тестированию осуществляется путем самостоятельной регистрации участников (например, через специальную форму на корпоративном сайте). Проведение такого тестирования не дает гарантий точной аутентификации и достоверности результатов тестирования, поэтому данная модель контроля удобна для консультативного тестирования, а в ситуации экспертизы она нуждается в перепроверке в ходе иных оценочных процедур. В этом случае, как правило, отсутствуют сеансы, и тестируемые вольны выбирать удобное для себя время выполнения теста(ов) в течение нескольких суток. Гарантия достоверности результатов отсутствует. Подробнее о самотестировании см. главу 8.

### **3.2. Подготовка инструкции**

3.2.1. Стандартная инструкция и ее модификации. Каждый тест снабжается стандартной инструкцией, которая должна быть строго идентичной и понятной для всех тестируемых. Стандартная инструкция может быть дополнена информацией, описывающей цели тестирования в конкретной организации, а также порядок доступа к результатам. Подобную информацию рекомендуется включать в иные документы (например, в предварительное информационное письмо, во внутренние «Правила тестирования» и «Памятку участника», в текст информированного согласия). Текст стандартной инструкции к психометрическому тесту может быть изменен только по согласованию с его разработчиком<sup>14</sup>.

В случае использования квазитестов или тестов, созданных внутри организации, модификация инструкции с целью приспособления к особым условиям признается допустимой. Такую инструкцию правильнее называть стандартизированной, а не стандартной. При использовании методики, существующей как в бланковой, так и в компьютерной форме, необходимы варианты инструкции, учитывающие различия в форме ее проведения.

3.2.2. Содержание инструкции. Стандартная инструкция включает в себя информацию о количестве заданий в тесте, времени их выполнения, способах выполнения заданий различной степени сложности (при их наличии) и их примеры. В отдельных случаях инструкция может включать в себя информацию о времени предоставления обратной связи и о решениях, которые могут быть приняты на основе результатов тестирования. Технически это достигается при использовании двух разноуровневых инструкций: частные инструкции к отдельным методикам (или субтестам) должны содержать только информацию о правилах выполнения теста. При этом общая инструкция к батарее методик может содержать организационную информацию. Таким образом, информация, которая предоставляется тестируемому в текстах стандартной инструкции и информированного согласия, может частично дублироваться (с целью повышения вероятности донесения до него важной информации), однако эти документы являются обособленными по статусу.

### **3.3. Подготовка бланков или буклетов**

Некомпьютеризированное, или бланковое тестирование (его относят также к типу «карандаш—бумага», но этот термин уже считается устаревшим) предполагает заблаговременную подготовку тестовых буклетов или бланков в необходимом количестве. Тестовый буклет представляет собой брошюру с тестовыми заданиями и полями для ответов, распо-

<sup>14</sup> Изменение инструкции с целью приспособления стандартного теста к особым условиям входит в комплекс работ по кастомизации теста, в котором в обязательном порядке должен принимать участие разработчик теста.

ложенными на тех же страницах. Тестовый бланк представляет собой отдельный лист с полями для ответов, заполняемый тестируемым и прилагающийся к брошюре с тестовыми заданиями. Применение лицензионных коммерческих тестов предполагает копирование буклетов и бланков в соответствии с рекомендациями разработчика: в ряде случаев разрешается свободное воспроизведение, но в большинстве случаев требуется использовать фирменные буклеты и бланки.

### **3.4. Подготовка администратора тестирования**

Администратор очного тестирования должен быть хорошо знаком с тестом, инструкцией к нему и, в случае если это предусмотрено процедурой тестирования, правильными ответами к примерам заданий, быть готовым к ответам на вопросы тестируемых обо всех аспектах процесса тестирования, но не создавать кому-либо из них преимуществ в прохождении теста и не подсказывать ход решения, а также при необходимости оказывать базовую техническую поддержку (подробнее об обязанностях администратора см. в главе 6).

### **3.5. Подготовка помещения для тестирования**

Очное тестирование в присутствии администратора предполагает специальную подготовку помещения.

3.5.1. Рекомендуемые параметры помещения, оборудованного для проведения тестирования:

- температура воздуха от 22 до 24°C (в холодный период года) или от 23 до 25°C (теплый период года);
- относительная влажность воздуха в пределах 40—60%;
- достаточное естественное и искусственное освещение (наиболее благоприятная освещенность составляет 100-200 люкс);
- размещение рабочих столов таким образом, чтобы они были ориентированы боковой стороной к световым проемам, и естественный свет падал преимущественно слева;
- расстояние от экрана мониторов до глаз тестируемых в пределах 600—700 мм, но не ближе 500 мм;
- шумовая изоляция (уровень шума, не превышающий 60 дБА);
- регулярное проветривание после каждого сеанса тестирования;
- удобство места для тестируемого (минимальная рекомендуемая площадь составляет 4,5 м<sup>2</sup> на одного участника тестирования, расстояние между стульями участников — не менее 0,5 м).

3.5.2. Важно удостовериться в функциональности необходимого оборудования и комплектности канцелярских принадлежностей (компьютер, калькулятор, бумага для записей, бланки или буклеты, письменные принадлежности и т.д.).

3.5.3. Следует проверить, чтобы были выключены телефоны и другие устройства, которые могут отвлекать внимание участников. Если корпоративная культура организации это позволяет, то лучше предоставить возможность участникам выложить все телефоны на время сеанса на стол администратора. Вход в помещение должен быть закрыт для посторонних. На двери рекомендуется повесить табличку «Не беспокоить, идёт тестирование». В тестовой аудитории желательно наличие настенных часов.

### **3.6. Проектирование сеансов тестирования**

3.6.1. При подготовке к очному и дистанционному тестированию следует заранее определить последовательность предъявления всей информации участнику. Вопросы, на которые должен ответить администратор тестирования:

— Необходимо ли предъявлять биографические (анкетные) вопросы и когда — в самом начале, в процессе или по окончании тестирования?

— Принципиальна ли строгая последовательность предъявления тестов и субтестов?

— Возможны ли перерывы между тестами? (Небольшие перерывы при выполнении серии тестов являются допустимыми, при тестировании продолжительностью более 1,5 часов — обязательными.)

— Имеется ли возможность получения обратной связи непосредственно после окончания сеанса тестирования?

3.6.2. Если число тестируемых велико и одновременное тестирование в одном помещении невозможно, целесообразно организовать предварительную запись на сеансы тестирования, заранее проинформировав об этом участников, что позволит избежать длительного ожидания в очереди, опозданий и т.п. (см. также главу 6).

3.6.3. Требования к продолжительности сеансов тестирования. В ходе одного сеанса тестируемому могут быть предъявлены несколько тестов и других оценочных методик (батарея или пакет методик). При этом необходимо учитывать естественные психофизиологические ограничения на длительность сеансов, обусловленные утомлением, колебаниями внимания, возникающими из-за монотонности, и другими факторами. Рекомендуются следующие ограничения на длительность сеансов тестирования:

— для тестирования на бланках — не более 3 часов (180 минут), в середине сеанса желателен перерыв на 5—10 минут;

— для компьютерного тестирования — не более 1 часа 30 минут (90 минут).

Если общая продолжительность выполнения батареи методик превышает указанные ограничения, следует разбить ее на несколько сеансов (два и более) и проводить их в разные дни. Из указанного ограничения вытекает методическое требование к тестовым методикам, затрагивающее и разработчика, и пользователей — они должны быть рассчитаны не более, чем на 90 минут непрерывной работы как в бланковом, так и в компьютерном режимах.

3.6.4. Требование к количеству отдельных тестовых заданий (вопросов). В различных тестовых методиках отдельные тестовые задания (вопросы) значительно отличаются друг от друга по трудоемкости. Короткие вопросы с ответом по типу «да/нет», занимающие не более одной строчки в 60 символов, в среднем требуют не более 5—6 секунд на ответ. Таким образом, большие перечни вопросов в тест-опросниках личностной диагностики (например, ММРП и его модификации) могут включать до 600 заданий, требуя на выполнение не более 1,5 часов. Ответы на вопросы кейсового типа, содержащие в условии развернутые описания проблемных ситуаций на 600 и более символов, могут потребовать от 2 до 5 минут, и в течение стандартного сеанса такие задания могут быть предъявлены в количестве 20—30 штук.

Тесты знаний (теория управления, экономика, юриспруденция, психология, география и т.п.) в большинстве случаев состоят из типичных вопросов с выбором ответа, которые содержат краткое условие (до 120 символов) и 4—5 кратких вариантов ответа (до 60 символов). Если эти задания не требуют расчетов по формулам, их решение обычно занимает от 30 до 60 секунд, и в оптимальном по продолжительности часовом сеансе можно использовать приблизительно от 60 до 100 подобных заданий, а в максимально допустимом сеансе на 1,5 часа — не более 150.

3.6.5. Число тестовых заданий и размах шкалы сырых баллов. Число заданий в тесте влияет не только на утомление тестируемого, но и на полноту охвата различных аспектов измеряемого свойства (полноту охвата области валидности). Если заданий слишком мало, то от этого больше страдает валидность и относительно меньше страдает надежность психометрической методики (см. об этих психометрических параметрах теста подробнее главу 4). Если методика, которая позиционируется разработчиком в качестве психометри-



ческой, содержит менее 20 градаций сырых баллов в одном субтесте (для измерения одного свойства по одной шкале), то психометрические свойства данного теста нуждаются в более внимательном отношении со стороны пользователя методики и более подробно должны быть описаны в руководстве пользователя со стороны ее разработчика<sup>15</sup>.

3.6.6. Хронометрический контроль. Для тестов знаний, созданных внутри организации, рекомендуется проведение предварительного хронометража на небольшой выборке тестируемых (до 10 человек) с целью определения оптимального временного лимита на выполнение теста в целом и/или отдельных субтестов, разделов и заданий. Грамотно заданная временная норма определяется тем, что не менее 70—80% тестируемых укладываются во временной лимит и выполняют все предложенные тестовые задания в течение сеанса. При обнаружении факта явной неадекватности временного лимита (например, если в него укладывается менее 50% тестируемых) в психометрическом тесте, ориентированном на скорость выполнения, необходимо обратиться к разработчику за разрешением смягчить временной лимит или за разработкой локальной версии теста.

### 3.7. Подготовка участника — информирование

Потенциальные тестируемые предварительно оповещаются о приглашении к тестированию, причем объем и качество предварительного информирования существенно влияют на отношение участников к процедуре и играют определяющую роль в обеспечении достоверности результатов.

#### 3.7.1. Цели информирования:

- собственно информировать участников;
- мотивировать участников на ответственное выполнение теста с соблюдением устных и письменных инструкций;
- снизить повышенный уровень тревожности отдельных участников.

3.7.2. При массовом тестировании предварительное информирование рекомендуется проводить в письменной форме. При тестировании сотрудников организации желательно, чтобы информационное письмо было написано от имени одного из их руководителей. Для обеспечения прозрачности и объективности тестирования все участники должны информироваться одинаково.

3.7.3. Сроки, в которые должно осуществляться предварительное информирование, зависят от целей и содержания тестирования. Если дата тестирования известна заранее, рекомендуется сообщить о тестировании не позднее, чем за 3 дня до его начала, что особенно важно при оценке профессиональных знаний. В случае дистанционного онлайн-тестирования момент прохождения тестов тестируемый, как правило, определяет самостоятельно; в этом случае в информационном письме следует предоставлять на выполнение тестовых методик не менее трех дней.

3.7.4. Если численность участников тестирования невелика и есть организационная возможность предварительно собрать всех участников в одном помещении, то оптимальная форма информирования представляет собой очную групповую консультацию, способствующую снятию тревожности в ходе задавания вопросов администратору и общения с иными потенциальными тестируемыми. При отсутствии соответствующей возможности очное взаимодействие может быть заменено общением по электронной почте и другим дистанционным средствам связи.

15 Изменение инструкции с целью приспособления стандартного теста к особым условиям входит в комплекс работ по кастомизации теста, в котором в обязательном порядке должен принимать участие разработчик теста.

3.7.5. В обязательном порядке тестируемые участники информируются о следующих составляющих процедуры тестирования:

- целях тестирования;
- форме проведения и типе контроля тестирования (например, бланковое, компьютерное в присутствии администратора или дистанционное онлайн-тестирование);
- дате и времени проведения тестирования (для бланкового тестирования или тестирования в присутствии администратора), допустимых временных рамках (датах), в пределах которых необходимо пройти тестирование (для дистанционного онлайн-тестирования), общей длительности выполнения тестов;
- обязательности прохождения (например, оно может проводиться по желанию);
- тематической направленности используемых тестов (без упоминания технических названий методик);
- особенностях использования результатов тестирования (перечне лиц, которые получают доступ к результатам тестирования) и сроках их хранения в организации;
- видах, способах и сроках предоставления обратной связи;
- лицах в организации, к которым можно обратиться за дополнительными разъяснениями относительно тестирования;
- об аппаратных и программных требованиях к персональным компьютерам для выполнения компьютерных тестов в режиме онлайн.

### **3.8. Дополнительное информирование**

3.8.1. В зависимости от ситуации и целей тестирования дополнительно рекомендуется, чтобы предварительное информационное письмо участникам содержало следующие пункты:

- польза, которую кандидат может получить от участия в тестировании;
- типы заданий, которые будут использованы в рамках тестирования (например, с контролем времени или без него, тесты с единственным правильным ответом или опросники с возможностью выбора ответа);
- рекомендации по подготовке к сеансу тестирования (например, в случае дистанционного тестирования способностей нужно указать, что соответствующие методики необходимо выполнять в оптимальном функциональном состоянии);
- обоснование выбора конкретных методик;
- предупреждение о возможности выборочного очного повторного тестирования при дистанционном выполнении компьютерного теста в режиме онлайн.

3.8.2. Примеры заданий в компьютеризированных тестах могут быть объединены в демоверсию, доступную для прохождения до начала реального тестирования и призванную познакомить участников тестирования с типами вопросов, которые в нем представлены. Тем не менее, принимая решение о доступности демоверсии, следует подумать о том, в какой степени тренировка тестируемых в выполнении конкретных типов заданий может повлиять на результаты тестирования и исказить их. Есть типы методик (например, методики на формирование «актуальной психологической установки»), в случае которых предварительное знакомство с методикой недопустимо.

3.8.3. В зависимости от политики организации и ожиданий от кандидатов в процессе информирования также может быть затронут вопрос ограниченных возможностей некоторых участников (например, наличие инвалидности, проблемы со зрением или слухом), что меняет формат оценочной процедуры для ряда из них вплоть до отмены тестовой формы и её замены интервьюированием по оцениваемым характеристикам.

## Глава 4. Требования к методикам тестирования

В настоящей главе основное внимание уделено психометрическим требованиям к методикам тестирования. Санитарно-гигиенические ограничения на продолжительность сеанса тестирования и требования к помещению для проведения тестирования приведены в главе 3. Организационно-технические требования рассмотрены в главах 3, 6 и 8.

### 4.1. Принципы объективности и стандартизации

Двумя базовыми требованиями к методикам тестирования выступают объективность и стандартизация. Объективность подразумевает исключение человеческого фактора из цикла тестирования: два пользователя методики должны получать для одного и того же тестируемого одинаковый результат. Стандартизация предполагает такие структуру и процедуру проведения методики тестирования, которые обеспечивают максимально возможное равенство условий для всех участников.

### 4.2. Оценка качества тестовой методики

4.2.1. Полностью универсальных тестовых методик не существует, более того, применение методики в неподходящей для нее ситуации может принести больше вреда, чем пользы. Пользователи методик обязаны уточнить возможности применения конкретной методики: а) в данных условиях (организационном контексте); б) в данной ситуации (ситуация консультирования или ситуация экспертизы); в) для решения данной практической задачи; г) для данных участников; д) для принятия определенных управленческих решений на основе ее результатов.

4.2.2. Для оценки качества тестовой методики и ограничений интерпретации ее результатов следует обратить внимание на требования к ее качеству. Требования к качеству тестов носят комплексный и контекстно-зависимый характер, то есть оценивать применимость методики тестирования следует для каждого конкретного оценочного мероприятия отдельно.

4.2.3. Существует четыре группы общих (основных) *требований к качеству тестовых методик* (психометрических требований): требования к надежности, валидности, репрезентативности и достоверности, причем последнее в большей степени относится к ситуации применения теста, чем к методике как таковой.

4.2.4. Существует также ряд специфичных требований, возникающих при решении определенного круга задач тестирования персонала в контексте конкретных организационных ситуаций. Например, требование рентабельности предполагает, что в некоторых ситуациях затраты на создание или приобретение психометрической методики не оправдываются в связи с небольшой величиной группы потенциальных тестируемых, отсутствием риска при зачислении неподготовленных кандидатов на испытательный срок и т.д.

4.2.5. Использование терминов «психометрические требования» и «психометрические тесты» требует введения в данный документ нескольких оговорок:

психометрические требования одновременно являются требованиями к качеству методики, которые необходимо применять не только по отношению к тестам, направленным на измерение психических свойств, но и по отношению к тестам профессиональных знаний и умений;

в ряде случаев тесты профессиональных знаний и умений создаются или модифицируются внутренними специалистами организаций для решения ситуационных задач и в

короткие сроки, что не позволяет им осуществлять проверку психометрических свойств данных методик;

указанные узконаправленные методики обладают рядом достоинств, позволяя учитывать в содержании заданий специфические или уникальные условия конкретной организации, а также описывать узнаваемые для ее сотрудников проблемные ситуации;

стандарт допускает, что от использования непсихометрических тестов (квизитестов) может быть полезный эффект, и не стремится препятствовать их созданию и/или применению;

при прочих равных условиях предпочтение следует отдавать тем тестам, которые в большей степени удовлетворяют психометрическим требованиям.

### **4.3. Вероятностный смысл результатов и погрешность измерения**

Тестовые измерения дают результаты с вероятностной точностью: истинный результат находится не в точке, а в определенном интервале на шкале, и содержит в себе заранее заданную степень погрешности (ошибку измерения). Перед проведением тестирования пользователь должен определиться с пониманием того, каким диапазоном точности (степенью погрешности) будет обладать полученная информация во избежание неправильного или излишне прямолинейного толкования результатов тестирования.

### **4.4. Требования к надежности**

Показателем степени погрешности измерения является надежность теста: чем выше надежность, тем ниже погрешность (меньше ошибка измерения). Формула расчёта ошибки измерения представлена в Приложении 3.

4.4.1. В прилагаемой разработчиком методической (технической) документации к методике тестирования должна быть приведена информация об исследованных видах надежности, полученных результатах и их следствиях для интерпретации тестовых баллов. Крайне желательна проверка надежности на уровне каждого отдельного тестового задания, то есть наличие в методической документации результатов анализа «коэффициентов дискриминативности» (различающей способности) для каждого задания. В отсутствие высокой дискриминативности отдельных заданий не могут быть достигнуты высокая надежность и валидность целого теста.

4.4.2. Минимально необходимая информация о надежности представляет собой данные о внутренней согласованности теста и отдельных тестовых шкал (коэффициент «альфа Кронбаха» или его эквиваленты). Для методик личностной диагностики рекомендуемый показатель надежности не должен быть ниже 0,6, для тестов способностей — 0,7, а для тестов профессиональных знаний и умений должен принимать значение в диапазоне 0,8—0,9. Надежность может превышать показатель 0,9 в случае использования узкоспециальных методик. Неправдоподобно высокие показатели надежности могут говорить о нарушениях в процедуре разработки (например, о слишком высокой однородности тестовых заданий, которые ограничивают область применения методики).

4.4.3. Для некоторых типов тестовых методик считаются приемлемыми более низкие показатели надежности. В таких случаях разработчик обязан представить обоснование либо с позиций результатов предшествующих научных исследований, либо с позиций практической необходимости, причем наличие такого обоснования не отменяет факт увеличения погрешности в результатах измерения.

4.4.4. Разработчик обязан указывать размер выборки, на которой исследовалась надежность. В общем случае при разработке тестовой методики размер выборки для исследова-

ния одного вида надежности не должен составлять меньше 100 человек. При этом чем более разнородна выборка (то есть чем больше возрастные, профессиональные и другие подгруппы различаются по измеряемой характеристике), тем большего размера она должна быть.

4.4.5. Если в тестовой методике используются параллельные формы (версии, одинаковые по структуре, но включающие разные по содержанию тестовые задания), разработчиком должны быть представлены количественные подтверждения эквивалентности этих форм.

4.4.6. При разработке тестовой методики крайне желательна также проверка ретестовой надежности: тот же самый тест проводится на той же самой выборке повторно через заданный промежуток времени, и исследуется корреляция (взаимосвязь) между результатами этих двух серий тестирования. В таком случае минимально приемлемой границей надежности является коэффициент 0.5. Близкие к границе 0.5 показатели допустимы для методик, оценивающих менее устойчивые качества, и при наличии большого интервала времени между двумя тестированиями (6 месяцев и более). Для менее длительных интервалов времени (в пределах 2—4 недель) и для более устойчивых качеств (например, для тестов способностей) рекомендуемый уровень ретестовой надежности составляет 0.6—0.7.

4.4.7. Если при разработке методики тестирования разработчиком используются психометрические подходы, альтернативные классической теории тестов (например, IRT), параметры надежности должны быть «переведены» на классический язык (например, должна быть оценена величина, эквивалентная надежности, или стандартная ошибка измерения, соответствующая параметру информативности теста в IRT). Так обеспечивается возможность ориентировки в уровне надежности измерения для широкого круга пользователей. Следствием этого положения является требование к сопутствующей технической (методической) документации, в которую разработчиком тестов рекомендуется либо включать информацию о психометрических свойствах разработанного теста на языке классических показателей, либо ограничивать круг пользователей данной методики теми лицами, которые готовы воспринимать термины и процедуры современной (неклассической) теории тестов.

#### 4.5. Требования к валидности

О валидной методике можно говорить, если методика измеряет именно то, для чего она предназначена. Валидность, как и надежность, никогда не достигает 100 процентов, поэтому можно говорить о более валидной или менее валидной методике. Некоторые надежные методики оказываются невалидными, если они применяются для тех целей, для которых они не предназначены (не пригодны): например, надежная методика на скорость устного счета не является валидным инструментом для измерения креативности (изобретательности).

4.5.1. В прилагаемой разработчиком методической (технической) документации к методике тестирования должна быть приведена информация об исследованных разработчиком видах валидности. Подтверждения валидности могут быть более или менее полными и убедительными, подразделяясь на три группы, которые обозначаются как содержательная, конструктивная и критериальная валидность.

4.5.2. **Содержательная валидность** в плане обоснования должна содержать отсылки на процедуру разработки и отвечать на следующие вопросы:

Какой конструкт (тестируемое свойство) лежит в основе методики тестирования?

Как это свойство концептуально соотносится с другими похожими свойствами?

Какова структура методики?

Какие компоненты входят в конструкт (какие субшкалы входят в шкалу измеряемого свойства)?

Почему были выделены именно эти компоненты (субшкалы)?

Какова содержательная область методики?

Какие виды деятельности тестируемого покрываются содержанием заданий тестовой методики?

Насколько полно задания методики отражают целевую содержательную область?

Существуют количественные процедуры, позволяющие измерить содержательную валидность путем применения коэффициента согласованности экспертных оценок, но чаще всего содержательная валидность обосновывается с помощью качественных характеристик.

4.5.3. Статистические доказательства, характеризующие **конструктивную валидность**, следует искать в результатах специально организованных исследований. Ими могут выступать: взаимосвязь результатов методики тестирования с результатами других тестовых методик (содержательно близких или, наоборот, отличных); внутренняя структура методики (статистическая структура взаимосвязей тестовых заданий между собой, коэффициенты интеркорреляций между шкалами); результаты исследований, проведенных с применением данной методики тестирования (эксперименты, результаты сравнения групп и т.д.).

4.5.4. При исследовании и описании указанных взаимосвязей необходимо соблюдение следующих принципов. Выборка для исследования валидности включает в себя не менее 100 человек (либо достаточность размера выборки для получения надежных результатов обосновывается с помощью специальных статистических процедур, таких как вычисление статистической мощности). Полученные коэффициенты корреляции (либо эквивалентные статистические показатели) при измерении конструктивной валидности должны быть сопоставлены с изначальными теоретическими ожиданиями, при этом корреляции 0.1—0.3 расцениваются как «слабые», 0.3—0.5 как «умеренные», 0.5—0.7 как «сильные», выше 0.7 — как «очень сильные» или даже «настораживающие». Очень высокие значения показателя валидности могут говорить о методических ошибках при ее проверке или о недобросовестности разработчика<sup>16</sup>. Если между результатами исследования и исходными теоретическими ожиданиями обнаружены расхождения, им должно быть дано содержательное объяснение, а также должны быть описаны следствия этих расхождений для интерпретации результатов методики в следующих случаях:

- если взаимосвязь теоретически предполагалась, но оказалась «слабой»;
- если взаимосвязь не предполагалась, но оказалась выше 0.3;
- если взаимосвязь оказалась выше 0.7 (в этом случае должно быть дано обоснование преимуществ разрабатываемой методики тестирования по отношению к уже существующей, с которой обнаруживается высокая корреляция).

4.5.5. В методической (технической) документации к тестовой методике должно содержаться описание соответствующих исследований (если они были проведены), а также должно быть описано, какие из наблюдаемых взаимосвязей с результатами других методик соответствуют теоретическим ожиданиям, а какие — нет. На основании этих исследований должно осуществляться взвешенное заключение об ограничениях в интерпретации тестового балла или тестового профиля.

4.5.6. **Критериальная, или прагматическая, валидность** методики тестирования представляет собой взаимосвязь ее результатов с внешними объективными критериями (показателями эффективности деятельности или КПЭ), в том числе отсроченными

16 Речь идет о наиболее общей интерпретации величин коэффициента корреляции. Для некоторых видов валидности, а также для некоторых классов тестовых методик и сфер их применения действуют иные нормативные значения: например, для тестов профориентации, на основе результатов которых строится долгосрочный прогноз в рамках широких областей деятельности (отрасли или профессии) «сильными» следует считать показатели валидности выше 0.4, а «очень сильными» — выше 0.5.

по времени. Для методик тестирования, используемых в оценке персонала, данный вид валидности признается наиболее важным.

4.5.7. Внешний критерий должен быть максимально объективным и максимально связанным с предполагаемыми целями тестирования. Например, в качестве критерия может выступать «процент годных изделий», «объем продаж», «уровень заработной платы при сделанных расценках», «количество новых клиентов (сделок)», «количество патентов» и другие показатели эффективности профессиональной деятельности. Ожидаемые взаимосвязи между критерием и результатами теста должны быть также обоснованы в методической (технической) документации.

4.5.8. Критерий должен быть независимым от процедуры проведения тестирования и от способа формирования выборки тестируемых. Использование критериев, не являющихся независимыми, приводит к ошибочному завышению показателя валидности. Например, если в качестве критерия выступает экспертная оценка сотрудника руководителем, и при этом экспертная оценка проводится по бланку, аналогичному тому, который получали сами сотрудники для самооценки, только переформулированному из первого лица в третье.

Наоборот, сужение выборки по критерию приводит к ошибочному занижению показателя валидности. Например, если был объявлен добровольный принцип участия в тестировании и согласились на него только успешные сотрудники, а низкоэффективные отказались принять в нем участие. В этом случае корреляция результатов теста с параметрами успешности в исследуемой группе окажется искусственно заниженной. Ещё одним случаем, когда показатель валидности ошибочно занижается, является случай, когда тестирование проходят кандидаты на позицию, а критериальные данные собираются только по успешно прошедшим тестирование и принятым на работу сотрудникам.

4.5.9. Необходимая для рентабельного профотбора величина взаимосвязи между методикой и внешним критерием определяется либо экспертным путем, либо с помощью расчетных формул. Для корректного эмпирического измерения критериальной (прагматической) валидности требуется обследовать корпоративную нормативную выборку объемом не менее 300 человек. При этом корреляции для одношкальных тестов в размере 0.15—0.25 расцениваются как «слабые», 0.25—0.35 как «умеренные», 0.35—0.5 как «сильные», выше 0.5 — как «очень сильные» и «настораживающие». Корреляции выше 0.5 чаще всего являются результатом расчета множественной корреляции для многошкальных (многофакторных) тестов. При высокой массовости применение тестов с валидностью 0.15 часто оказывается оправданным (рентабельным), так как применять более глубокие методы оценки в этих условиях экономически нецелесообразно.

При оценке и интерпретации величины корреляции между методикой и внешним критерием следует опираться на сравнение с данными опубликованных мета-анализов (усредненных результатов нескольких независимых исследований) о взаимосвязях результатов родственных тестовых методик с объективными критериями эффективности профессиональной деятельности.

4.5.10. Следует различать ситуации, в которых внешний критерий измерялся одновременно с проведением тестирования, и ситуации, в которых он измерялся позже (например, через несколько лет). В первом случае результаты исследования не могут являться подтверждением того, что методика тестирования способна спрогнозировать эффективность будущей деятельности и измерить профессиональный потенциал сотрудника, а во втором случае подобная интерпретация возможна. При этом существуют ситуации тестирования, в которых прогноз эффективности профессиональной деятельности не требуется, и большее значение имеет оценка достигнутого уровня (например, при аттестации или оценке профессионального соответствия).

4.5.11. Данные о критериальной валидности тестовой методики часто не могут быть предоставлены разработчиком теста в документации сразу, но могут быть получены с участием пользователей методики только спустя некоторое время после ее фактического применения (см. раздел 9.4). При наличии организационных возможностей пользователям методики рекомендуется продумывать программы и технологии сбора данных совместно с разработчиком, что позволяет производить отсроченную проверку критериальной валидности теста.

4.5.12. Особенно ценные данные о критериальной валидности дает схема проверки «прогностической валидности»: сначала проводится тест, а критериальная информация о протестированных (например, о КПЭ) собирается спустя несколько месяцев или даже лет. Реализация такой схемы в серьезных статистических масштабах оказывается невозможной без тесного сотрудничества разработчиков и пользователей теста.

#### **4.6. Требования к тестовым нормам и репрезентативности**

4.6.1. В методической (технической) документации к тесту должны быть указаны тестовые нормы. *Тестовые нормы* рассматриваются как критические точки или интервалы на шкале тестовых баллов, которые разделяют статистически различные группы тестируемых (то есть группы, к которым целесообразно применять различные управленческие решения). В документации к тесту должно содержаться подробное описание выборки нормирования, с результатами которой будут сопоставляться результаты участника тестирования.

4.6.2. *Выборка нормирования* (или выборка стандартизации) должна быть представительной (репрезентативной) по отношению к множеству потенциальных участников тестирования (то есть соответствовать основным существенным особенностям контингента участников по демографическим параметрам: по полу, возрасту, образованию, принадлежности к профессиональной или языковой группе, государственно-национальной принадлежности, уровню должности или социо-экономическому статусу).

4.6.3. Репрезентативность выборки нормирования обосновывается как концептуально (через описание состава этой выборки с точки зрения основных демографических параметров), так и эмпирически. Эмпирические процедуры проверки репрезентативности могут включать в себя исследование устойчивости тестовых норм при расщеплении выборки на две независимые половины и последующее сравнение норм для этих двух половин.

4.6.4. Необходимый размер выборки нормирования для стандартизованных методик тестирования зависит от степени однородности предполагаемой группы участников тестирования. Чем более разнородны участники тестирования, тем больше необходима выборка нормирования<sup>17</sup>. В общем случае минимально приемлемая выборка нормирования состоит из 200 человек, а рекомендуемый размер выборки нормирования для универсальных тестов, претендующих на возможность применения в различных организациях и в различных отраслях, составляет 500 человек.

4.6.5. Для тестов, ориентированных на локальное применение в конкретной организации, выборка нормирования может быть ограничена ее сотрудниками. В этом случае участники тестирования сравниваются внутри своей группы, что определяет специфику интерпретации результатов. Результаты подобной локальной методики при ее использовании в другой организации не могут быть признаны без проведения дополнительных исследований.

4.6.6. Тестовые нормы должны периодически обновляться и уточняться как разработчиками тестов, так и пользователями тестов. Рекомендуемая частота обновления тестовых норм составляет один раз в три года.

<sup>17</sup> Например, если рассматривать группы менеджеров по работе с клиентами и диспетчеров метро, становится очевидно, что вторая группа является заведомо более однородной. Поэтому для менеджеров по работе с клиентами требуется более объемная выборка нормирования.



#### 4.7. Требования к достоверности

Достоверность является требованием, которое в большей степени характеризует не тест, а процедуру его применения. Нарушения в процедуре тестирования не могут непосредственно контролироваться разработчиком, поэтому соблюдение требования достоверности в основном обеспечивается пользователем методики. Но разработчик должен предупредить пользователя, в каких ситуациях возникают искажения, и как их обнаружить.

4.7.1. Разработчиком методики в методической (технической) документации должны быть описаны основные риски, присущие методике с точки зрения тактики выполнения теста участниками. Также должны быть описаны предпринятые способы снижения этих рисков.

4.7.2. К нарушениям достоверности тестирования в тестах-опросниках могут относиться мотивационные искажения (стремление выглядеть в лучшем свете, или социальная желательность, либо, напротив, стремление выглядеть в худшем свете — аггравация), установка на согласие, установка на нейтральный ответ, установка на крайние, а также случайные ответы.

4.7.3. К нарушениям достоверности тестирования в тестах знаний или способностей относится угадывание ответов, которое может провоцироваться структурой теста (например, правильные ответы содержат самые длинные формулировки). Но самый серьезный риск при тестировании знаний несет рассекречивание правильных ответов (использование шпаргалок).

4.7.4. Требования к достоверности методики тестирования становятся особенно актуальными в ситуации экспертизы, предполагающей принятие серьезных управленческих решений. В данной ситуации у участников тестирования появляются реальные мотивы к фальсификации результатов. Попытки фальсификации результатов в ситуации консультирования возникают реже в силу личной заинтересованности тестируемого в достоверности результатов.

4.7.5. Достоверность методики тестирования может учитываться, обеспечиваться и контролироваться с помощью:

— особых способов контроля, заложенных в структуру методики (например, сочетания «прямых» и «обратных» утверждений, шкал лжи, повторяющихся одинаковых вопросов, ипсативной технологии, специально нацеленной на то, чтобы смоделировать вынужденный выбор между суждениями, имеющими одинаковый балл на шкале социальной желательности);

— специальных исследований (например, сравнения с результатами тестирования, проведенного с инструкцией на намеренное искажение результатов);

— технических и статистических средств (например, видеонаблюдения, формирования случайной выборки заданий из широкого банка заданий, обнаружения статистически нетипичного паттерна ответов);

— сопоставления результатов тестирования с результатами, полученными в рамках иных оценочных процедур (например, интервью или деловых игр).

4.7.6. Помимо описания всех способов обеспечения и контроля достоверности методики тестирования, при чтении тестовой документации пользователю рекомендуется обращать внимание на возможные уровни ошибки (погрешности) измерения, указанные разработчиком как ожидаемые в определенных ситуациях тестирования в связи с нарушениями достоверности.

#### 4.8. Требования к лицензионной чистоте методик тестирования

Особым требованием к тестовым методикам выступает требование лицензионной чистоты. Пользователям тестов следует избегать применения: а) тестовых методик, авторы которых неизвестны; б) тестовых методик, размещенных на пиратском сайте в Интернете или опубликованных в нелегальном печатном или электронном сборнике тестов; в) тестовых методик, авторы которых не приводят в сопутствующей методической (технической) документации подтверждений осуществления проверки их психометрических свойств.

Нарушение прав интеллектуальной собственности при использовании тестовой методики может быть оспорено реальным правообладателем в суде. Без письменного запроса и письменного разрешения автора возможно использование методик, материалы которых опубликованы в научно-методической литературе с пометкой «методика для свободного распространения» или снабжены символом СС (creative commons). Наличие символа копирайта © в отсутствие символа СС следует трактовать как разрешение правообладателя на знакомство с методикой и наличие запрета на ее массовое использование. Тиражирование тестовых материалов с символом копирайта © возможно только с разрешения правообладателя. Тиражированием в данном случае называется изготовление десяти (и более) копий с материалами тестовой методики для тестирования десяти (и более) человек. При этом число 10 следует считать условным пороговым значением. Для судебного иска истцу достаточно предъявить в суд один незаконный (контрафактный) экземпляр методики, после чего ответчик должен предоставить для уничтожения все контрафактные экземпляры.

### Глава 5. Требования к квалификации исполнителей

В работе с тестовыми методиками можно выделить несколько ролей исполнителей. Роли, связанные с *разработкой тестовых методик*:

- разработчики тестовых методик;
- эксперты в предметной области, на оценку которой направлены тестовые методики;
- программисты-разработчики компьютерных программ тестирования.

Основные роли, обеспечивающие *проведение тестирования в организации*:

- пользователь, отвечающий за выбор тестовой методики, разработку программы тестирования, анализ полученных результатов и формирование итоговых отчетов с результатами, выводами и возможными рекомендациями по итогам тестирования;
- администратор тестирования, обеспечивающий его организацию и проведение;
- консультант-наставник, обучающий работе с тестовой методикой.

Возможные *вспомогательные роли* соисполнителей при массовом тестировании:

- менеджеры-организаторы, обеспечивающие оповещение и явку участников;
- системные администраторы, инженеры компьютерных сетей и баз данных;
- операторы-ассистенты в аудиториях;
- операторы по вводу данных с бланка в базы данных (при бланковом тестировании).

При наличии в организации собственного подразделения по оценке персонала значительная часть вышеперечисленных ролей закрепляется за сотрудниками этого подразделения, которое должно находиться в тесном функционально-проектном взаимодействии с другими подразделениями по управлению персоналом (по отбору и ротации, обучению и развитию, кадровому администрированию и др.).

Внедрение и эксплуатация системы тестирования выступает отдельным внутриорганизационным проектом, который требует адекватного подбора исполнителей, оптимального распределения ролей и тесного функционального взаимодействия сотрудников нескольких

подразделений. В том случае, когда функции оценки персонала выполняет подразделение по обучению персонала, необходимо еще более тщательно отслеживать соответствие квалификации и функциональной подготовки сотрудников данного подразделения требованиям настоящего стандарта.

### **5.1. Пользователи тестовых методик**

Пользователи тестовых методик отвечают за выбор методики тестирования, соответствующей целям и условиям оценки, за разработку программы тестирования и корректный анализ результатов. В случае, когда к тестированию привлекаются соисполнители разных профилей (например, линейные руководители, сотрудники IT-службы), пользователи методик часто оказываются в роли менеджеров-координаторов всего комплекса мероприятий, связанных с подготовкой, проведением и обработкой результатов тестирования.

#### **5.1.1. Пользователи тестовых методик должны:**

- обладать высшим образованием;
- обладать базовыми знаниями в области теории тестов и психодиагностики (в частности, знать классификации методов, понятия «шкала», «норма», «корреляция», «валидность», «надежность» в объеме настоящего Стандарта);
- уметь производить элементарные статистические расчеты (в частности, расчет среднего и стандартного отклонения, рангов, перцентилей, корреляций в электронных таблицах);
- уметь правильно подбирать тестовые методики под конкретную задачу, знать области применения и ограничения применяемой методики;
- знать правила применения и интерпретации конкретной методики, принципы формирования итогового результата;
- обладать навыками предоставления обратной связи (в частности, пройти соответствующее обучение);
- уметь пользоваться компьютерной программой тестирования (если это предполагает технология тестирования).

В случае, когда тестирование является частью комплексной оценочной процедуры, специалисты, отвечающие за методически грамотное внедрение и эксплуатацию системы тестирования в организации, могут выступать в роли с более широким кругом организационных прав — в роли специалистов по оценке персонала.

5.1.2. К пользователю некоторых психометрических тестов предъявляется требование наличия высшего психологического образования, а также рекомендуется выполнение программы обучающего сертифицирующего семинара в профильной организации под руководством дипломированных психологов<sup>18</sup>. Необходимость такой особой подготовки пользователей, как правило, должна быть зафиксирована в аннотации к методике.

В случае отсутствия в подразделении по работе с персоналом работников с высшим психологическим образованием в организации может быть закреплена роль администратора тестирования, а роль пользователя психометрических тестов может быть закреплена во внешней специализированной организации (например, в организации-подрядчике, оказывающей услуги по оценке персонала).

### **5.2. Администраторы сеансов тестирования**

5.2.1. В случаях, когда в организации пользователи методик и администраторы оказываются разными лицами, задачи администраторов сводятся к контролю проведения тести-

<sup>18</sup> Особенно это касается сложных интеллектуальных тестов и многофакторных личностных тестов, оперирующих сложными конструктами, то есть особыми научно-психологическими терминами, обозначающими диагностические шкалы и категории.

рования в отдельных аудиториях (или на рабочих местах). На практике администраторы сеансов тестирования могут называться либо менеджерами-организаторами, либо наблюдателями (супервизорами, инспекторами), либо операторами-ассистентами (см. главу 8), либо экзаменаторами (когда речь идет о тестовом экзамене). В некоторых случаях роли менеджеров-организаторов, супервизоров и операторов могут быть специфичными.

Особые требования к администраторам-организаторам возникают при одновременном проведении сеансов тестирования в различных подразделениях многофилиальных организаций при массовом дистанционном тестировании кандидатов в режиме онлайн и других случаях массового тестирования. Для этих случаев в организации необходим внутренний распорядительный документ «Регламент тестирования», в котором роли четко определены и распределены между исполнителями (см. подробнее раздел 6.2).

#### 5.2.2. Администраторы тестовых методик должны:

- знать правила администрирования конкретных тестовых методик (в частности, пройти инструктаж или соответствующее обучение с контролем усвоения);
- обладать необходимыми компетенциями в области личной эффективности (в частности, знать смежных исполнителей в своей организации, ориентироваться в системе коммуникации и схемах информационных потоков, принятых при проведении сеансов тестирования и обработке тестовых результатов);
- уметь устанавливать и эксплуатировать компьютерные программы тестирования и/или уметь проводить тестирование на бланках.

### 5.3. Консультанты-наставники или обучающие работе с тестовой методикой

Консультантами-наставниками чаще всего выступают внештатные специалисты (например, представители компаний-разработчиков тестовых методик), которые привлекаются для подготовки пользователей и администраторов, а также для помощи во внедрении и эксплуатации целостной тестовой системы.

Консультанты-наставники должны:

- обладать высшим образованием;
- обладать дополнительным образованием по работе с психометрическими методиками;
- обладать углубленными знаниями в сфере теории тестов, психодиагностики, статистики, психометрики, а также в предметной области, соответствующей применяемой тестовой методике;
- иметь стаж работы психологом или специалистом по работе с персоналом, опыт разработки и/или адаптации тестовых (оценочных) процедур;
- иметь опыт работы с конкретной методикой;
- обладать навыками индивидуального консультирования и групповой модерации.

### 5.4. Разработчики тестовых методик

Разработчики тестовых методик должны:

- обладать образованием не ниже степени специалиста по направлению «Психология»;
- обладать углубленными знаниями в сфере теории тестов, психодиагностики, статистики, психометрики, а также в предметной области, соответствующей разрабатываемой тестовой методике;
- иметь опыт разработки и адаптации тестовых (оценочных) процедур;
- иметь опыт статистического анализа многомерных данных.

Пользователям методик рекомендуется обращать внимание на такие признаки, подтверждающие квалификацию разработчика, как наличие публикаций, независимые

рецензии (отзывы других пользователей), наличие документов о сертификации разработанных методик.

### **5.5. Эксперты в предметной области**

Эксперты в предметной области, участвующие в разработке тестовых методик (в качестве соисполнителей разработчика), должны:

- обладать высшим или средним специальным профессиональным образованием, соответствующим предметной области разрабатываемой тестовой методики;
- обладать опытом работы в предметной области, соответствующей разрабатываемой тестовой методике;
- продемонстрировать способность к словесному описанию своей профессиональной деятельности и ключевых проблемных ситуаций в ней.

Пользователям методик рекомендуется обращать внимание на такие признаки, подтверждающие квалификацию экспертов, как научные публикации, наличие ученых степеней, членство в экспертных сообществах, преподавательский опыт. Конечным показателем квалификации эксперта как автора тестовых вопросов является высокое качество созданных им заданий, подтвержденное в психометрическом (статистическом) исследовании.

### **5.6. Программисты-разработчики компьютерных программ тестирования**

Программисты-разработчики компьютерных программ тестирования должны (подробнее о компьютерном тестировании см. главу 8):

- иметь высшее или среднее специальное образование по направлениям «Прикладная математика и информатика», «Программная инженерия», «Математика и компьютерные науки», знания и умения в области компьютерных технологий, связанных с установкой и эксплуатацией компьютерных диалоговых систем, а также баз данных;
- обладать базовыми знаниями в области теории тестов (пройти обучающий сертифицирующий семинар с контролем усвоения);
- иметь практический опыт установки и эксплуатации компьютерных систем, связанных с обработкой массивов персональных данных.

## **Глава 6. Проведение тестирования**

Целесообразно различать требования к проведению тестирования в очной и заочно-дистанционной форме, а также в ситуациях консультирования и экспертизы. Приведенный в данной главе перечень требований в большей степени касается очного административного тестирования в аудиториях и компьютерных классах, однако часть из них применима в том числе в ситуации консультирования. Некоторые параграфы главы 6 повторяют содержание главы 3, чтобы сориентировать организаторов тестирования в подготовке и проведении тестирования и помочь им не упустить из виду значимые нюансы и того, и другого этапа работы. В тех случаях, когда в организациях различаются роли организаторов подготовки системы тестирования и администраторов отдельных сеансов, первой группе исполнителей следует читать и главу 3, и главу 6, а второй — только главу 6. Ряд требований к ситуации адресного дистанционного онлайн-тестирования (см. об этом понятие главу 3) введены в главу 8<sup>19</sup>.

<sup>19</sup> Следует констатировать, что по поводу данной ситуации в профессиональном сообществе специалистов-тестологов в настоящее время не сложилось консенсуса: будет ли эта ситуация трактоваться по правилам «ситуации экспертизы» или по правилам «ситуации консультирования». Это зависит от поставленных задач и действий специалистов по оценке в рамках конкретной организации.

### **6.1. Регламент тестирования**

В ситуации экспертизы порядок проведения тестирования должен быть формализован с помощью особого документа — регламента, утвержденного руководителем организации, осуществляющей тестирование. В регламенте должны быть детально освещены обязательные действия должностных лиц, от которых зависит достоверность результатов тестирования, а также правила поведения тестируемых при выполнении тестов. Регламент описывает желательную схему взаимодействия исполнителей и передачи между ними информации. В случае компьютерного тестирования ключевая роль в реализации регламента может принадлежать IT-подразделению компании.

### **6.2. Предварительная запись на сеансы тестирования**

Если число тестируемых превышает возможности для одновременного проведения тестирования в одном помещении, целесообразно заранее проинформировать тестируемых о месте и времени проведения тестирования, чтобы исключить длительное ожидание в очереди, опоздания на сеанс и т.п., а также организовать процедуру записи на сеансы тестирования. Рекомендуемая продолжительность сеансов, на которую следует ориентироваться, в том числе при организации предварительной записи на них, не превышает 90 минут (см. параграф 3.6.3).

### **6.3. Предварительное информирование и консультация**

Перед тестированием необходимо провести предварительное информирование участников о целях тестирования (зачем оно проводится, и какие решения могут быть приняты по результатам), форме проведения тестирования, сроках проведения и обработки результатов, дать информацию о конфиденциальности результатов или о правах доступа к ним со стороны разных групп лиц (см. подробнее раздел 3.7)

### **6.4. Условия проведения тестирования**

6.4.1. *Условия проведения очного административного тестирования на бланках.* Помещение должно быть подготовлено к проведению тестирования (см. параграф 3.5.1). Желательным является обеспечение изолированного рабочего места для каждого из участников тестирования в виде отдельного стола. На каждом из столов должны находиться только идентичные тестовые материалы и канцелярские принадлежности. В некоторых случаях допускается использование стандартных справочных материалов, калькуляторов и других вспомогательных предметов, но использование электронных запоминающих устройств и мобильных средств связи в большинстве случаев запрещается. Об этих и иных особенностях процедуры (например, наличии в помещении системы видеонаблюдения) тестируемые должны быть оповещены заранее.

6.4.2. *Условия проведения очного административного тестирования на компьютерах.* В данном случае наибольшее внимание уделяется подготовке идентичных компьютеров или иных устройств, на базе которых проводится процедура тестирования и которые должны в равной степени удовлетворять требованиям программы тестирования. Использование разнотипных персональных компьютеров (например, ноутбуков, планшетов и смартфонов) может создавать отдельным участникам неконтролируемые со стороны организаторов преимущества. В ряде случаев необходимо предусмотреть обеспечение участников особыми техническими устройствами (например, наушниками при наличии в тесте мультимедийного стимульного материала). Более детальные требования к компьютерному тестированию описаны в главе 8.

6.4.3. *Списки тестируемых участников.* Администраторы сеансов тестирования должны получить в свое распоряжение результаты предварительной записи участников для контроля доступа участников к сеансам, а также обеспечить аутентификацию (проверку личности) тестируемых согласно предварительным спискам. Удобно совместить подобные списки с ведомостями, на которых заранее предусмотрены графы для вписывания и визирования результатов тестирования каждым из участников.

### **6.5. Правила поведения для участника тестирования**

Тестируемые должны быть ознакомлены с правилами поведения во время тестирования, представляющими собой документ с перечнем пунктов, отвечающих на следующие вопросы:

— как будет производиться проверка документов, удостоверяющих личность тестируемого (аутентификация), и какой документ должен предъявить тестируемый для допуска к тестированию?;

— разрешенные и запрещенные предметы при проведении тестирования (например, невозможность использования мобильных средств связи и разрешение на использование калькулятора);

— недопустимые действия в аудитории (например, переговоры с соседями, перезапуск программы тестирования на компьютере, фотографирование тестовых заданий, сознательная порча бланков);

— санкции за нарушения правил поведения тестируемого (в каких случаях делается предупреждение, а в каких тестируемый отстраняется от дальнейшего участия в тестировании?);

— действия тестируемого в случае технических сбоев (например, обнаружение неполноты тестовых материалов, опечаток, программных сбоев во время компьютерного тестирования);

— действия тестируемого при обнаружении некорректных, по его мнению, тестовых заданий (например, фотографировать спорное задание, чтобы не тратить время на его разбор во время тестирования, оповестить о правилах апелляции по содержанию процедуры после теста);

— действия тестируемого при досрочном и штатном завершении сеанса тестирования; необходимость визирования тестируемым своего протокола или предварительных результатов тестирования;

— разрешенные для выноса из аудитории материалы тестирования (например, невозможность выноса тестовых буклетов).

### **6.6. Информированное согласие**

В начале сеанса в ряде случаев целесообразно получить от тестируемого подписанное «Информированное согласие» (см. подробнее об этом документе в главе 11), подтверждающее факт ознакомления участника тестирования с его правилами, готовность их соблюдать, а также являющееся выражением согласия на обработку персональных данных.

### **6.7. Действия и обязанности администратора тестирования**

Администратор сеансов осуществляет контроль хода тестирования. В ходе сеанса тестирования администратор (наблюдатель-супервизор) обязан следить за поведением тестируемых, чтобы исключить действия, которые могут снизить достоверность результатов. Все нарушения должны фиксироваться в особом протоколе, позволяющем аннулировать недостоверные результаты тестируемого, нарушившего правила.

Администратор тестирования не имеет права подсказывать участникам способы выполнения теста; в случае обращения участников за разъяснениями администратор обязан перефразировать те или иные пункты стандартной инструкции, не выходя за ее пределы, а также оказывать помощь в преодолении технических затруднений (например, правила заполнения бланков, разрешенный возврат к предыдущим заданиям теста, повторное обращение к письменной инструкции и т.п.).

### **6.8. Права администратора (наблюдателя)**

Согласно регламенту, администратор наделяется правом осуществления оперативного вмешательства в ход тестирования, вынесения предупреждений и отстранения тестируемых от прохождения теста в случае неоднократного нарушения ими правил или при возникновении критических ситуаций. При отстранении участников от тестирования необходимо оповещать их об учете их прав, а также о возможности выполнения теста в дополнительные дни, если пропуск основного сеанса обусловлен уважительной причиной. Администратор должен быть готов к стрессовым реакциям со стороны отдельных участников тестирования, стараться сохранять хладнокровие и вежливость в случае конфликтных реакций. Обсуждение претензий участников предпочтительно производить после завершения сеанса тестирования.

### **6.9. Проведение дистанционного онлайн-тестирования**

Главная сложность при проведении дистанционного онлайн-тестирования состоит в надежной аутентификации участников (установлении личности), а также в контроле числа запусков сеансов тестирования. Эти вопросы подробно освещаются в главе 8 и разделе 3.1., где дается классификация различных типов контроля в данной ситуации.

### **6.10. Завершение сеанса тестирования**

По завершении каждого сеанса тестирования или сессии, состоящей из нескольких сеансов, важно произвести комплекс завершающих мероприятий. При массовом тестировании целесообразно описать в регламенте этот комплекс мероприятий, среди которых наиболее распространенными являются следующие.

6.10.1. При проведении тестирования в ситуации экспертизы, не предусматривающей анонимность участников, каждый из них должен визировать заполненный бланк подписью или вносить в соответствующую графу персональный шифрокод. Данная процедура обеспечивает объективность в случае подачи апелляции и более востребована на первых этапах внедрения системы тестирования в организации, но не является обязательной, если система хорошо отлажена и доверие участников к ней высоко. При проведении компьютерного тестирования аналогом визирования является подпись протестированного участника под предварительным результатом в специальной ведомости.

6.10.2. По завершении бланкового тестирования заполненные бланки должны быть запечатаны в конверты; по каждому конверту должна быть составлена опись с подписью лица, отвечающего за проведение тестирования в данном помещении (администратора сеанса). Такое оформление результатов необходимо в случае обработки результатов тестирования иными лицами.

6.10.3. Многоцветные тестовые материалы (буклеты) должны быть собраны в соответствующие пакеты для обеспечения их секретности. Недопустимо оставлять их в распоряжении тестируемых.



6.10.4. Одноразовые вспомогательные материалы (например, черновики) должны быть либо уничтожены, либо заархивированы в персональных папках с грифом, ограничивающим доступ к ним.

6.10.5. Компьютерные программы тестирования, если они размещаются на локальных неспециализированных компьютерах, по завершении сеанса или сессии, состоящей из нескольких сеансов, должны быть удалены с этих компьютеров. Доступ к онлайн-системам тестирования после завершения должен быть приостановлен.

### **6.11. Облегченные требования к консультативному тестированию**

В ситуации консультирования главным требованием к организации тестирования является стандартность условий его проведения для всех участников, то есть соблюдение заданной процедуры. В силу доминирования у тестируемых мотивации самопознания можно в меньшей степени опасаться утечки информации по ключам, преднамеренных фальсификаций и иных источников снижения достоверности результатов.

### **6.12. Развивающее испытание, но не прямое обучение**

Любое тестовое испытание (как и любое экзаменационное испытание) несет в себе для человека определенный опыт, является тренировкой познавательных и волевых качеств и тем самым имеет развивающий эффект. Но это другой вид воздействия на человека, чем прямое обучающее воздействие. При проведении тестирования администратор не имеет права помогать тестируемому приобретать новые знания в ходе тестирования, задавать тестируемому вспомогательные «наводящие» вопросы, выражать поощрение действий тестируемого, упрощать смысл инструкции или, напротив, затруднять правильное понимание легких заданий, чтобы «оптимизировать» трудность упражнения. Администратор не имеет права превращать анализ результатов тестирования в обучающее мероприятие. Разбор правильных или ошибочных ответов по результатам тестирования нельзя ни обещать участникам, ни планировать; ключевые (правильные) ответы должны оставаться в секрете.

Тестовые материалы и процедуры могут применяться в целях обучения и развития сотрудников, однако данные мероприятия должны быть заявлены соответствующим образом и преследовать релевантные цели. Цель тестирования — это измерение показателей, а не улучшение их в ходе измерения. Интерпретация результатов, как правило, применяется в отношении стандартизированных баллов, поэтому различие сырых и стандартизированных баллов является одним из важных условий квалифицированного использования тестовых методик в оценке персонала.

## **Глава 7. Обработка, интерпретация и анализ результатов**

Интерпретация результатов, как правило, применяется в отношении стандартизированных баллов, поэтому различие сырых и стандартизированных баллов является одним из важных условий квалифицированного использования тестовых методик в оценке персонала.

### **7.1. «Сырые баллы»**

«Сырые баллы» — это первичные показатели, вычисляемые согласно правилам, указанным разработчиками тестовой методики.

7.1.1. Простейшим случаем сырого балла является количество правильных ответов в тесте способностей или тесте знаний. В этом случае понятие «ключевой ответ» совпадает с «правильным ответом». Если тестовая методика содержит несколько шкал, то сырой балл рассчитывается для каждой из шкал по специальным правилам («подсчет по ключам»).

7.1.2. В тест-опросниках ключевыми называются ответы, соответствующие максимальной выраженности тестируемого показателя. При этом чем больше его выраженность, тем выше балл.

7.1.3. Выполнение компьютерных тестов, как правило, предполагает автоматический подсчет сырого балла. В этом случае ключи к тесту в большинстве случаев неизвестны не только тестируемым, но и пользователям методики.

7.1.4. Для методик, предполагающих проведение очного бланкового тестирования и последующий ручной подсчет по ключам, разработчики обязаны указать правила расчёта сырых баллов; при отсутствии ключей бланковый тест не может использоваться в практике тестирования персонала.

## 7.2. Стандартизация тестовой шкалы

Пересчёт сырых баллов в стандартизированные баллы позволяет сравнить результаты данного тестируемого с результатами других тестируемых, а также результаты данного тестируемого по разным тестовым шкалам.

7.2.1. Пересчет сырых баллов в проценты к максимальному возможному баллу *не является стандартизацией*, хотя и приводит к унифицированному масштабированию шкал, так как один и тот же результат (например, 70%) в данном случае может означать как высокий, так и низкий результат для различных по трудности тестов.

7.2.2. Пересчёт сырых баллов в стандартизированные производится путём сравнения сырых баллов респондента с показателями нормативной группы (см. раздел 4.6). Если разработчиками теста предлагаются для сравнения несколько нормативных групп, для определения стандартизированного балла конкретного тестируемого необходимо выбрать нормативную группу, соответствующую его демографическим характеристикам и цели данного тестирования.

7.2.3. Для методик, предполагающих проведение очного бланкового тестирования с последующей самостоятельной обработкой результатов, разработчики обязаны указать:

- рекомендуемую стандартную шкалу;
- среднее значение сырого балла каждой шкалы по каждой из нормативных групп,
- стандартное отклонение каждой шкалы для каждой из нормативных групп;
- формулу или таблицу перевода сырых баллов в стандартизированные (см. приложение 3).

7.2.4. Компьютерные тестовые программы в ряде случаев автоматически рассчитывают стандартизированный балл по всей группе лиц, протестированных в данной организации. Если при этом выявляются значимые различия в показателях средних значений для нормативной выборки и обследованной группы, то применять стандартную интерпретацию, предложенную разработчиком в руководстве, следует с осторожностью: возможно, более правильным для организационных целей является использование новых тестовых норм или их корректировка (см. подробнее раздел 7.5).

## 7.3. Стандартные шкалы

Двумя наиболее часто используемыми стандартными шкалами являются стэны (от англ. StandardTen — стандартная десятка) и Т-баллы. В тестах интеллекта часто используют

IQ-шкалу. Реже в практике оценки персонала используются стэнайны (от англ. StandardNine — стандартная девятка), тысячебалльная Т-шкала и Z-шкала. Вместе с основными категориями результатов данные шкалы представлены в Таблице 7.1.

Таблица 7.1. Основные стандартные шкалы и основные категории результатов по ним

Шкала	Среднее	Стандартное откл.	Могут принимать значения	Результаты				
				Низкие	Ниже среднего	Средние	Выше среднего	Высокие
Стэны	5,5	2	1 — 10	1	2—3	4 — 7	8 — 9	10
Стэнайны	5	2	1 — 9	1	2—3	4—6	7—8	9
Т-баллы	50	10	0 — 100	0 — 29	30 — 39	40 — 60	61 — 70	71 — 100
1000-балльная Т-шкала	500	100	0 — 1000	0 — 299	300 — 399	400 — 600	601 — 700	701 — 1000
IQ-шкала	100	15	больше 0	40 — 70	71 — 85	86 — 115	116 — 130	131 — 160
Z-баллы	0	1	любые	Меньше (-2)	(-2) — (-1)	(-1) — (+1)	(+1) — (+2)	Больше (+2)

#### 7.4. Процентили

7.4.1. Использование стандартных шкал основывается на предположении, что оцениваемый признак (например, способности к работе с различной информацией, уровень знаний, личностная черта) распределен в соответствии с законом нормального распределения. В тех случаях, когда разработчики теста предполагают, что оцениваемый признак распределен в выборке нормирования иначе, может использоваться шкала процентилей.

Процентиль<sup>20</sup> — это процент протестированных респондентов из нормативной группы, показавших результат ниже или такой же по сравнению с данным респондентом. Использование процентилей влечёт за собой определённые ограничения в интерпретации и анализе результатов. Например, вычисление среднего арифметического процентиля по нескольким шкалам не является корректной процедурой, в то время как допустимо вычисление среднего арифметического стандартного балла по нескольким шкалам, когда показатели по всем шкалам измерены в одноименных стандартных баллах. Процентильные границы часто используются для определения «крайних групп» на шкале тестовых баллов — «низкой» и «высокой» (по 25% или 33% от нормативной выборки).

7.4.2. В большинстве случаев шкала процентилей имеет вспомогательное значение и служит промежуточным звеном для перевода сырых баллов в стандартизированные в соответствии с так называемыми конверсионными таблицами.

7.4.3. Шкала процентилей часто оказывается наиболее понятной для тестируемого по сравнению с другими стандартными шкалами и в некоторых случаях рекомендуется для предоставления обратной связи (в частности, по тестам способностей).

#### 7.5. Использование локальных нормативных групп

7.5.1. В практике тестирования персонала в крупных организациях периодически используются так называемые корпоративные нормативные группы, состоящие только из сотрудников данной организации. Корпоративная нормативная группа должна включать не менее

20 Иногда в тестологической литературе вместо термина «процентиль» используется термин «процентаж».

100 сотрудников (оптимальное количество — 300 и более). Использование корпоративных нормативных групп величиной менее 100 человек признается нежелательным в силу отсутствия возможности обеспечить статистически достоверные результаты и выводы.

7.5.2. *Преимуществом* использования корпоративной нормативной группы является бóльшая простота и, во многих случаях, точность принятия решений по итогам тестирования (например, кандидат на позицию сравнивается с результатами сотрудников, уже работающих на этой позиции).

7.5.3. *Ограничением* использования корпоративной нормативной группы является риск неверной интерпретации абсолютного уровня результатов<sup>21</sup>. Переход на корпоративные и иные локальные нормы в настоящем стандарте рекомендуется производить не автоматически, но путем их анализа и сравнения с нормами, которые получены разработчиком теста на нормативной выборке и сообщаются в методической (технической) документации.

## 7.6. Погрешность измерения

Точность выводов при тестировании ограничена степенью погрешности (или ошибкой измерения), которая должна учитываться пользователем методики (см. Приложение 3). В зависимости от качества методики она может варьироваться от минимальной до достаточно серьёзной (например, более одного балла по шкале стэнов) и оказывать существенное воздействие на правомерность тех или иных решений, принимаемых на основе результатов тестирования<sup>22</sup>.

## 7.7. Интерпретация результатов тестирования

Интерпретация предполагает построение содержательных выводов на базе стандартизированных баллов по отдельным шкалам и их сочетаниям.

7.7.1. В практике тестирования персонала существует два типа выводов, относящихся к двум основным этапам интерпретации. Первый тип: оценка степени выраженности измеряемой характеристики у респондента (например, «обладает более высокими способностями к работе с числовой информацией, чем 8% менеджеров аналогичного уровня»; «склонен брать на себя лидерские функции в той же мере, что и большинство других людей»; «усвоил 90% знаний по вопросам, связанным с проведением гидроразрыва пласта»). Второй тип: прогноз рабочего поведения респондента на основе степени выраженности измеряемой характеристики (например, «может допускать серьёзные ошибки при работе с числовой информацией»; «будет эпизодически брать на себя лидерские функции в знакомой обстановке и при отсутствии других претендентов»; «будет успешно моделировать проведение гидроразрыва пласта за исключением особо сложных случаев»).

Выводы первого типа базируются на знании пользователем теста свойств используемых стандартных шкал и прямо следуют из результатов теста с поправкой на возможную погрешность измерения. Выводы второго типа требуют аналитической работы со стороны пользователя теста и могут опираться либо на его личный опыт и квалификацию, либо на шаблоны интерпретации, предоставленные разработчиками теста, либо на статистические данные о том, как справляются с работой обладатели определенного тестового балла. Наличие развёр-

21 Например, сотрудник, показавший семь стэнов по шкале «инициативность» на корпоративной нормативной выборке, состоящей из сотрудников компании с ярко выраженным командным стилем управления (стилем, подавляющим инициативу), на самом деле может быть не столь инициативным (по сравнению с нормой для всего региона, рынка, отрасли и т.п.), и его стандартизированный балл на самом деле оказывается завышен из-за того, что занижена база сравнения.

22 Например, если порог отсева по тесту профессиональных знаний и умений устанавливается в точке 80%, а погрешность измерения составляет 5%, то неправильно снимать с дальнейшего конкурса тестируемых с показателями в 76%, так как они также попадают в доверительный интервал. В этом случае реальный порог отсева следует понизить до 75%, а в тесте с погрешностью в 10% — до 70%. По этим же причинам практически не бывает тестов с порогом отбора в 100% правильных ответов, так как погрешность измерения в 5% позволяет считать фактически равными результаты на уровне 96% и 100% правильных ответов.

нутых шаблонов интерпретации позволяет обеспечить унификацию выводов всех пользователей теста, сократить число неверных выводов и ошибок в интерпретациях.

При принятии решений на базе результатов тестирования необходимо чётко разделять выводы двух указанных типов, а для выводов второго типа четко разделять:

- выводы, которые основаны на личном суждении пользователя;
- выводы, которые основаны на рекомендациях разработчиков теста;
- выводы, которые базируются на исследованиях и реальных производственных показателях.

7.7.2. Качество интерпретации существенно повышается при прохождении пользователями сертификационного обучения на право использования методики, проводимого разработчиком теста. Некоторые наиболее сложные тесты способностей и многофакторные личностные тест-опросники не могут использоваться без прохождения такого обучения.

7.7.3. Выводы по большинству тестов выносятся на базе стандартизированных, а не сырых баллов. Исключением могут служить тесты знаний и умений, для которых указаны пороговые значения в абсолютных показателях — в сырых баллах, масштабированных к 100 процентам (подробнее о критериально-ориентированном тестировании см. в разделе 2.3).

## **7.8. Различные подходы к анализу результатов тестирования**

Анализ результатов тестирования предполагает построение содержательных выводов и сравнение тестируемых между собой. Как правило, основной целью такого анализа является составление рекомендаций на основе всей совокупности имеющейся информации.

7.8.1. Существует несколько основных подходов к анализу результатов тестирования:

- качественный подход, при котором результаты тестируемых сравниваются между собой без применения каких-либо статистических методов;
- расчёт интегрального рейтинг-балла (по одной или нескольким методикам) для каждого тестируемого и ранжирование всех тестируемых в соответствии с их рейтинг-баллами;
- разделение тестируемых на группы в соответствии с результатами тестирования (по одной или нескольким методикам);
- сравнение профиля тестируемого с предварительно составленным «идеальным профилем».

7.8.2. Вне зависимости от применяемого подхода точность и качество анализа могут существенно повышаться, если наряду с результатами тестирования учитывается информация о тестируемом, полученная из других источников. В частности, результаты оценки сотрудников рекомендуется дополнять объективными данными о текущей эффективности их работы, экспертными (субъективными) оценками со стороны руководителей, других сотрудников, специалистов по оценке и т.п. (см. главу 9).

## **7.9. Расчёт интегрального рейтинг-балла**

Для принятия решений по итогам тестирования часто вычисляют интегральный рейтинг-балл на основе результатов тестируемого по различным тестовым методикам и шкалам внутри одной методики.

7.9.1. Интегральный рейтинг-балл представляет собой сумму результатов тестируемого по разным методикам, выраженным в одной и той же стандартизированной шкале, или усредненный стандартизированный балл, что облегчает его интерпретацию. Результатам тестирования по разным методикам могут присваиваться разные веса, которые зависят от: а) цели тестирования; б) специфики профессиональной деятельности тестируемого; в) производ-

ственных задач или обобщенных должностных функций тестируемого; г) типа выбранных для тестирования методик и их качества.

7.9.2. Расчёт интегрального (суммарного) рейтинг-балла является корректным только при использовании в качестве слагаемых стандартизированных баллов на линейных шкалах (стэны, Т-баллы, IQ-баллы, Z-баллы и т.д.). При этом важно, чтобы все показатели были измерены в одноименных шкалах. Расчет интегрального рейтинга путем суммирования и последующего усреднения некорректен при использовании процентов или процентилей (и тем более при использовании «процентovaných сырых баллов» — баллов, выраженных в процентах к максимальному баллу).

### **7.10. Разделение тестируемых на группы**

В случае использования процентилей наиболее корректным приемом является их перевод в оценочные категории (например, градации — А, В, С, D). Например, если D — это 25 процентов с самыми низкими баллами, С — это следующая четверть от нормативной выборки (от 26 до 50 перцентиля), то после перевода результатов нескольких методик в оценочные категории можно их скомбинировать (сочетание из двух категорий А и одной категории В может давать общую категорию А по трем тестам и т.д.). Указанным комбинациям показателей по нескольким методикам будут соответствовать различные интерпретации и рекомендации. Правила перевода первичных результатов в оценочные категории определяются в соответствии с целями и задачами тестирования и при участии экспертов. Более подробно о схемах сопряжения результатов различных методик см. раздел 9.2.

### **7.11. Сравнение с «идеальным профилем»**

Зачастую при анализе результатов тестирования используется так называемый «идеальный профиль», описывающий усредненные показатели эффективного работника для данной должности.

7.11.1. Идеальный профиль состоит из набора конкретных целевых показателей или областей целевых значений и может составляться несколькими способами: а) на основе результатов тестирования наиболее эффективных работников; б) на основе сравнения результатов тестирования наиболее и наименее эффективных работников; в) на основе экспертной оценки значимости различных характеристик.

7.11.2. Идеальный профиль может представлять собой:

- набор точечных целевых значений по различным шкалам и/или методикам;
- диапазоны целевых значений по различным шкалам и/или методикам;
- диапазоны недопустимых значений по различным шкалам и/или методикам.

7.11.3. В случае использования идеального профиля с ним сравниваются результаты каждого тестируемого. В зависимости от типа профиля могут использоваться как более простые способы определения сходства с ним (например, подсчёт количества попаданий в коридор из диапазонов допустимых значений), так и более точные статистические показатели. В последнем случае также могут использоваться весовые коэффициенты для различных шкал и/или методик.

## **Глава 8. Информационные технологии в тестировании**

Тестирование персонала может производиться как в бланковой форме, так и на компьютерах. Компьютерная форма тестирования является предпочтительной, так как экономит рабочее время персонала и другие ресурсы на проведение тестирования и обработку резуль-

татов, а также обеспечивает более высокий уровень достоверности, надежности и валидности. Иногда компьютерная форма требует более высоких затрат на подготовку, однако впоследствии эти затраты быстро окупаются. Глава 8 повторяет ряд моментов, освещавшихся в главах 3 и 6. Повторы сохраняются в тексте с целью обеспечения полноты информации для IT-специалистов, которые являются основными адресатами данной главы.

### **8.1. Условия для выбора компьютерной формы**

Выбор компьютерной формы тестирования обоснован при следующих условиях:

— большинство потенциальных тестируемых являются уверенными пользователями компьютерной техники, и компьютерная форма не будет создавать для них избыточный стресс;

— в организации имеются квалифицированные пользователи компьютерных программ тестирования, способные осуществить их грамотный выбор, а также выполнять предписания настоящего стандарта;

— организация либо располагает собственной технической базой и обученным персоналом для администрирования сеансов тестирования, либо заказывает у провайдеров не только программное обеспечение, но и комплексную услугу по реализации компьютерного тестирования в ситуации экспертизы.

В ситуации консультативного тестирования или в ситуации дистанционного онлайн-тестирования кандидатов фактически не требуется наличие собственной технической базы. Достаточным является договор с провайдером тестовых технологий, так как необходимое программное обеспечение размещается (хостируется) на серверах провайдеров, а тестируемые сами запускают и проходят онлайн-тесты в обычных браузерах.

### **8.2. Виды компьютерного тестирования и типы ситуаций**

В случае с компьютерным тестированием в особенности целесообразно различать тестирование под очным контролирующим наблюдением со стороны администратора (под супервизией) и дистанционное тестирование, которое, в свою очередь, может происходить при дистанционном наблюдении и без наблюдения со стороны администратора (самотестирование). В главе 3 основные виды дистанционного тестирования подробно определены в параграфах 3.1.2—3.1.4. Данные виды тестирования по способу контроля, как правило, сопряжены с различиями в правах принятия решений по результатам тестирования (администрация или сам тестируемый), а также с разными моделями отбора (жесткими и мягкими). В данном случае речь идет о необходимости различения ситуации экспертизы и ситуации консультирования (см. раздел 2.7). Зона интенсивных поисков и развития — это применение все новых технологических схем не непосредственного, а дистанционного контролирующего наблюдения и персонализированного доступа к тестам, основанного на адресных рассылках (см. об этом раздел 3.1).

8.2.1. *Административное тестирование* (или тестовый экзамен) в рамках схемы сопряжения оценочных методик по модели «жесткая воронка с отсевом» (подробнее см. в главе 9) требует строгого наблюдения за поведением тестируемых. В таких случаях тестовый экзамен наиболее целесообразно проводить в режиме очного присутствия администратора в компьютерной аудитории. При проведении административного тестирования в рамках схемы «мягкое баллирование без отсева» возможно использование дистанционного наблюдения (в частности, с помощью веб-камер), но подобная форма контроля признается более сложной и неоднозначной. Современные автоматизированные системы, использующие биометрические данные для аутентификации тестируемого (например, по отпечаткам

пальцев или радужной оболочке глаза), не дают гарантии отсутствия подсказок от лиц, находящихся в одном помещении с тестируемым, а также возможности дистанционного управления рабочим столом компьютера, за которым находится тестируемый, что может порождать ряд спорных ситуаций<sup>23</sup>.

8.2.2. При *консультативном тестировании* обследуемый участник наделяется правом самостоятельного принятия решений по результатам тестирования и является лицом, заинтересованным в достоверности результатов, поэтому административный контроль, как правило, в данном случае отсутствует.

Ниже приводится таблица 8.1, призванная проиллюстрировать возможные комбинации и рекомендуемые соответствия между формами наблюдения за поведением тестируемого, с одной стороны, и моделями принятия решения, — с другой. Эти возможные комбинации призваны помочь сетевому администратору в правильном проектировании разграничения доступа к различным материалам и результатам тестирования для разных сетевых пользователей в корпоративной сети Интранет. При определенной модели принятия решения возникает потребность во внедрении в компьютерную систему роли «супервизора», осуществляющего административный контроль.

Таблица 8.1 Соответствие форм наблюдения и моделями принятия решений

Агенты и модели принятия решения	Административный контроль (супервизор в системе тестирования)	
	Есть	Нет
Сам тестируемый и модель «мягкого баллирования»		V*
Администрация и модель «мягкого баллирования»	V	
Администрация и модель «жесткой воронки»	V	

\* **Примечание:** галочки указывают на рекомендуемые комбинации факторов.

Комбинация «административный контроль в сочетании с «мягким баллированием» означает, что тестируемый при сколь угодно низком показателе по тесту допускается до участия в других оценочных мероприятиях, однако распоряжается этой информацией о низком показателе не сам тестируемый, а администрация организации. Комбинация «сам тестируемый и модель «жесткой воронки» в таблице не указана, так как она не имеет содержательного смысла.

### 8.3. Особенности подготовки участников к компьютерному тестированию

В случае выполнения теста за компьютером все тестируемые должны быть в равной степени знакомы с основами компьютерной грамотности. Для тестируемых, не знакомых с компьютерным тестированием, за несколько дней до сеанса должна быть создана и предоставлена в свободном доступе (например, на корпоративном сайте) демонстрационная версия компьютерного теста, в точности имитирующая все действия тестируемого — управление клавиатурой, мышью, сенсорной панелью «тачпад» или экраном типа «тачскрин», а также другими устройствами ввода, если их предполагается использовать в основном тесте.

<sup>23</sup> Определенный компромиссный принцип, которому целесообразно следовать в этой спорной ситуации, может быть сформулирован так: низкие и средние результаты по тестам способностей и профессиональных знаний, полученные с использованием биометрической аутентификации, не требуют перепроверки, а высокие результаты требуют перепроверки под очным наблюдением за поведением тестируемого.



#### 8.4. Подготовка оборудования

При проведении компьютерного тестирования особой задачей выступает подготовка оборудования.

Следует позаботиться о том, чтобы участники тестирования не могли получать подсказки из параллельных окон на экране компьютера (например, из поисковых систем и карманных компьютерных энциклопедий), не пользовались переговорными и сетевыми устройствами, то есть отвечали на вопросы теста исключительно на основании собственных знаний, умений и способностей.

Следует проверить надежность оборудования по энергопитанию, чтобы сбои в энергоснабжении не приводили к остановкам и перезагрузкам компьютерной техники, а также устойчивость сетевых соединений в случае передачи тестовых заданий по сети.

Для обеспечения сохранности результатов тестирования следует проверить, дает ли программная система возможность сохранения уже имеющихся ответов участника в случае технических сбоев и разрывов в соединении (то есть возможность возобновить тест с того тестового задания, на котором произошел сетевой сбой).

В случае предъявления участникам тестирования звуковых стимулов (речь, музыка, шумовая имитация) каждый тестируемый должен быть снабжен головной гарнитурой с наушниками.

#### 8.5. Режим секретности при компьютерном тестировании

При проведении компьютерного тестирования пользователь теста (совместно с сетевым администратором) должен позаботиться о соблюдении режима секретности<sup>24</sup> в отношении тестовых материалов. Следует подчеркнуть, что без режима «тестовой безопасности» организация как коллективный пользователь теста не может выполнить перед разработчиком теста одно из важных договорных обязательств — о нераспространении лицензионных тестовых материалов (иначе организация понесет юридическую материальную ответственность за ущерб, нанесенный разработчику теста). Тестовые материалы в ходе тестирования неизбежно «раскрываются» и лишаются значительной части секретности, однако возможность предварительного изучения тестовых материалов, а тем более их копирования (тиражирования) должна отсутствовать. Современные требования к секретности (тестовой безопасности) можно обобщить в следующих тезисах:

— недопустима примитивная организация компьютерного тестирования с передачей лицензионных тестовых материалов по электронной почте; в случае проведения административного тестирования сетевая передача данных должна быть обеспечена только по защищенным протоколам (типа https, но не http);

— режим «разбор ошибок», при котором тестируемые получают доступ к тестовым заданиям и правильным ответам на них, имеет отношение к компьютерному обучению, но не может применяться в случае компьютерного тестирования в ситуации экспертизы;

— не следует смешивать режим секретности в отношении тестовых материалов и режим конфиденциальности при распространении тестовых данных (результатов): доступ к лицензионным (секретным) тестовым материалам и к результатам тестирования обычно предоставляется разным лицам, что должно быть закреплено в регламенте;

— категорически не рекомендуется публиковать персональные тестовые результаты на общедоступных ресурсах в Интернете;

<sup>24</sup> Понятия «режим секретности» и «информационная безопасность» являются пересекающимися, но не тождественными; в случае тестирования будет более точным термин «тестовая безопасность». «Тестовая безопасность» относится в равной степени к персональным данным участников и к тестовым заданиям.

— уровень требований к секретности тестовых материалов снижается, если они представляют собой демонстрационные и обучающие версии, специально предназначенные разработчиками для свободного распространения, а также при использовании квазитестов различного назначения (авторских, игровых, ознакомительных и т.п.), разработанных в организации для самотестирования сотрудников и/или кандидатов;

— в случае использования для самотестирования лицензионных тестовых методик пользователь должен соблюдать режим секретности в распространении тестовых материалов, выполняя договорные обязательства перед разработчиком тестовых технологий.

### **8.6. Аутентификация и авторизация**

Компьютерное тестирование персонала в режиме онлайн должно быть обеспечено эффективными средствами установления личности — аутентификации тестируемых<sup>25</sup>. При неправильной организации процедуры тестирования или при отсутствии контроля со стороны администратора после авторизации в контакте с компьютерной программой (и с самим тестируемым) могут оказаться третьи лица, проходящие тест под чужими логином и паролем или снабжающие тестируемого подсказками. Наиболее надежным средством аутентификации является присутствие в аудитории очного администратора, мотивированного на сотрудничество с организаторами тестирования, а наиболее ответственные административные тестовые процедуры не могут полагаться только на средства дистанционной аутентификации (биометрию, веб-камеры, сеансы связи по системам типа Skype, Viber и др.).

В отсутствии очных администраторов (супервизоров), принимающих на себя личную ответственность за достоверность результатов, тестовые результаты, полученные путем дистанционного тестирования, могут оказывать влияние на принятие предварительных решений, но не на окончательные управленческие решения.

### **8.7. Технические возможности в профессиональной автоматизированной системе тестирования**

При проведении компьютерного тестирования в ситуации экспертизы (тестового экзамена) администратор сеансов тестирования<sup>26</sup> должен располагать определенным набором технических возможностей для управления процессом компьютерного тестирования:

— средствами запуска процедуры тестирования в локальном или сетевом режиме, которые должны предполагать возможность быстрого включения и отключения группового сеанса тестирования с рабочего места администратора (в частности, необходимо иметь возможность отключить тест при появлении признаков массовой фальсификации результатов по причине утечки ключевых ответов);

— средствами дифференцированного по ролям парольного доступа к количественным и качественным результатам тестирования; это особенно необходимо, если конфиденциальность результатов гарантирована тестируемым в рамках подписанного ими информированного согласия (в случае компьютерного дистанционного тестирования аналогичный документ часто называют «договором-офертой», но точнее его называть «информированное согласие»);

25 Не следует смешивать процессы авторизации и аутентификации: авторизация сводится к использованию логинов и паролей, обеспечивает определенные права пользователя в компьютерной системе и не гарантирует аутентификацию тестируемого и достоверность результатов тестирования.

26 Несмотря на терминологическое сходство, не следует смешивать роли администратора (или оператора) сеансов тестирования и сетевого администратора-программиста. Второй — это, как правило, представитель ИТ-подразделения, который обеспечивает условия работы для первого — представителя HR-подразделения. Первый является пользователем компьютерной системы тестирования, второй — ее настройщиком и хранителем базы данных.

— средствами надежного долгосрочного хранения результатов тестирования в форматах, гарантирующих защиту от несанкционированного доступа к персональным данным (в частности, под персональными шифрокодами, или пин-кодами); система компьютерного тестирования не должна допускать технической возможности распространения тестовой информации на общедоступных ресурсах в Интернете, передачи по электронной почте посторонним лицам и т.п.

### **8.8. Требование лицензионной чистоты**

Компьютерное тестирование в организациях должно осуществляться с помощью лицензионно чистого программного обеспечения. Недопустимо использовать для этих целей программные средства, полученные «пиратским способом», а также средства, выпущенные безымянными разработчиками программного обеспечения (последнее, как правило, ведет за собой потерю ответственности за надежность хранения и адресной точности в распространении тестовых инструментов и тестовой информации). Вместо пиратских копий лицензионных программ пользователь может применять свободно распространяемые (бесплатные) обучающие программные системы с открытым кодом<sup>27</sup>.

### **8.9. Соглашение о конфиденциальности с соисполнителями**

В случае проведения компьютерного тестирования организатор тестирования (пользователь компьютерной тестовой методики) должен лично разобраться в схеме информационных потоков для точного понимания прав и сроков доступа к тестовым материалам и к результатам тестирования. Технические специалисты и ассистенты-операторы (как и администраторы-супервизоры на местах проведения тестирования) оказывают техническую поддержку и консультативную помощь, но не принимают на себя ответственность за вынесение окончательных организационно-технических решений. В силу обладания ими доступом к различным компонентам компьютерной системы тестирования, организатор должен заключить с ними «соглашение о конфиденциальности» (подписанное полномочным представителем тестирующей организации) или взять «обязательство по обеспечению конфиденциальности» (односторонний документ), предполагающие наличие должностной и юридической ответственности этих лиц.

Крайне желательно ограничить число лиц, получающих подобный доступ, исключив тех, кто может быть заинтересован в нарушении режима секретности в распространении тестовых материалов, а также режима конфиденциальности в распространении результатов тестирования. Подписание юридически обязывающих документов является необходимым, но не достаточным условием лояльности со стороны всех соисполнителей, то есть не дает полных гарантий информационной безопасности. Однако в отсутствие таких документов нарушения оказываются слишком вероятными по причине элементарной неосведомленности ИТ-специалистов в том, что какие-то тестовые материалы являются секретными.

### **8.10. Действия организатора в случае нарушения информационной безопасности**

Организатор компьютерного тестирования должен осознавать наличие рисков утечки тестовых данных в результате различных неправомерных действий (например, хакерских атак, фотосъемки тестовых материалов с последующей их публикацией и т.п.). В случае обнаружения утечки тестовых материалов организатор должен:

<sup>27</sup> Программные системы с открытым кодом более уязвимы для взлома и последующей утечки лицензионных тестовых материалов, так как подсистемы шифрования в них развиты слабо или отсутствуют, о чем необходимо помнить всем пользователям компьютерных тестовых методик.

- быстро остановить процесс тестирования, сократив тем самым масштабы утечки и возможной фальсификации данных тестирования;
- официально обратиться к разработчику (или правообладателю) тестовой системы (а в ряде массовых и опасных нарушений в соответствующие подразделения правоохранительных органов Российской Федерации) с просьбой оказать действенную помощь в обнаружении источников утечки и предотвращении дальнейшей утечки;
- в случае обнаружения в открытых (социальных) сетях или на публичных сайтах в Интернете секретной или конфиденциальной тестовой информации организатор должен обратиться за помощью к провайдером, контролирующим сервера этих социальных сетей, с просьбой остановить утечку и помочь в выявлении лиц, ответственных за утечку.

### **8.11. Автоматизированные обучающие системы и автоматизированные системы тестирования**

В общем случае целесообразно различать автоматизированные обучающие системы (АОС, или СДО — системы дистанционного обучения) и специализированные автоматизированные системы тестирования (АСТ). Наиболее мощные АОС включают в себя АСТ как подсистемы. Большинство АОС, как правило, могут быть использованы только для самотестирования кандидатов и выполнения тестируемыми демоверсий тестов для самоподготовки и самообучения работников. Недопустимо ставить знак равенства по уровню необходимого режима секретности между материалами, предназначенными для обучения, и материалами, предназначенными для тестового контроля в ситуации экспертизы (тестового экзамена). Для целей административного тестирования предпочтительнее использовать специализированные системы АСТ или развитые подсистемы АСТ в рамках мощных АОС.

### **8.12. Технический регламент**

Для внедрения компьютерной системы административного тестирования организатор должен разработать собственный технический регламент или адаптировать для своей организации типовой регламент. В ряде случаев данный документ может называться «организационно-техническим регламентом». Этот документ должен соответствовать не только настоящему Стандарту тестирования персонала, но и международным и национальным правовым нормам, в частности, Федеральному закону №152 РФ «О защите персональных данных» и другим подобным документам. Для внедрения компьютерной системы в целях консультативного тестирования технический регламент не требуется.

## **Глава 9. Сопряжение с другими оценочными процедурами и контроль качества тестирования**

В практике управления персоналом организации для решения оценочных задач могут применяться как тестовые, так и нетестовые методы оценки (например, моделирующие упражнения, интервью, бизнес-симуляции, решение кейсов, обследование с использованием полиграфа, 360 градусов, ключевые показатели эффективности, деловые игры). Результаты одного и того же человека по разным методам оценки целесообразно сравнивать, а случаи расхождения анализировать.

## 9.1. Схемы сопряжения

Различные методы оценки применяются как независимо друг от друга, так и в комбинации, образуя комплексные оценочные технологии (например, метод «центр оценки»). Чаще всего они применяются для оценки разных компетенций, но могут быть использованы и для оценки одних и тех же компетенций. В обоих случаях возникает задача присвоения каждому из методов тех или иных весов при подсчете интегрального рейтинг-балла тестируемого.

9.1.1. *Различные комбинации тестовых и нетестовых технологий.* В комплексных оценочных технологиях, как правило, экспертные оценки имеют больший вес, чем данные по тестам<sup>28</sup>. Однако в общем случае возможные комбинации тестовых и нетестовых технологий, а также присваиваемых им весов, могут быть различными:

а) «мягкое баллирование» (или «комплексное баллирование») с преобладанием экспертных оценок (например, экспертные оценки весят 66% или 75% в интегральном рейтинге, а тестовые технологии 33% или 25%);

б) «мягкое баллирование» с преобладанием тестовых данных (это бывает оправдано в тех случаях, когда тесты носят характер объективного испытания, а экспертами выступают внутренние сотрудники организации, не обладающие серьезным опытом в экспертных оценках);

в) «воронка отбора» с первичным фильтром в виде экспертной оценки (к тестированию допускаются только те, кто прошел экспертный отбор);

г) «воронка отбора» с первичным фильтром в виде тестовой процедуры (к экспертным оценкам, например, к интервью, допускаются только те, кто прошел тестовый отбор).

«Воронка отбора», в свою очередь, может быть «вынужденной» (в ситуации экспертизы) или «добровольной» (в ситуации консультирования). Во втором случае сам тестируемый принимает решение об участии в дальнейших этапах отбора.

9.1.2. *Ситуации обязательного сопряжения.* Тестовые методики, как правило, не могут быть использованы как единственный источник информации при принятии таких кадровых решений, как прием на работу, включение в кадровый резерв, аттестация на соответствие занимаемой должности, увольнение и т.п. При этом для ряда особых видов профессиональной деятельности существуют объективные тесты, которые могут быть использованы как достаточный критерий при решении вопроса о профессиональном несоответствии работника (например, тесты на цветоразличение для профессиональных водителей). В этом случае возникает асимметричная схема сопряжения методик: низкие баллы по тесту приводят к отсеву, а высокие баллы говорят лишь о необходимости дальнейшего обследования. Таким образом, возникает обязательная «воронка отбора» с тестом в виде первичного фильтра (случай 9.1.1г).

9.1.3. *Ситуации автономного использования.* Тестовые методики могут быть использованы как единственный источник информации при принятии таких кадровых решений, как определение потребности в обучении, оперативный допуск на рабочее место (в начале рабочего дня), квалификационные испытания по окончании образовательной программы и т.п.

## 9.2. Согласование результатов при сопряжении

Для повышения качества принимаемых решений в особо ответственных ситуациях рекомендуется оценивать одни и те же компетенции и тестовыми, и нетестовыми методами. Это

28 Например, результаты устного и письменного экзамена в комплексных процедурах вузовской аттестации или профессиональной сертификации традиционно имеют больший вес, чем результаты тестового экзамена.

повышает валидность оценки в целом и снижает вероятность ошибок, которые могут быть вызваны, например, хорошей подготовкой участников оценки к прохождению какой-либо конкретной оценочной процедуры или утечкой информации по ключам теста.

В ситуации «мягкого баллирования» при оценивании одних и тех же компетенций тестовые и нетестовые методы могут давать несовпадающие результаты. На допустимость или недопустимость величины подобных расхождений указывает величина коэффициента корреляции. В обычном (допустимом) случае коэффициент корреляции должен превышать значение в 0,3. При получении значений коэффициента корреляции ниже 0,3 необходимо провести углубленный анализ причин таких расхождений — выяснить возможность методических ошибок и организационных просчетов в проведении тестирования или нетестовых методов, различий в трактовке оцениваемой компетенции у экспертов и у разработчиков тестов и т.п.

### 9.3. Выборочный контроль

При получении крайне низких или крайне высоких результатов по тестам рекомендуется проверить полученные данные с помощью других методов для исключения возможных артефактов в виде несанкционированного доступа к ключам теста, «списывания», неправильно понятой инструкции или непонимания значимости тестирования. Например, неправдоподобно высокие (близкие к 100%) результаты сразу нескольких человек из одного сеанса тестирования в одном помещении позволяют допустить отсутствие должного контроля со стороны администратора и являются основанием для повторного вызова данных участников тестирования с целью очного выполнения ими аналогичных тестовых испытаний. В ряде ситуаций массового тестирования (когда в короткий промежуток времени обследуются сотни или тысячи человек) подобная проверка может носить выборочный характер. Если итоги этой выборочной проверки говорят о невалидности тестовой методики или недостоверности тестовой технологии, следует пересмотреть схему категоризации тестовых результатов и принятия решений вплоть до отказа от учета результатов тестирования.

### 9.4. Проверка валидности тестирования

Использование комбинации тестовых и нетестовых методов в рамках пилотных проектов позволяет дополнительно валидизировать<sup>29</sup> (проверить на пригодность) тестовую методику и тестовую технологию применительно к конкретной организации и управленческой задаче. Выбор в пользу психометрических тестов, отвечающих всем требованиям к качеству, не снимает с пользователей ответственности за корректность их применения при решении тех или иных управленческих задач, а также за обеспечение качества тестовой технологии. Дополнительная валидизация тестирования предполагает осуществление стандартного корреляционного анализа связей между тестовыми и нетестовыми данными. Этим обеспечивается возможность получения количественного результата (коэффициента корреляции), сопоставимого с заданными нормативами валидности (подробнее см. главу 4). При отсутствии в организации специалистов, готовых произвести такой анализ на должном уровне, целесообразным является обращение в специализированные консалтинговые компании или научные организации за помощью в его проведении.

<sup>29</sup> Проверка валидности по содержанию выполняется, как правило, разработчиком теста и обозначается термином «валидация», в то время как проверка критериальной (прагматической) валидности выполняется часто пользователем теста уже в ходе практического применения методики и обозначается термином «валидизация». Валидизация привносит в изначальные сведения о валидности теста важные дополнительные сведения.

### 9.5. Принцип превентивного (упреждающего) тестирования

При проведении комплексной оценки, включающей тестовые и нетестовые методы, тестирование хронологически целесообразно производить раньше других оценочных процедур<sup>30</sup>. При этом необходимо принять во внимание ряд соображений:

— тестовые методы менее трудоемки при проведении и более дешевы по себестоимости (в пересчете на одного обследованного);

— тестовые данные в рамках комплексных технологий оценки могут быть основанием для формулирования оценочных (диагностических) гипотез, которые впоследствии будут проверены другими методами;

— устная обратная связь по тестовым методикам может быть интегрирована в интервью или другие подобные методы оценки;

— применение «мягкого баллирования» более предпочтительно, нежели применение «жесткой воронки», во всех случаях, где это возможно по организационным и экономическим соображениям (см. главу 2 и параграф 9.1.2);

— большинство тестовых процедур эффективно для решения задач отсева однозначно неподходящих, «слабых» кандидатов и неэффективно для отделения «сильных» от «средних»;

— применение «жесткой воронки» может повлечь за собой ошибочный отсев подходящих кандидатов в случаях, когда диагностируемые тестом критерии (например, конкретные интеллектуальные способности) не являются критичными для успешности в должности, на которую ведется отбор;

— обоснованный процент отсева кандидатов при использовании «жесткой воронки», как правильно, не превышает 30%<sup>31</sup>;

— сложные интегральные компетенции (например, профессиональная компетентность или управленческий потенциал) не могут диагностироваться только тестовыми методами.

### 9.6. Количественная и ранговая модели сопряжения с нетестовыми методиками

При использовании такой схемы сопряжения, как «мягкое баллирование», для подсчета интегральных рейтинг-баллов стандартизированные баллы по тестовым методикам, с одной стороны, и экспертные оценки по нетестовым методикам, с другой стороны, могут быть объединены. Экспертные оценки часто выносятся с помощью рангов (грейдов), которые сопровождаются словесными или буквенными обозначениями, например, А — высокое значение компетенции, В — выше среднего, С — ниже среднего, D — низкое значение. Интеграция результатов может при этом производиться несколькими способами.

Во-первых, путем перевода экспертных оценок в те стандартизированные шкалы, которые используются для стандартизированных тестов, с последующим суммированием экспертных и тестовых оценок. Например, в шкалу стэнов экспертные оценки по традиционной четырехбалльной шкале могут быть переведены следующим образом: оценка А «отлично» — балл 9, оценка В «хорошо» — балл 6,5, оценка С «удовлетворительно» — балл 4,5, оценка D «плохо» — балл 2 (см. пример 1 в Таблице 9.1).

30 Исключение могут составлять случаи, когда для поддержания имиджа работодателя и/или интереса внешних кандидатов к вакантной должности (руководители или высококвалифицированные специалисты) оказывается целесообразным первыми проводить оценочные мероприятия, связанные с непосредственным общением (например, интервью).

31 Особенно это касается оценки кандидатов на управленческие позиции, для которых тестовые показатели редко гарантируют высокое качество работы. Более жесткий отсев (на уровне 50% или более) может быть приемлем в ситуации очень высокого конкурса на массовые низовые позиции или на позиции узких специалистов (свыше пяти кандидатов на одну вакансию), а также при наборе недопустимо низких баллов по тесту на профессиональные знания и умения (например, невладение десятипальцевым методом печати у претендента на позицию секретаря-письмоводителя).

Таблица 9.1. Примеры возможных вариантов соответствия десятибалльной шкалы стэнов и четырех оценочных категорий

Категория	D	C	B	A
Оценка	Неудовлетворительно	Удовлетворительно	Хорошо	Отлично
Пример 1	1 — 3	4 — 5	6 — 7	8 — 10
Пример 2	1 — 5	6 — 7	8 — 9	10

Во-вторых, путем перевода стандартизированных баллов по тестам в оценочные категории (градации А, В, С, D) с последующим составлением таблиц для указания возможных комбинаций оценочных категорий по тестам и экспертным оценкам (например, комбинация из категории А по тесту и категории В по экспертным оценкам дает Решение 2). См. Таблицу 9.2.

Таблица 9.2. Пример перевода комбинации оценок по двум методикам в разные управленческие решения

Методика 1 (тестовая)	Методика 2 (нетестовая, экспертная)			
	Категория D	Категория C	Категория B	Категория A
Категория А	Решение 3	Решение 2	Решение 1	Решение 1
Категория В	Решение 4	Решение 4	Решение 3	Решение 3
Категория С	Решение 4	Решение 4	Решение 4	Решение 3
Категория D	Решение 4	Решение 4	Решение 4	Решение 4

*Примечание.* Решение 1 — допуск к самостоятельной работе в искомой должности. Решение 2 — допуск к работе с участием наставника-супервизора. Решение 3 — допуск к производственному обучению. Решение 4 — недопуск к обучению на искомую должность.

## Глава 10. Обратная связь и принятие решений по результатам тестирования

Обратная связь (ОС) является заключительным этапом процедуры тестирования и представляет собой информацию для тестируемых по результатам проведенной диагностики. Информация для заказчиков и информация для лиц, принимающих решения (ЛПР), предоставляемая им после проведения тестирования, в данном стандарте различаются терминологически в силу возможного несовпадения данных ролей в крупных организациях.

### 10.1. Необходимость обратной связи

Предоставление тестируемым обратной связи в той или иной форме является *обязательным и необходимым* как в ситуации консультирования, так и в ситуации экспертизы. Отсутствие обратной связи признается ошибочным как в силу несоответствия морально-этическим нормам, так и по причине повышения вероятности принятия ошибочных управленческих решений и/или отсутствия развивающего эффекта от проведенного тестирования.

### 10.2. Функции и цели обратной связи для тестируемых

Функции и цели обратной связи:

- информационная — открытое информирование тестируемого о результатах тестирования;
- этическая — тактичное раскрытие тестируемому информации о нем;



- верификационная — уточнение и дополнение результатов тестирования;
- мотивирующая — формирование конструктивного отношения к результатам тестирования и позитивного восприятия аналогичных оценочных процедур в дальнейшем;
- развивающая — стимулирование тестируемого к целенаправленному развитию в указанных направлениях;
- организационная — оповещение тестируемого об организационно-управленческих решениях, принимаемых на основе результатов (например, приглашение успешного кандидата на следующий этап отбора).

### 10.3. Формы обратной связи

Обратная связь может предоставляться в следующих формах.

10.3.1. *Краткой или развернутой* (в виде короткого утверждения о степени соответствия результатов тестирования требованиям организации или развернутого сообщения о потенциале тестируемого, его преимуществах, недостатках, факторах риска и т.п.).

10.3.2. *Устной или письменной* (в виде очного сообщения, сообщения по телефону, видеосвязи или автоматизированного письменного отчета по результатам компьютерного тестирования).

10.3.3. *Немедленной или отсроченной* (сразу по завершении сеанса тестирования или по прошествии некоторого времени с момента его завершения). Немедленная обратная связь является более предпочтительной, однако в ее отсутствии необходимо сориентировать тестируемого по срокам получения отсроченной обратной связи. Отсрочка не должна превышать один месяц. В случае задержек в обработке результатов тестирования обратная связь может предоставляться поэтапно в разных объемах: 1) краткая немедленная; 2) развернутая отсроченная.

Обратная связь по результатам тестирования может предоставляться вместе с обратной связью по результатам других методов оценки в случае использования комплексных оценочных технологий (например, Центра оценки).

### 10.4. Содержание обратной связи

По содержанию обратная связь предполагает сбалансированную качественную интерпретацию, содержащую указание как на позитивные, так и на негативные результаты с точки зрения их возможного воздействия на профессиональную успешность. Вне зависимости от функций, целей и формы она должна включать в себя следующее:

- информацию об измеряемых характеристиках (оцениваемых свойствах);
- информацию о принципах обработки результатов тестирования;
- информацию об особенностях нормативной группы, с которой сравнивается тестируемый (общая выборка по популяции, определенная профессиональная группа или локальная группа сотрудников организации);
- качественную и/или количественную информацию о результатах тестируемого.

В случае автоматизированной письменной обратной связи предпочтительно текстовое описание количественных результатов во избежание их ошибочной интерпретации. Методическая и техническая информация должна быть представлена в понятной тестируемому форме. В зависимости от особенностей тестирования обратная связь может также включать:

- информацию о качественных решениях (например, «сдано / не сдано», «пригоден / не пригоден»);

- объяснение взаимосвязи оцениваемых качеств с управленческими решениями, которые будут приниматься на основе полученного результата;
- рекомендации по развитию оцениваемых характеристик;
- информацию о возможностях и ограничениях использования результатов тестирования.

### **10.5. Требования к исполнителям и технической реализации**

Лица, предоставляющие устную неформализованную обратную связь (например, администратор тестирования или эксперт-консультант), должны обладать навыками предоставления обратной связи и интерпретации результатов конкретной методики. В случае использования коммерческой лицензированной методики под этими навыками понимается прохождение специальной подготовки и обучения на сертификат пользователя данной методики.

10.5.1. В случае использования квазитестов профессиональных знаний и ситуационных умений, не являющихся психометрическими, обратная связь на первом этапе должна включать в себя информацию о проценте правильных ответов, а на втором — информацию об относительном результате тестируемого в сравнении с другими (процентильном балле, — см. раздел 7.4.). Причем последний этап желательно реализовать в конфиденциальном режиме программными средствами и обеспечить тестируемому возможность многократного доступа к своим результатам.

10.5.2. Содержание автоматизированной письменной обратной связи, предоставляемой компьютерной системой тестирования, находится в зоне ответственности ее разработчиков (требования к разработчику см. в главе 5), хотя окончательное решение о ее адекватности в данной ситуации обследования принимает пользователь методики. Обратная связь не должна носить жесткий или безоговорочный характер и во многих ситуациях должна иметь возможность дополнительной настройки сообщений с учетом особенностей тестируемых, корпоративной культуры организации и т.п.

10.5.3. Особые требования к ОС возникают, когда получены настолько низкие результаты тестирования (неблагоприятные в организационном контексте), что появляются основания ожидать возможный психотравмирующий эффект обратной связи. В этом случае предпочтительнее организовать не отчужденную письменную передачу информации, а живой человеческий диалог, в ходе которого можно выявить, насколько драматично воспринимает тестируемый свои результаты и как именно его можно ободрить.

Правила хранения и предоставления информации, содержащейся в индивидуальных письменных отчетах, описаны в главах 7, 8 и 9.

### **10.6. Информирование заказчиков тестирования и лиц, принимающих решения**

Ситуация экспертизы характеризуется получением доступа к персональным результатам тестирования со стороны заказчиков и лиц, принимающих решения, которые могут быть представлены им в различной форме, но при обязательном соблюдении ряда принципов.

10.6.1. Результатами тестирования для заказчиков и/или лиц, принимающих решения, могут являться:

- профили по тестовым шкалам совместно с информацией о пороговых значениях (точках отсечения при отборе и т.п.);
- содержательная интерпретация тестовых баллов (в том числе с указанием на взаимосвязи с успешностью в профессиональной деятельности);

— интегральные рейтинг-листы по степени соответствия тестируемых требованиям должности;

— диаграммы и графики по результатам тестирования отдельных участников или их группы, а также иные формы визуализации;

— выводы и рекомендации (например, индивидуальные и/или групповые экспертные заключения, письменные отчеты).

10.6.2. Содержание результатов тестирования, форма их фиксации и представления должны быть согласованы с заказчиками на предварительном этапе.

10.6.3. Смысл тестовых шкал и полученных результатов должен специально разъясняться лицам, принимающим решения, во избежание возможных недоразумений. В отдельных случаях целесообразно запрашивать их понимание результатов на конкретных примерах.

10.6.4. Информация для принятия решений предоставляется только уполномоченным лицам, непосредственно задействованным в принятии управленческих решений по результатам тестирования; правила доступа этих лиц к результатам должны быть заранее оговорены во внутренних документах компании и в информированном согласии тестируемых.

10.6.5. Участники тестирования до выполнения теста должны быть проинформированы о круге лиц, принимающих решения (с точностью до должностей). В отсутствие такого информирования организатор тестирования уполномочен проводить только консультативное, но не административное тестирование. Подробнее об информировании тестируемых см. раздел 3.7.

10.6.6. Любая информация личного характера о тестируемом, не соответствующая заявленным целям тестирования, сохраняет статус конфиденциальной для лиц, принимающих решения.

10.6.7. Заказчик должен быть проинформирован о возможностях и ограничениях использования результатов тестирования при принятии на их основе управленческих решений (см. главы 2, 7 и 9). Заказчик должен быть сориентирован, что любой прогноз успешности работника на основе результатов тестирования носит вероятностный характер, и сохраняется серьезная вероятность ошибки (как правило, не ниже нескольких десятков процентов, но в каждом случае лучше указывать точную конкретную цифру).

10.6.8. При проведении административного тестирования информация для лиц, принимающих решения, содержит персональные данные тестируемых и их результаты, о недопустимости обнародования которых они должны быть проинформированы. Исключениями являются случаи, когда публикация информации оговаривается со всеми участниками процедуры заранее (например, в образовательном тестировании).

## **10.7. Принятие решений на основе результатов тестирования**

Рекомендации по принятию решений могут быть включены организаторами тестирования в отчетную информацию для лиц, принимающих решения. При этом должны быть соблюдены перечисленные ниже принципы.

10.7.1. Результаты тестирования используются главным образом при принятии решений, связанных с наймом и с улучшением условий профессиональной деятельности сотрудников организации в следующих случаях:

- продвижение;
- ротация;
- присвоение или подтверждение квалификации;
- включение сотрудников в особые категории (кадровый резерв, «высокий потенциал», или HiPo, и т.п.);

- направление на обучение.

10.7.2. Результат тестирования как единственной оценочной процедуры в ситуации экспертизы не является основанием для принятия решений, ухудшающих условия профессиональной деятельности сотрудников организации (например, понижение в должности, лишение премии, увольнение), а в ситуации консультирования тестируемый сам принимает решение по неудовлетворительным результатам тестирования.

10.7.3. Результат тестирования в ситуации экспертизы используется третьими лицами для принятия решений при соблюдении следующих принципов:

- позитивные решения (например, о повышении в должности) принимаются на основе комплекса оценочных мероприятий;
- негативные решения (например, о недопуске к какому-либо виду деятельности или этапу оценочных испытаний) принимаются только при заведомо низких результатах тестирования с заранее объявленными пороговыми значениями;
- решение об увольнении сотрудника по причине несоответствия требованиям занимаемой должности может приниматься только на основе комплекса мероприятий по аттестации с участием аттестационной комиссии в соответствии с действующими положениями ТК РФ.

10.7.4. Результат тестирования используется для принятия решений, заранее известных участникам. О типах решений по результатам тестирования в разных ситуациях см. также главы 2, 7 и 9.

## **10.8. Использование и распространение результатов тестирования**

10.8.1. Следует считать некорректным использование результатов тестирования в целях, отличных от изначально заявленных для участников тестирования.

10.8.2. Список лиц, имеющих доступ к результатам тестирования, должен быть зафиксирован во внутренних документах организации. Сотрудники организации, имеющие доступ к результатам тестирования, обязаны обеспечить их корректное использование в соответствии со статусом результатов тестирования (подробнее о техническом регламенте и условиях доступа к результатам тестирования см. главу 8).

10.8.3. Недопустимо распространение результатов, дискредитирующих участников тестирования внутри или вне организации, а также нарушающее их права.

10.8.4. Публичное распространение результатов тестирования возможно только с согласия тестируемого, оформленного документально в виде подписанного тестируемым информированного согласия (см. главы 6, 8, 11).

10.8.5. В зависимости от специфики оцениваемых свойств результаты тестирования имеют различные сроки использования, которые определяются разработчиком. Для большинства тестов профессиональных знаний и умений эти сроки составляют один год, для большинства собственно психологических тестов — два года. По истечении указанных сроков результаты тестирования могут использоваться только в деперсонализированном виде и только в исследовательских целях.

## **Глава 11. Этические нормы в тестировании персонала**

В настоящей заключительной главе содержатся предписания и рекомендации, которые могут послужить опорой для анализа жалоб в ходе работы конфликтных комиссий. Эти рекомендации частично пересекаются с содержанием предыдущих глав, но формулируются и комментируются с точки зрения этических норм в тестировании персонала, а также возможного столкновения прав и интересов различных групп лиц.

Пользователи методик, под которыми обобщенно понимаются все категории исполнителей тестирования внутри организации, являются субъектами определенных взаимоотношений, возникающих в ходе реализации программы тестирования. Говорить об этих отношениях, не объединяя исполнителей тестовых и нетестовых оценочных методик, неправильно. За специалистами по оценке персонала, реализующими программы тестирования в организации, в данной главе закреплен термин «пользователь».

Жалобы и претензии со стороны физических лиц и организаций в отношении деятельности пользователей тестовых методик принимаются к рассмотрению в комиссиях по этике, созданных в рамках профильных общественно-профессиональных объединений (таких как Российское психологическое общество или Федерация оценки персонала). Настоящая глава суммирует принципы, на которых должна базироваться деятельность подобных комиссий.

### **11.1. Этика, право и этико-правовые дилеммы**

11.1.1. Согласно принципу этической и юридической правомочности, любая полученная в ходе процедуры оценки информация может быть использована только в рамках действующего законодательства.

11.1.2. В случае, если этические стандарты вступают в противоречие с требованиями закона (законодательными актами РФ и международным правом), пользователь должен сообщить лицам, принимающим решения, и/или разработчикам тестовой методики об обязательствах и ограничениях, связанных с их исполнением, и предпринять шаги для ответственного разрешения этико-правового конфликта вплоть до аргументированного отказа от выполнения определенного вида работ.

11.1.3. Если ни закон, ни этика профессионального поведения не проясняют, как действовать в определенной ситуации, пользователь должен принимать во внимание другие материалы (требования к процедуре, к инструментарию, к деловой коммуникации и т.д.) и по необходимости обращаться за консультацией к другим специалистам и/или разработчикам тестовой методики.

11.1.4. Если пользователь ограничен в праве предоставления тестируемому информации по результатам тестирования в той или иной форме, то последний должен быть оповещен об этом до начала тестирования в ходе консультации и по факту подписания информированного согласия на участие в процедуре и обработку персональных данных.

11.1.5. Пользователям следует принимать во внимание отсутствие готовых правовых норм для ряда ситуаций, связанных с тестированием персонала. Например, это касается конфликтных ситуаций и/или обнаруженных фактов нарушения этических предписаний со стороны третьих лиц<sup>32</sup>. В ситуации экспертизы интересы организации преобладают над интересами тестируемого, что во многом определяет логику действия тех или иных этических нормативов, а также разрешения возникающих дилемм.

### **11.2. Компетентность как этическая норма**

За внедрение и эксплуатацию системы тестирования в организации несут ответственность специалисты разной квалификации, выступающие в разных ролях и наделенные различными обязательствами и полномочиями (более подробно см. главу 5). В этом разделе в определенной степени повторяются главные требования к этим специалистам, сформули-

<sup>32</sup> Правовую неопределенность в данных случаях призваны компенсировать базовые принципы профессиональной этики, сформулированные, в частности, в Этическом кодексе психолога (материалы Президиума Российского психологического общества, см.: <http://рпо.рф/рпо/documentation/ethics.php>).

рованные в главе 6, но добавляется принцип общественно-профессионального контроля за квалификационным уровнем (см. параграф 11.2.6).

11.2.1. Специалисты, роль которых связана с проведением тестирования в организации, должны поддерживать необходимый уровень компетентности для осуществления своих профессиональных функций и повышать собственную квалификацию в ходе специальных обучающих мероприятий, а также путем обмена опытом с коллегами, развивать навыки профессионального общения (в частности, предоставления обратной связи по результатам тестирования) и демонстрировать профессиональную зрелость.

11.2.2. Специалисты без высшего образования при проведении тестирования могут выступать только во вспомогательной роли администраторов тестирования (как операторы компьютерной техники и исполнители по обработке протоколов ответов, подготовке аппаратуры и тиражированию стимульного материала при бланковом тестировании, осуществлению контроля поведения тестируемых при выполнении тестов, общему техническому сопровождению процедуры оценки).

11.2.3. Специалисты с высшим образованием, обладающие базовыми знаниями в области теории тестов, психодиагностики и статистики, могут выступать во всех ролях, связанных с проведением тестирования, но не с разработкой тестов и/или их батарей, а также с проектированием схем и моделей их сопряжения с иными оценочными процедурами. Последние виды работ требуют не базового, а полного высшего образования в области психологических или педагогических измерений.

11.2.4. Специалисты по оценке могут делегировать своим сотрудникам и стажерам только те полномочия, связанные с проведением тестирования, компетентного выполнения которых можно от них ожидать в силу их образования, подготовки и опыта (при участии супервизора или без него).

11.2.5. Специалисты по оценке не должны допускать использования в той или иной форме тестовых методик лицами, не имеющими необходимой квалификации.

11.2.6. При рассмотрении жалоб в комиссиях по этике (этических комитетах) может быть поставлен вопрос об оценке квалификации пользователей тестовых методик, допустивших конфликтную ситуацию, и об оценке уровня их профессиональных знаний в объеме настоящего стандарта.

### **11.3. Этические нормы, связанные с преимуществами особой информированности**

Лица, вовлеченные в проведение тестирования в организации, получают в силу своей профессиональной позиции определенные преимущества в виде особой информированности о применяемых методиках, о тестируемых и о возможных последствиях внедрения тех или иных методик в организации, что накладывает на них дополнительные морально-этические ограничения.

11.3.1. Пользователи тестовых методик воздерживаются от участия в процедуре тестирования, если знают, что их личные особенности и/или проблемы могут нанести ущерб или вред участникам тестирования или другим лицам, по отношению к которым они принимают на себя профессиональные обязательства. Хотя тестовые технологии являются более объективированными, чем большинство иных оценочных процедур, однако в случае необходимости принятия нестандартных решений по результатам тестирования и в ряде других ситуаций личность пользователя также оказывает влияние на процесс тестирования.

11.3.2. На разных этапах подготовки, проведения и анализа результатов тестирования пользователи тестовых методик стараются исключить влияние различных социальных, финансовых, организационных или политических факторов, не имеющих отношения

к осуществляемой ими профессиональной деятельности, на судьбу людей, участвующих в тестировании в статусе оцениваемых.

11.3.3. Пользователи тестовых методик аргументированно отказываются от проведения тестирования, если им становится известным, что велика вероятность злоупотребления полученными данными и/или манипулирования результатами с целью продвижения несправедливого решения.

11.3.4. При выборе методик пользователи воздерживаются от использования и принятия решений по тестам, которые признаны устаревшими, потерявшими диагностическую ценность или не пригодными для реализации поставленных целей.

#### **11.4. Этика деловой коммуникации (взаимоотношений) в ситуации оценки**

11.4.1. При взаимодействии с тестируемым администраторам тестовых методик следует:

- корректно, на условиях сотрудничества соблюдать в отношениях дистанцию, позволяющую тактично сохранять ролевые различия, четко исполнять все взятые на себя обязательства, ограничивая их сферой собственной компетентности;
- не оказывать на тестируемого психологического воздействия силой убеждения, избегать предвзятости и присвоения ярлыков, соблюдать принцип свободы выбора;
- проявлять открытое доброжелательное отношение, уважать труд тестируемого, рационально использовать время;
- терпеливо и уважительно относиться к возможным фактам отказа от участия в тестировании;
- избегать конфликтов при разрешении разногласий.

11.4.2. При взаимодействии с лицами, принимающими решения, пользователям тестовых методик следует:

- до начала оценочных процедур прояснить характер ситуации тестирования, особенности требуемой информации и ограничения, связанные с требованием конфиденциальности и защитой персональных данных;
- соблюдать строгую конфиденциальность в отношении информации, полученной в процессе работы;
- не допускать преднамеренного искажения или утаивания имеющейся информации о низких результатах тестирования (тестовой оценки).

11.4.3. При взаимодействии с коллегами пользователю тестовых методик необходимо:

- поддерживать профессиональную кооперацию, включая кооперацию с коллегами, применяющими нетестовые методы оценки (возможность обмена информацией в целях повышения качества методической и организационной практики, а также тестовой оценки испытуемых);
- поддерживать уважительные отношения, оказывать помощь и поддержку коллегам в методически и/или этически сложных ситуациях;
- не предпринимать действий, в результате которых может пострадать профессиональная репутация коллег.

#### **11.5. Национально-культурные особенности**

11.5.1. Специалист по оценке персонала (пользователь методик, тестолог) в своей профессиональной деятельности:

- уважает права других людей придерживаться обычаев, ценностей, установок, мнений, отличных от его собственных, если они не противоречат действующему законодательству;

— при поиске и отборе соискателей не допускает дискриминации по половым, расовым, национальным, религиозным, возрастным или политическим признакам и не участвует в любых мероприятиях, включающих несправедливую оценку, стремится учитывать языковые предпочтения, компетентность и культурный уровень участников тестирования.

11.5.2. В случае, если национально-культурные (прежде всего, языковые) различия существенны и могут влиять на справедливость отбора кандидатов, от пользователя тестовых методик требуется заказать работы по адаптации (локализации, модификации) как процедуры, так и используемого инструментария у разработчика тестовой методики, включая культурно-языковую адаптацию тестовых заданий с последующим получением новых данных по валидности, надежности и тестовым нормам. Если исполнение такого заказа оказывается невозможным в приемлемые сроки (или из-за финансовых ограничений), то необходимо осуществить замену неадекватной процедуры тестирования на иную оценочную процедуру.

### **11.6. Информированное согласие и использование результатов тестирования**

Информированное согласие как один из ключевых организационных документов при проведении тестирования уже неоднократно описывалось в главах 6, 8 и 10. В данном разделе акцент делается на понимании этического смысла профессиональных нормативов, связанных с применением данного инструмента.

11.6.1. В информированном согласии должно быть в явном виде обозначено требование к тестируемому в ситуации экспертизы (отборочного конкурса, тестового экзамена) выполнить тестовую процедуру до конца; с этим требованием связано предупреждение о том, что выполнение только части тестового испытания ведет к присвоению нулевого или низкого тестового балла<sup>33</sup>).

11.6.2. Информированное согласие основано на полноценном информировании тестируемых о:

- возможности или невозможности добровольного отказа от участия в процедуре оценки на любом из этапов и вероятных последствиях такого отказа;
- целях тестирования при сохранении за организаторами права определять необходимость раскрытия измеряемых параметров и общего характера тестовых заданий;
- факторах, которые могут повлиять на готовность тестируемых участвовать в процедуре тестирования (в частности, о риске, дискомфорте или границах конфиденциальности);
- персонале, осуществляющем тестирование, в том числе об опыте и квалификации специалистов;
- доступности информации (анонимность, частичная или полная открытость полученных результатов для лиц, принимающих решения, включая персональную информацию), предполагающиеся варианты огласки результатов тестирования;
- возможностях использования результатов (вариантах их участия в последующих процедурах оценивания, предназначенных участнику: например, условия найма, продвижения, развития);
- лицах, принимающих решения с точностью до должностей, правилах и правах их доступа к результатам тестирования;

<sup>33</sup> В ряде случаев «Информированное согласие» призвано дать понять тестируемому, что он собирается выступать в роли профессионального исполнителя, а значит, отказывается от части своих прав в пользу организации и тех лиц, которых организация обслуживает, а тест в этих условиях является частью системы профессиональных испытаний (например, фактически знаменует начало испытательного срока). Например, отменяется право человека хранить в тайне информацию о своих физических недостатках, если эти недостатки заведомо мешают выполнять профессиональную деятельность (дефекты зрения, слуха и другие очевидные недостатки) — см. подробнее рекомендуемое руководство «Практическая тестология» (Шмелев, 2013).



- критериях оценки, процедуре сбора, порядке и сроках хранения полученных материалов, их использовании, а также о контактном лице, ответственном за администрирование, фиксацию, хранение и архивирование документов;
- возможностях, формах, сроках получения и содержании обратной связи;
- процедуре повторного тестового обследования (если она предусмотрена).

Организаторы вправе дополнять «Информированное согласие» особыми пунктами и соображениями с учетом специфики организационной задачи и применяемой методики тестирования.

11.6.3. При бланковом проведении теста каждый из пунктов информированного согласия на участие в тестировании должен быть изложен на бумажном носителе, который участник тестирования подписывает лично. При компьютерном и дистанционном онлайн-тестировании документ об информированном согласии предьявляется в электронной форме — как текст, подобный договору-оферте, снабженный интерактивным элементом для выражения согласия с его содержанием. Если в ходе информирования у участников появляются вопросы, не включенные в текст информированного согласия, необходимо разъяснить все аспекты, о которых спрашивают предполагаемые участники, до начала процедуры тестирования. При использовании во время тестирования видео- и аудиозаписи необходимо дополнительно получить от участников информированное согласие на осуществление подобной записи.

11.6.4. Если в информированном согласии не указано обратное, участник тестирования имеет право на получение обратной связи по результатам тестирования (см. главу 10).

11.6.5. При появлении после тестирования новых целей, отличных от заявленных ранее, тестируемые должны быть проинформированы об этом с получением дополнительного согласия на использование имеющихся результатов в новом контексте.

### **11.7. Конфиденциальность и работа с полученными данными**

11.7.1. В целях защиты информации тестовые задания и интерпретирующие заключения (автоматизированные или рукописные отчеты экспертов) сохраняют статус секретных и конфиденциальных. В то же время по требованию участника ему может быть предоставлена информация об используемом методическом инструментарии, основаниях для вынесения рекомендаций, заключений и управленческих решений.

11.7.2. Пользователи тестовых методик и администраторы тестирования должны принимать специальные меры по обеспечению сохранности, секретности и конфиденциальности документов, связанных с процедурными аспектами тестирования. Нарушением конфиденциальности не являются обсуждения конфиденциальной информации на формальных консультациях с коллегами или лицами, принимающими решения.

11.7.3. Вся информация о результатах тестирования также является конфиденциальной. Пользователь тестовых методик не имеет права делиться конфиденциальной информацией с лицами, не имеющими прямого отношения к проведению тестирования и выработке решений по его результатам. Если планируется использование полученных данных в исследовательских и иных целях, это должно быть сообщено участнику тестирования заранее и отражено в информированном согласии.

11.7.4. Если пользователь тестовых методик использует конфиденциальный материал результатов тестирования для обсуждения с коллегами в целях обмена опытом, данный материал должен быть изменен таким образом, чтобы не дать возможность идентифицировать тестируемого.

## Заключение

Данная версия документа «Стандарт тестирования персонала» 1.3 предназначена для обсуждения в профессиональном сообществе психологов-диагностов, разработчиков оценочных методик, практических специалистов по оценке персонала. В период такого обсуждения данная версия 1.3 имеет статус проекта и не может рассматриваться как завершённый нормативный документ. Все вопросы, замечания и предложения, снабженные ссылками на главы и параграфы документа, следует публиковать на форуме «Экспертного Интернет-сообщества психологов» на сайте [forum.ht-line.ru](http://forum.ht-line.ru) в разделе «Разработка профессиональных стандартов».

### ПРИЛОЖЕНИЕ 1. Перечень рекомендуемой литературы

1. Анастаси А., Урбина С. Психологическое тестирование. 7-е международное издание. СПб: Питер, 2001. 688 с.
2. Батулин Н.А., Мельникова Н.Н. Технология разработки психодиагностических методик. Челябинск: Издательский центр ЮУрГУ, 2012. 135 с.
3. Фер М., Бакарк В. Психометрика. Введение. Пер с англ. Челябинск: Издательский центр ЮУрГУ, 2010. 445 с.
4. Шмелев А.Г. Практическая тестология. М.: Маска, 2013. 688 с.

### ПРИЛОЖЕНИЕ 2. Глоссарий основных терминов и сокращений

*HR (Human Resources, дословно «человеческие ресурсы») — в настоящее время аббревиатура HR широко используется для обозначения всей сферы управления персоналом (в том числе оценки).*

*IQ-баллы — баллы на стандартной шкале с параметрами: среднее равно 100, стандартное отклонение равно 15; в первых тестах интеллекта балл 100 интерпретировался как 100-процентное соответствие умственного возраста хронологическому.*

*IRT (Item Response Theory) — теория ответов на тестовые задания, или современная математическая теория тестов.*

*Авторизация пользователя — ввод логина и пароля в компьютерную (автоматизированную) систему тестирования (АСТ). Не следует смешивать авторизацию, регистрацию и аутентификацию.*

*Авторский тест — см. Непсихометрический тест.*

*Административное тестирование — тестирование кандидатов и работников с целью принятия управленческих (кадровых) решений руководителями организации. Синонимы — тестовый экзамен, ассесмент-тестирование, аттестационное тестирование, high stakes assessment, ситуация внешней оценки, ситуация экспертизы, профессиональное испытание. АТ предъявляет более высокие требования к качеству тестовых материалов (психометрическим параметрам теста), а также уровню информационной безопасности (административный контроль поведения тестируемых). От АТ следует отличать самотестирование кандидатов, консультативное тестирование, профориентационное или карьерное тестирование, проблемно-психологическое тестирование и другие виды тестирования, в которых решение по результатам принимает сам тестируемый, а оценивающий специалист предоставляет ему лишь консультативную информацию.*

*Администратор* — специалист в организации, отвечающий за проведение тестирования в отдельных аудиториях или на рабочих местах. В отдельных случаях, когда число участников тестирования и тестовых сеансов невелико, функции администратора может выполнять пользователь тестовых методик. Следует учесть, что эта терминологическая традиция в настоящее время входит в определенное противоречие с тем, что в автоматизированных системах «администратор» — это прежде всего IT-специалист.

*Администрирование теста* — управление процессами подготовки тестирования, проведения и обработки результатов теста.

*Администратор сеансов тестирования* — специалист в организации, отвечающий за проведение тестирования в отдельных аудиториях или на рабочих местах. На практике функции администраторов могут выполнять пользователи тестовых методик, наблюдатели (супервизоры, инспектора), операторы-ассистенты либо экзаменаторы (когда речь идет о тестовом экзамене).

*Альфа-коэффициент (альфа Кронбаха)* — коэффициент внутренней надежности-согласованности тестовых заданий, относящихся к одной шкале — направленных на измерение одного свойства.

*Анкета* — нетестовый опросник с открытыми или смешанными вопросами, который позволяет получать информацию от кандидата (респондента) о нем самом. Почти во все тесты (программы тестирования) включается либо очень краткий (например, ФИО, пол, возраст), либо расширенный анкетный блок.

*АОС* — компьютеризированная система управления обучением (англ. эквивалент LMS — Learning Management System). От русского слова «Автоматизированная обучающая система».

*Ассесмент* — совокупность различных методов оценки персонала, в которых решающая роль принадлежит компетентности оценивающего специалиста (ассессора). Близкие термины — ассесмент-центр, центр оценки, метод экспертной оценки. Нередко на практике метод ассесмента фактически сводится к методу центра оценки, но в общем случае понятие «ассесмент» шире понятия «центр оценки».

*АСТ* — автоматизированная система тестирования (общеупотребительный англоязычный аналог — Test Management System).

*Аутентификация* — процедура проверки подлинности, в случае компьютерного тестирования это проверка подлинности ФИО участника (тестируемого) в дистанционном режиме. Аутентификация, как правило, включает в себя авторизацию пользователя (по логину и паролю), но не ограничивается ей и дополнительно использует биометрические данные, веб-камеры, сеансы видеосвязи и другие средства для подтверждения подлинности личности участника. Не следует путать с регистрацией пользователя.

*Баллирование* — процедура подсчета сырого балла по ключам и перевода сырого балла в стандартный (шкалированный) балл. См. также «мягкое баллирование».

*Банк тестовых заданий* — компьютерное хранилище, в котором собраны тестовые задания по определенной тематике; на основе банка заданий современные программы автоматически формируют в сеансе тестирования варианты для предъявления конкретному тестируемому.

*Валидизация* — проверка валидности. Другой вариант русского перевода англоязычного термина validation — валидация. Перевод «валидизация» чаще используется для обозначения дополнительной работы по выявлению критериальной валидности теста уже в ходе его практического применения, тогда как «валидация» обозначает деятельность разработчика до внедрения теста.

*Валидность* — соответствие результатов тестирования заявленной цели тестирования, в частности, тому свойству (или свойствам), которое измеряется. В — одно из основных

психометрических свойств теста, характеризующих пригодность и обоснованность применения теста в конкретных условиях.

*Внешний критерий валидности* — показатель эффективности производственной (профессиональной) деятельности, относительно которого проверяется валидность теста. Близкие термины — показатель эффективности деятельности, ключевые показатели эффективности, эффективность на рабочем месте.

*Воронка отбора* — один из вариантов учета результатов этапа отбора (в частности, предварительного тестирования), при котором на следующий этап допускаются лишь те участники-кандидаты, которые набрали балл не ниже некоторого порогового значения («точки отсечения»). Другой вариант учета — «мягкое баллирование» (см ниже).

*Выборка нормирования (стандартизации)* — группа тестируемых, на которых рассчитываются тестовые нормы, или производится калибровка шкалы тестовых баллов, то есть устанавливается соответствие между сырыми и стандартизированными баллами.

*Высокая группа* — тестируемые, получившие наиболее высокие баллы по тесту (выше установленной границы). Объем высокой группы и/или граница могут определяться различными способами. Чаще всего это треть (квантиль 0.66) или четверть обследованной выборки (квантиль 0.75, или верхний квантиль).

*Грейды* — качественные градации (ступени, ранги), соответствующие определенным содержательно интерпретируемым интервалам на шкале тестовых баллов. Школьные (или вузовские) оценки — это частный случай грейдов.

*Диагностическая оценка* — отнесение обследуемого к определенной диагностической категории по набору значимых признаков; диагностическую оценку следует отличать от рейтинговой оценки — на одномерной шкале рейтинга.

*Дистанционное тестирование* — в настоящем стандарте это компьютерное сетевое тестирование, при котором процедура тестирования производится на локальном компьютере тестируемого, а обработка результатов — на сервере провайдера тестовой технологии (в частном случае, на сервере разработчика). В настоящее время дистанционное тестирование производится главным образом в режиме «онлайн». Важная разновидность дистанционного тестирования (обеспечивающая надежную аутентификацию тестируемых) — это распределенное многофилиальное тестирование, когда процедура тестирования производится в локальной сети под наблюдением локального администратора группового сеанса тестирования, а обработка результатов производится на сервере в некотором центральном офисе.

*Дистрактор* — неправильный ответ в тестовом задании с готовыми правильными и неправильными вариантами ответов (как правило, в тестах способностей, знаний и профессиональных навыков). Дистракторами также являются варианты ответов в ипсативном опроснике.

*Достоверность* — защищенность тестовой методики от фальсификации и мотивационных искажений; Д — одно из четырех основных психометрических свойств теста.

*Идеальный профиль* — графическая или числовая модель, отражающая наиболее желательное соотношение измеряемых свойств, например, усредненные показатели работника с высокой производительностью труда.

*Информированное согласие* — документ, объясняющий цели, принципы и условия планирующегося обследования, который участник подписывает до начала тестирования.

*Кандидат (соискатель)* — физическое лицо, участвующее в серии оценочных процедур в рамках конкурса на вакансию. В контексте отбора персонала и рекрутинга термин «кандидат» превращается в синоним термина «участник тестирования (тестируемый)», но в

общем случае термин «участник тестирования» шире, так как может относиться к действующим работникам.

*Кастомизированный тест* — универсальный тест, прошедший определенную адаптацию, настройку или модификацию для специализированного использования в определенной организации. Кастомизированные тесты следует отличать от локальных корпоративных тестов, созданных изначально внутри организации сотрудниками этой организации с целью применения только в рамках данной организации.

*Квалификационные тесты* — тесты на профессиональные знания и умения. КТ не являются психодиагностическими тестами, но входят в область практической тестологии — практического тестирования персонала.

*Квазитест* — см. Непсихометрический тест.

*Кейс-тест* — тест, в котором тестовые задания содержат развернутые описания жизненных (профессиональных) проблемных ситуаций (случаев), а варианты ответа представляют собой возможные способы действия в этих ситуациях или оценки этих ситуаций. На Западе чаще используется название «ситуационный тест» (situational judgement test, или SJT).

*Ключи к тесту* — правила, связывающие ответы на задания и тестовые шкалы; с помощью ключей осуществляется подсчет тестовых баллов по шкалам.

*Компетенции* — эталоны (образцы) успешного профессионального поведения или значимые для эффективной работы характеристики работника, включающие его знания, умения, навыки, способности, мотивационную направленность и личностные черты, — все свойства, которые оказываются важными для эффективной профессиональной деятельности в определенной организации на определенной позиции (должности).

*Комплексная технология оценки* — совокупность оценочных процедур, в которой сочетаются тестирование и различные виды экспертных процедур (интервью, деловые игры, имитационные упражнения и т.п.). Близкие термины — центр оценки (ЦО), ассесмент. Но в отличие от ЦО и ассесмента, в КТО экспертные оценки необязательно играют решающую роль: тестирование (как и показатели эффективности реальной деятельности) может входить в комплексную оценку с тем же весом, что и экспертные оценки.

*Компьютерное тестирование* — выполнение теста в режиме диалога с компьютером.

*Конверсионная таблица* — таблица перевода сырых баллов в стандартизированные баллы, оценки или грейды в том случае, когда линейное нормирование (перевод в z-баллы) не является корректной процедурой).

*Консультативное тестирование* — тестирование, которое проводится по инициативе самого тестируемого и сам тестируемый распоряжается результатами тестирования. Важно различать КТ и так называемое «административное тестирование» (см).

*Конструктивная валидность* — количественная мера валидности (пригодности) теста, полученная путем подсчета коэффициентов корреляции теста с другими тестами или оценочными процедурами. Следует отличать КВ от критериальной (прагматической) валидности теста, когда подсчитывается корреляция теста прежде всего с «ключевыми показателями эффективности» (КПЭ).

*Корпоративная нормативная выборка* — выборка протестированных работников (или кандидатов), на которых рассчитываются (пересчитываются) локальные тестовые нормы. Желательная численность КНБ — более 300 человек. Минимальная (для малочисленных профессиональных групп) — 100 человек.

*Корпоративный тест* — локальный тест, разработанный или адаптированный (кастомизированный) внутри производственной компании (учреждения, общества), не специали-

зирующей на разработке тестов, и предназначенный для внутреннего применения (но не для коммерческого распространения) — для кандидатов и работников данной компании. К КТ близки по своему статусу и уровню психометрической разработки так называемые кастомизированные тесты. Другой родственный термин — непсихометрические тесты.

*Корреляция / коэффициент корреляции* — мера статистической связи двух показателей. КК получает значения на отрезке от  $-1$  до  $+1$ , что означает либо максимально сильную отрицательную связь, либо максимально сильную положительную связь. Например, высокая (положительная или отрицательная) корреляция между результатами теста и показателями КПЭ говорит о возможности практического применения теста для предсказания эффективности деятельности работника (кандидата).

*Коэффициент дискриминативности (КД)* — коэффициент, указывающий на качество (надежность) отдельного тестового задания; КД лежит в пределах от  $-1$  до  $1$  и указывает на степень связи между правильным решением конкретного задания и успешностью выполнения теста в целом.

*КПЭ* — ключевые показатели эффективности (от англ. KPI — Key Performance Indicators). Часто выступают основным критерием для оценки практической валидности (пригодности) тестов. Близкий синоним — «показатель эффективности деятельности» (ПЭД).

*Лицензионный тест (коммерческий тест)* — тестовая методика, разработанная, как правило, профессиональным разработчиком, право на которую пользователь теста приобретает согласно лицензионному договору. Большинство психометрических (измерительных) тестов являются лицензионными, так как предполагают серьезные затраты разработчика (правообладателя) на их создание. От лицензионных тестов следует отличать «свободно распространяемые», а также корпоративные тесты.

*Локализация* — приспособление теста к местным условиям (включая языковой перевод) силами разработчика теста. Близкое понятие — «кастомизация», но в случае кастомизации эту работу могут выполнять совместно как разработчики, так и сами пользователи теста

*Логит* — единица измерения трудности тестовых заданий и способности тестируемого в современной теории тестов (IRT).

*ЛПР* — лица, принимающие решения, или руководители в организациях.

*Мета-анализ* — определенный метод интеграции статистических результатов, полученных разными исследователями на разных выборках.

*Мягкое баллирование* — учет результатов тестирования в интегральном рейтинге наряду с другими оценочными процедурами (путем суммирования тестовых баллов с определенным весом). Другой вариант учета результатов тестирования — «воронка отбора».

*Надежность* — мера защищенности результатов теста от случайных колебаний; Н — одно из четырех основных психометрических свойств теста.

*Непсихометрический тест* — тест, который базируется на экспертной интуиции автора и не проходит проверку на соответствие психометрическим требованиям надежности и валидности. Другие и родственные термины — квазитест, авторский тест, самодеятельный тест, интуитивный тест.

*Нормальное распределение* — частотное распределение тестовых баллов, соответствующее нормальному (гауссовскому закону): кривая нормального распределения, которая ставит в соответствие величину тестового балла (по абсциссе) и частоту встречаемости (по ординате), имеет характерную колоколообразную форму.

*Обратная связь* — информация о результатах тестирования и/или других оценочных процедур, предоставляемая участнику.

*Опрос* — эмпирическая вопросно-ответная процедура, позволяющая собрать статистику ответов по группе респондентов, но в общем случае не позволяющая приписать числовую характеристику индивидуальному респонденту. Не следует путать с тест-опросником.

*Очное административное тестирование* — тестирование, которое происходит под контролирующим наблюдением администратора, присутствующего в аудитории. Следует различать ОАТ и дистанционное тестирование с удаленным контролем со стороны администратора (через систему удаленного видеонаблюдения и т.п.).

*Ошибка измерения* — диапазон разброса в тестовых показателях, возникающий при повторном выполнении одного и того же теста одним и тем же тестируемым.

*Показатель эффективности деятельности* — независимая от результатов тестирования оценка участника, которая рассматривается как объективная в соответствующей профессиональной среде.

*Пользователь тестовых методик* — это специалист по оценке персонала, отвечающий за выбор методик тестирования, соответствующих целям и условиям оценки, разработку программы тестирования и анализ результатов; часто в его задачи также входит организация тестирования, внедрение и отладка комплексных тестовых систем в организации.

*Прогностическая валидность* — способность теста предсказать будущую эффективную работу, то есть выделить группу работников (кандидатов), которые будут успешными исполнителями.

*Процентаж* — доля нормативной выборки тестируемых, получивших тестовые баллы на определенном интервале тестовой шкалы. Например, процентаж «высокой группы» составляет чаще всего либо 25%, либо 33% (правая или верхняя часть нормативной выборки).

*Процентильный балл* — доля выборки в процентах, то есть процент лиц из числа протестированных, набравших более низкий сырой балл по тесту, чем набрал данный тестируемый. Близкий термин — «процентаж».

*Психодиагностика* — наука о распознавании выраженности психических свойств у человека; часть тестов, направленных на оценку персонала в организации, являются психодиагностическими, но «квалификационные тесты» чаще всего не являются психодиагностическими.

*Психометрические свойства теста (или психометрические требования)* — надежность, валидность, репрезентативность и достоверность.

*Психометрический тест* — тестовая методика, разработчик которой выполнил проверку психометрических свойств и показал удовлетворительные значения надежности, валидности и репрезентативности.

*Разработчик тестовых методик* — прикладной математик или психолог, разрабатывающий современные измерительные (психометрические) тесты с использованием специализированных компьютерных инструментов конструирования тестов.

*Распределение тестовых баллов* — частоты встречаемости тестовых баллов среди всех прошедших тестирование (в нормативной выборке). Распределение может описываться в табличной и графической форме.

*Регистрация пользователя* — создание учетной записи пользователя в компьютерной системе (АСТ). Регистрация предусматривает создание логина и пароля и предшествует авторизации (вводу готовых логина и пароля). Регистрация задает роль пользователя в системе АСТ (например, роль тестируемого, или роль тестирующего администратора, или роль лица, принимающего решение и имеющего доступ к результатам и т.п.).

*Режим секретности* — режим распространения тестовых материалов (тестовых заданий), исключающий преждевременное знакомство тестируемых с этими заданиями.

Режим секретности обеспечивается техническими и административно-юридическими инструментами. В юридической литературе «режим секретности» часто именуется как «режим ноу-хау» (гражданский кодекс РФ).

*Режим конфиденциальности* — ограниченный доступ к результатам тестирования только для лиц, подписавших протокол конфиденциальности, либо самих тестируемых, подписавших информированное согласие. Этот режим, в частности, исключает публикацию результатов тестирования под фамилией тестируемого без его согласия.

*Рейтинг-балл* — итоговый результат комплекса оценочных процедур, определяющий место каждого обследованного в ранговом ряду (в рейтинг-листе). Простейший способ получения РБ — суммирование нормированных результатов, полученных по каждой оценочной процедуре. РБ — это способ интеграции тестовых и нетестовых методик. РБ — это показатель на рейтинговой шкале.

*Рейтинговая шкала* — система значимых для организации критериев, которая используется для упорядочивания объектов оценивания (в частности, работников, учащихся и т.п.) в соответствии с их ценностью относительно этой системы. Следует различать рейтинговые шкалы и диагностические оценки.

*Репрезентативность* — мера соответствия тестовых норм, полученных на нормативной выборке (выборке стандартизации), тому контингенту, на котором проводится тест; Р — одно из четырех основных психометрических свойств теста.

*Респондент* — лицо, которое отвечает на анкету, опросник или тест-опросник; в случае тестов знаний или способности говорят о тестируемом, или участнике тестирования. Правильнее говорить, что тестовые задания выполняют, а не «отвечают на вопросы», ибо в тесте задания часто не имеют формата вопросов (например, текст с пропусками для заполнения).

*Ретестовая надежность* — устойчивость результатов тестирования во времени; она измеряется путем расчета корреляции между результатами первого и повторного тестирования одной и той же группы тестируемых (методика измерения надежности «тест-ретест»).

*РПО* — Российское психологическое общество. Являясь наиболее авторитетной национальной профессиональной ассоциацией, РПО обеспечивает независимую сертификацию (оценку качества) психодиагностических тестовых методик. Временный регламент требований к психодиагностическим методикам разработан специальной Комиссией РПО.

*Самодетальный тест* — см. Непсихометрический тест.

*Самотестирование* — процедура тестирования, получившая распространение в недавнее время — с началом массового применения дистанционного онлайн-тестирования в глобальных сетях (в Интернете). Эта процедура, как правило, не требует высоких стандартов информационной безопасности, так как кандидат может проходить тестирование под псевдонимом, может сам принимать решения на основе своих результатов, может сам отказаться от использования своих результатов.

*Система сертификации* — совокупность организаций и мероприятий, которые подтверждают соответствие тестовых методик профессиональным стандартам; сертификация пользователей тестов подтверждает квалификацию специалистов в области профессионального применения тестов.

*Ситуация консультирования* — ситуация обследования, которая возникает по инициативе самого обследуемого, являющегося одновременно и владельцем результатов тестирования и лицом, принимающим решение. Родственные термины: ситуация клиента, конфиденциальное тестирование.



*Ситуация экспертизы* — ситуация, когда обследование проводится по инициативе третьих лиц, и главным пользователем его результатов оказываются лица, принимающие решения в интересах организации, а не в интересах обследуемого. Родственные термины: административное тестирование, аттестационное тестирование, high-stakes assessment.

*Содержательная валидность* — один из ключевых видов валидности теста, связанный с экспертными оценками соответствия между содержанием тестовых заданий и профессиональной деятельностью тестируемых. Авторское обоснование такого соответствия следует рассматривать как частный случай экспертных оценок такого рода (автор — это один из экспертов).

*Социальная желательность* — это установка большинства тестируемых (респондентов) давать при выполнении тест-опросников так называемые «социально-одобряемые ответы».

*Стандартизация теста* — сбор обширного банка данных с целью измерения репрезентативных тестовых норм; после стандартизации в процедуру теста уже запрещается вносить модификации — вплоть до пересмотра и создания новой версии методики.

*Стандартизированные баллы* — это тестовые баллы на стандартных шкалах, соответствующие определенным сырым баллам, полученным определенными участниками тестирования. Стандартизированные баллы являются лучшим приближением к истинным баллам (теоретически-возможным с идеальной точностью), но отличаются от истинных баллов в силу ошибки измерения.

*Стандартные шкалы* — шкалы, для которых искусственно заданы фиксированные значения среднего и стандартного отклонения, а распределение тестовых баллов подчинено нормальному закону. Примеры стандартных шкал — шкале стэнов от 1 до 10, шкала Т-баллов, шкала IQ. Тестовые баллы на стандартных шкалах называются стандартизированными или нормированными.

*Стэны* (от англ. *sten = standard ten*) — баллы на стандартной шкале «стандартная десятка» (от 1 до 10). Среднее значение равно 5.5, стандартное отклонение — 2.

*Сужение выборки по критерию* (англ. *range restriction*) - в том случае, если критериальная валидность теста проверяется только по группе «высокопроизводительных работников», а «низкая группа по критерию» никак не рассматривается, результаты проверки критериальной валидности оказываются заниженными.

*Супервизор* — в настоящем стандарте уполномоченный, подготовленный наблюдатель в аудитории, где выполняется тестирование.

*Сырой балл* — суммарный балл, полученный тестируемым согласно его или ее ответам, рассчитанный в соответствии с ключами. Сырой балл невозможно интерпретировать без сравнения с тестовыми нормами или без перевода в стандартную тестовую шкалу.

*Т-баллы* — баллы на стандартной шкале с параметрами: среднее равно 50, а стандартное отклонение равно 10. Часто используется тысячебалльный вариант Т-шкалы, на которой среднее равно 500, а стандартное отклонение — 100.

*Тест / тестовая методика / методика тестирования* - методика оценки индивидуальных свойств работника (кандидата), которая производится по стандартизированной процедуре (опирается на стандартный, одинаковый для всех набор заданий) и выдает результат в количественной форме — в виде тестового балла на количественной шкале (шкалах).

*Тестовая шкала / шкала теста / измерительная шкала* — шкала количественных результатов тестирования (тестовых баллов). Важно различать шкалу «сырых баллов» и шкалу стандартизированных баллов. В случае легкого теста средний показатель на шкале «сырых баллов» смещается к правому краю шкалы, а в случае трудного — к левому. На стандартной шкале средний балл всегда оказывается в центре шкалы, независимо от трудности теста.

*Тестовые материалы* — материалы теста, включающие тестовые задания, предтестовую инструкцию, параметры количественной обработки (ключи и нормы), тексты интерпретирующих сообщений (систему интерпретации). На тестовые материалы распространяется режим секретности (не путать с режимом конфиденциальности).

*Тестовые нормы* — граничные значения на шкале тестовых баллов, отделяющие интервалы с различной интерпретацией.

*Тест-опросник* — это тестовая процедура, в которой индивидуальные результаты выражаются на шкале тестовых баллов, а тестовые задания формулируются как вопросы, касающиеся характера, привычек, установок и предпочтений человека; в ТО, в отличие от тестов на способности и знания, не бывает однозначно правильных или неправильных ответов, а баллы за ответ начисляются в соответствии с близостью к ключевому ответу.

*Точка отсечения (англ. cut-off score)* — точка, разделяющая всех тестируемых, выполнивших тест, на две категории (например, успешные и неуспешные). Синонимичные названия для ТО — «порог отсева» или «порог отбора».

*Участник тестирования* — в тексте настоящего стандарта физическое лицо, которое выполняет тест, то есть тестируемый. Использование термина «участник» в контексте оценки персонала является предпочтительным и более корректным, чем термин «испытываемый». В частном случае уместно называть тестируемого термином «респондент» (в случае применения тест-опросников).

*Центр оценки (ЦО)* — комплексный метод диагностики потенциальной успешности в профессиональной деятельности, включающий набор различных методов, основой которого является оценка поведения участников группой наблюдателей-экспертов в моделирующих упражнениях. Федерация оценки персонала разработала отдельный «Стандарт центра оценки», который наряду с данным «Стандартом тестирования персонала» является основой для профессиональной деятельности специалистов по оценке персонала.

*Шкала лжи* — особый количественный инструмент для проверки уровня достоверности каждого индивидуального протокола. Чаще всего шкала лжи в тест-опросниках состоит из заданий, провоцирующих социально-желательные ответы. Высокий балл по шкале лжи позволяет интерпретировать часть протоколов тестирования как недостоверные.

*Шкалирование* — это процедура, в результате которой устанавливается правило отображения сырых баллов в стандартные на некой стандартной шкале. Близкий термин — стандартизация. Другое возможное значение термина «шкалирование» связано с применением особых «методик шкалирования»: респондент в этом случае дает субъективные оценки определенным объектам по определенному набору шкал.

*Шифрокод* — буквенно-цифровой код, который может быть использован в автоматизированной системе тестирования (АСТ) вместо фамилии и инициалов тестируемого. В отличие от паролей шифрокоды могут быть опубликованы (например, с целью конфиденциальной групповой обратной связи о результатах тестирования для самих тестируемых; такова многолетняя практика вступительных экзаменов во многих вузах). При этом соответствие шифрокода и ФИО тестируемого является секретом для всех пользователей АСТ, кроме тех, кто имеет специальные полномочия — владеет таблицей соответствия, которая может храниться вне самой АСТ. Использование шифрокодов снимает с организации, являющейся внешним хостинг-провайдером АСТ, необходимость обработки идентификационных персональных данных (ФИО), которые остаются в кадровых системах организации. На шифрокодах может базироваться анонимное дистанционное самотестирование. Аналогом шифрокода является логин, не позволяющий идентифицировать пользователя (маскирующий логин). Не следует путать шифрокоды с пин-кодами (персо-

нальный идентификационный код), так как пин-коды часто используются в банковской системе как функциональный аналог пароля. Подробнее об опыте применения шифрокодов в АСТ см. рекомендуемую книгу «Практическая тестология» (Шмелев, 2013).

*Экспертная оценка* — оценочная процедура, в ходе которой либо свойства человека, либо эффективность его деятельности оцениваются людьми, знающими оцениваемого или наблюдающими его поведение. Определенный субъективизм ЭО компенсируется участием нескольких независимых друг от друга экспертов. Очень часто ЭО выставляются по итогам интервью, деловых игр, имитационных упражнений, и в этом качестве ЭО дополняют и результаты тестирования и позволяют их перепроверить.

### ПРИЛОЖЕНИЕ 3. Некоторые важнейшие расчетные формулы

#### 1. Формула перевода сырых баллов в стандартизированные

Формулы перевода в общем случае выглядят так (для упрощения разобьем одну формулу на две простые, применяемые последовательно):

$$\frac{z-x-\bar{x}}{s} \quad (1.1)$$

$$Y = Z \cdot St + Mt \quad (1.2)$$

где  $X$  — сырой балл тестируемого,  $\bar{X}$  — эмпирическое среднее арифметическое сырых баллов на нормативной выборке (должно быть указано в описании теста),  $S$  — эмпирическое стандартное отклонение на нормативной выборке (должно быть указано в описании теста),  $Z$  — стандартизированный балл испытуемого на  $Z$ -шкале (шкале с параметрами: среднее  $Mz = 0$ , отклонение  $Sz = 1$ ),  $Mt$  — теоретическое среднее на избранной стандартизированной шкале (см. табл. 7.1),  $St$  — теоретическое стандартное отклонение на избранной стандартизированной шкале (см. табл. 7.1),  $Y$  — стандартизированный балл тестируемого в избранной шкале.

#### 2. Расчёт ошибки измерения на основе надёжности

Стандартная ошибка измерения — относительная доля случайного изменения (дисперсии) измеряемого показателя по отношению к совокупному изменению этого показателя (общей дисперсии). Из общей логики, чем меньше строгость измерения (т.е. допустима большая ошибка измерения), тем ниже надёжность. Рассмотрим формулу, связывающую эти два показателя:

$$R = 1 - \frac{S_e^2}{S^2} \quad (2.1)$$

$R$  — надёжность измерения (должна быть указана в описании теста),  $S$  — эмпирическое стандартное отклонение на нормативной выборке (должно быть указано в описании теста),  $S_e$  — стандартная ошибка измерения.

Исходя из этой формулы можно выразить стандартную ошибку измерения через надёжность и стандартное отклонение.

$$S_e = S\sqrt{1-R} \quad (2.2)$$

С вероятностью 95% истинное значение тестового балла тестируемого лежит в интервале

$$(X - 1,96 \sqrt{1-R}; X + 1,96 \sqrt{1-R})$$

Для определения стандартной ошибки разности двух показателей 1 и 2 необходимо пользоваться формулой

$$S_{e_{12}} = \sqrt{S_{e_1}^2 + S_{e_2}^2} \quad (2.3)$$

Где  $S_{e_i}$  — стандартная ошибка измерения для показателя  $i$  ( $i = 1$  или  $2$ ).

Поскольку чтобы сравнивать два показателя, необходимо перевести их в одну стандартную шкалу, имеющую одинаковое стандартное отклонение, то воспользовавшись формулой 2.2, можно представить стандартную ошибку разности двух показателей как

$$S_{e_{12}} = S \sqrt{2 - R_1 - R_2}$$

где  $R_i$  — надежность измерения показателя  $i$  ( $i = 1$  или  $2$ ).



## **Russian Standard for Personnel Testing (Interim version, designed for a discussion)**

**Nikolai BATURIN** (*Psytest, Chelyabinsk, Russia*)

**Evgenii VUCHETICH** (*Exect Partners Group, Moscow, Russia*)

**Svetlana KOSTROMINA** (*St. Petersburg State University, St. Petersburg, Russia*)

**Boris KUKARKIN** (*Exect Partners Group, Moscow, Russia*)

**Evgeny KUPRIYANOV** (*Higher School of Economics, Moscow, Russia*)

**Evgeny LOURIE** (*ECOPSY, Moscow, Russia*)

**Olga MITINA** (*Lomonosov Moscow State University, Russia*)

**Anna NAUMENKO** (*Independent Consultant, London, UK*)

**Ekaterina OREL** (*Higher School of Economics, Moscow, Russia*)

**Yulia POLETAEVA** (*SHL, Moscow, Russia*)

**Alexey POPOV** (*ECOPSY, Moscow, Russia*)

**Alex POTAPKIN** (*Independent consultant, Moscow, Russia*)

**Svetlana SIMONENKO** (*Detech, Moscow, Russia*)

**Julia SINITSINA** (*TalentQ, Moscow, Russia*)

**Alexander SHMELEV** (*Humanitarian technologies, Moscow, Russia*)

**Absctact.** In recent decades Russian companies have seen a widespread use of psychometric testing as part of personnel assessment procedures; however, until now there have been no single, universally accepted document containing the list of requirements and criteria for assessing the quality of testing. Russian standard for personnel testing was established in 2014—2015, on the initiative of the National Confederation of “Human Capital Development” and supported by the expert community. This standard is intended to clarify and popularize the Russian speaking space of one of the basic principles of shared responsibility: the test is responsible for the quality of the developer, and for the quality of the test — the user of techniques. This standard is developed in the period of information and computer revolution, which has led to the proliferation of new types of testing — especially remote online testing has become particularly relevant for large, geographically distributed companies. The standard is designed to solve this problem at the level common to all of the professional community of methodical, technological and ethical requirements for the testing procedures and regulations for the use of their results. In the development of the Russian standard takes into account the recommendations of national and international guidelines for the development and application of test methods. This first edition of the Russian standard for personnel testing — a generally applicable practice-oriented methodological guidance. It lays the foundation for further development of more detailed and highly specialized standards, and is the basis for “professional standards for personnel evaluator” along with the standards of other valuation methods and procedures.

**Keywords:** Russian standard for personnel testing; testing; online testing; test methods; psychological testing; evaluation of staff; psychometric tests; reliability; validity; professional selection; Tests of professional knowledge.

## **Chapter 1. Characteristics of the Russian Standard for Personnel Testing**

### **1.1. Rationale for developing the Standard**

In recent decades Russian companies have seen a widespread use of psychometric testing as part of personnel assessment procedures; however, until now there have been no single, universally accepted document containing the list of requirements and criteria for assessing the quality of testing. It is obvious that incorrect use of test methods leads to serious negative consequences both for test takers and for organizations using test results in making managerial decisions of various scale and significance. Realization of the potential of specific test methods depends on how well and appropriately they are being used; errors in the application invalidate the results even when quality methods are being used.

The Standard is designed to explain and promote one of the basic principles of shared responsibility: the developer is responsible for the quality of the testing instrument and the test user is responsible for the quality of test administration. This standard is being developed in the era of information and computer revolution, which has already led to the proliferation of new testing technologies, mainly, remote online testing becoming particularly relevant for large, geographically distributed companies. The natural emergence of terminology to describe new processes, procedures and types of testing (e.g., remote online testing) marks the urgent need for unification of the concepts used in order to improve mutual understanding between test developers and test users.

The Standard is designed to educate test users at the level of methodical, technological and ethical requirements for the testing procedures and regulations for the use of test results, common to all of the professional community.

### **1.2. Prototypes used**

The Russian Standard took into account national and international guidelines for the development and use of test methods, such as:

- Interim Standard for Psychological Tests accepted in 2011 by the Psychometric Testing Section of the Russian Psychological Society (Baturin, Melnikova, 2012);
- International Test Commission. ITC Guidelines on Test Use (2013);
- British Psychological Society. Guidelines for the Development and Use of Computer-Based Assessments (2002);
- European Federation of Psychologists' Associations. EFPA Review model for the description and evaluation of psychological and educational tests — Test review form and notes for reviewers (2013);
- New Zealand Psychologists Board. Guidelines on the use of Psychometric Tests (2013);
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for educational and psychological testing (1999).
- Considered standards testify the experience and practice of test use gained in Europe, USA, Russia and other countries.

This first edition of the Russian Standard for Personnel Testing is a generally applicable, practice oriented methodological guidance. It lays the foundation for further development of more detailed and highly specialized standards and forms the basis for "Professional Standard for Personnel Assessment Professionals" along with other standards of assessment methods and procedures.

### **1.3. Status of the current Standard**

The Standard is a scientifically sound document. The standard represents not a legal, but a professional, cultural and ethical norm, i.e. professional community is invited to accept the rules

and requirements of the Standard voluntarily. All provisions constitute guidelines for practitioners assessing candidates and employees within an organization or providing personnel assessment services to other organizations. The document contains recommendations which, if ignored, can lead to distortion or loss of quality of the test results obtained in organizations.

The first version of the Standard was developed in 2014—2015 by a special commission created within the Personnel Assessment Federation at NC HCD (National Confederation of Human Capital Development). The commission included both leading academic experts and representatives of business organizations, developing and distributing psychological tests and providing assessment services with the use of test technologies. Currently, the document is adopted as an intercorporate standard for all assessment providers that are members of the Personnel Assessment Federation. In the future, the document is expected to acquire national status.

#### **1.4. Purpose and objectives of the Standard**

Formation and prevalence of modern scientifically sound concepts of psychometric testing<sup>34</sup> and correct use of test results.

Regulation of test use activities.

Providing information about the rules increasing efficiency of psychometric testing.

Regulatory support of test use activities and professional arbitration<sup>35</sup>.

Informational support of the quality assessment of psychometric tests.

Informational support for decision-makers considering using psychometric tests.

Promoting applied research validating the effectiveness of psychometric testing.

#### **1.5. Target audience of the Standard**

1.5.1. The Standard is addressed to those who:

- are test users;
- make decisions about the use of psychometric testing;
- advise and examine the development and application of psychometric tests;
- sell psychometric tests as a service, that is conduct projects based on testing;
- develop and sell psychometric tests as a product;
- research the effectiveness of psychometric tests;
- provide organizational and technical conditions for testing;
- sit psychometric tests.

1.5.2. The chapters of the document intentionally contain a number of repetitions, as the Standard addresses not only to a small group of expert psychometricians but also to other professionals who are not expected to read the entire text of the document. The repetitions allow referring to the individual chapters without having to read the entire Standard. For example, Chapter 8 is addressed not only to psychometricians, but also to IT specialists and it covers the issues discussed in Chapter 6 from a slightly different angle. Some paragraphs of Chapters 3 and 6 are also interrelated and contain repetitions.

In cases when test developers and local test administrators are essentially different groups of people test developers are advised to start with Chapter 3 and test administrators — with Chapter 6. Chapter 11 is addressed to the members of conflict commissions set to deal with possible complaints of test takers, and this chapter overlaps with Chapters 3, 5 and 6.

<sup>34</sup> Hereinafter, by default testing refers to candidates and employees. In other cases, special notes or comments are included.

<sup>35</sup> “Professional arbitration” is the settlement of disputes arising between test users (clients) and test developers, as well as between employees sitting tests and organizations that make decisions based on test results.

The repeats also facilitate the application of the Standard requirements at different stages of test delivery: when preparing testing programme it is enough to reread Chapter 3, when actually administering tests it is OK to read Chapter 6, when giving feedback — Chapter 10, etc. Most of the recommendations on reading order are summarized in Table 1.1.

Table 1.1 Reading recommendations for the Standard for Personnel Testing

Professional Groups	Recommended chapters
Test users	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11
Test administrators	3, 5, 6, 8
Supervisors	1, 2, 3, 4, 7, 9, 10, 11
Test developers	1, 2, 4, 5, 6, 7
Subject matter experts	2, 5, 10
Software developers	3, 6, 8, 9
IT specialists (system administrators, network administrators)	8
HR decision makers	9, 10, 11

### 1.6. Implementation of the Standard provisions: mandatory and desirable

Most of the provisions of the Standard should be seen as guidelines, that is as desirable but not mandatory professional standards. Exceptions are those which are directly dictated by the current Russian legislation. Following these guidelines improves the quality of testing procedures and obtained results, and ignoring them reduces the quality of testing down to a critically low level.

This standard does not imply that tests and testing activities could be deemed as either “suitable” or “unsuitable”. This standard is intended to help divide psychometric tests and testing activities into a higher number of qualitative categories: the more requirements of the Standard are implemented, the higher the quality of the test and relevant testing programme. Therefore, this Standard does not prohibit the use of tests of low quality, but it helps to rank tests by quality. To assess the quality of a test in question it is recommended to at least roughly estimate the percentage of test properties conforming to this Standard. If more than half of the properties does not comply with the Standard requirements it might be considered as a serious argument to discontinue the contract with the test provider.

When given a choice between several potential test providers, each of which complies with the majority (more than 50%) of the Standard requirements, in terms of quality the one who complies with the higher number of requirements (i.e. demonstrating more congruence with the Standard) should be preferred.