



EFPA

STANDING COMMITTEE ON

TESTS AND TESTING (SCTT)

**МОДЕЛЬ РЕЦЕНЗИРОВАНИЯ (ОПИСАНИЯ И ОЦЕНКИ)
ПСИХОЛОГИЧЕСКИХ ТЕСТОВ, РАЗРАБОТАННАЯ EFPA**

**ОБРАЗЦЫ ДОКУМЕНТОВ (ФОРМЫ) И ПОЯСНЕНИЯ
ДЛЯ РЕЦЕНЗЕНТОВ**

ВЕРСИЯ 3.42

Модель рецензирования (описания и оценки) психологических тестов, разработанная EFPA ¹

Образцы документов (формы) и пояснения для рецензентов ²

**ВАЖНО, чтобы рецензент прочитал внимательно все пояснения в
процессе выполнения рецензии.**

Оригинальная версия составлена и отредактирована Дейвом Бартрамом (Dave Bartram)
Настоящая версия – результат пересмотра и корректировки оригинальной версии,
произведенных Патрицией Линдли (Patricia Lindley), Дейвом Бартрамом (Dave Bartram) и
Натали Кеннеди
(Natalie Kennedy) ³
Русский перевод – Н.А. Батурин, С.В. Устьянцева
Версия 3.42: 2005г.

¹ Постоянный комитет по тестам и тестированию EFPA одобрил и эти пояснения (заметки) для рецензентов и соответствующий формат рецензирования. Члены психологических ассоциаций могут использовать их как основу для рецензирования собственного инструментария. Желание сделать это широко доступным вдохновляет на то, чтобы привести в соответствие оценочные процедуры и критерии, сделав их едиными для всей Европы. Комментарии по поводу этих документов приветствуются: мы надеемся, что опыт рецензентов будет способствовать улучшению и прояснению процессов.

² Этот документ составлен на основе ряда источников, включая «Модель Рецензирования и Оценки Тестов», созданную BPS (the BPS Test Review Evaluation Form) (NPAL, и Комитет по Тестовым Стандартам); «Испанский Вопросник для Оценки Психометрических Тестов» (the Spanish Questionnaire for the Evaluation of Psychometric Tests) (Испанская Психологическая Ассоциация) и Система Оценки Качества Теста (the Rating System for Test Quality) (Комитет по Тестированию Ассоциации Психологов Дании). Часть содержания заимствована и адаптирована с разрешения BPS: Books Reviews of Level B Assessment Instruments for use in Occupational Assessment, Notes for Reviewers: Version 3.1. December 1998: Copyright © NPAL, 1989, 1993, 1998. (по словам так: BPS книги обзоры (рецензии) уровня В оценивание инструменты для использования в профессиональном рецензировании. Заметки для рецензентов: Версия 3.1. Декабрь 1998: Авторское право © NPAL, 1989, 1993, 1998.)

³ Настоящая версия является объединением двух отдельных документов (Форма (для рецензирования и Заметки для рецензентов). Вдобавок содержание было отредактировано и усовершенствовано с учетом опыта использования этой версии BPS для рецензирования тестов в режиме «Online».

В случае каких-либо вопросов, пожалуйста, обращайтесь к местному редактору-консультанту.

По поводу деталей процедуры рецензирования или каких-либо местных изменений (модификаций) содержания и критериев рецензирования обращайтесь в вашу местную Психологическую Ассоциацию.

EFPA сообщает, что оценочные процедуры в этой модели рецензирования ориентированы скорее на квалифицированных практических пользователей тестов, чем на психологов академической направленности или специалистов в области психометрики и психологического тестирования.

EFPA сообщает будущим пользователям следующую информации о происхождении данного документа:

“Критерии оценки тестов EFPA были большей частью разработаны на основе форм и содержания критериев оценки тестов Британского Психологического Общества (BPS) и критериев, разработанных Комитетом по Тестам (COTAN) Ассоциации Психологов Дании (NIP). Первоначально Дейв Бартрам и Патриция Линдлей разработали критерии и оценочные процедуры для тестов Службы Занятости Объединенного Королевства и позднее они были приняты к применению Британским Психологическим Обществом. Арн Эверс (Arne Evers) отредактировал Датскую систему ранжирования тестов с точки зрения качества. EFPA благодарна BPS и NIP за разрешение использовать их критерии в целях разработки Европейской модели. EFPA также благодарна Дейву Бартраму, Арну Эверсу и Патриции Линдлей за их вклад в создание этой модели. Все права интеллектуальной собственности на оригинальные критерии признаются и сохраняются за этими организациями.”



.....

Раздел 1:

Описание инструмента: Общая информация и классификация

.....

В первом разделе следует представить общую информацию, необходимую для идентификации инструмента и сообщить о том, где можно достать этот инструмент. Следует дать название инструмента, сообщить имя издателя и/или дистрибьютора, имя автора (авторов), дату публикации оригинала и дату той версии, которая рецензируется.

Разделы с 1.1 по 1.9 должны быть простыми по стилю. Они содержат фактическую информацию, хотя потребуются подумать о том, как представить полную информацию.

EFPA 3.2 ссылка

	Рецензент 1:	
	Рецензент 2:	
	Редактор-консультант:	
	Старший редактор:	
	Старший корректор: (требуется только для новых модификаций)	
	Корректор: (требуется только для новых модификаций)	
	Дата настоящей рецензии:	
1.1	1.1. Название инструмента (локальная версия):	
	Краткая версия названия теста:	
1.2	Название теста-оригинала (если локальная версия является адаптацией):	
1.4	Авторы теста-оригинала:	
1.3	Авторы локальной адаптации:	
1.7	Локальный дистрибьютор/издатель:	
1.8	Издатель оригинальной версии теста (на тот случай, если издатель настоящей версии другое лицо):	
1.9.1	Дата публикации настоящей новой	

	редакции (исправленной версии):	
1.9.2	Дата публикации версии, адаптированной для локального применения:	
1.9.3	Дата публикации теста-оригинала:	



Общее описание инструмента. Самостоятельное неоценочное описание (200-600 слов)

Этот раздел должен содержать краткое неоценочное описание инструмента. Описание должно представить читателю ясную идею того, на что тест претендует, что он содержит, какие шкалы собирается измерять и т. д. Описание должно быть как можно более нейтральным по стилю. Следует описать тест таким, какой он есть, описать шкалы, которые он измеряет, основные точки интереса или необычные характеристики и дать какие-либо значимые сведения из истории теста. Этот раздел может быть совсем коротким (200-300 слов). Однако для более сложных, многошкальных инструментов необходимо более длинное описание (300-600 слов). Это должно быть отдельно стоящее (самостоятельное) описание инструмента. Как следствие, оно может воспроизводить часть более специфичной информации, представленной в разделах с 1.1 по 1.29.

На этот пункт следует отвечать исходя из информации, предоставляемой издателем и проверенной на точность.



.....

Раздел 2:

Классификация

.....

1.10.1	<p>Содержание (Области содержания) <i>(Пожалуйста, отметьте то, что относится к тесту).</i></p> <p>Вам следует идентифицировать области содержания, специфицированные издателем. Там, где области содержания не ясны, необходимо их указать. Выбирая наиболее подходящие ответы на пункт 1.10.1, следует исходить из информации, представленной в руководстве (выборки стандартизации, применение, валидизация и т.д.)</p>	<ul style="list-style-type: none"><input type="checkbox"/> Учебные достижения<input type="checkbox"/> Общие способности<input type="checkbox"/> Вербальные способности<input type="checkbox"/> Математические способности<input type="checkbox"/> Пространственное мышление (Пространственные способности)<input type="checkbox"/> Невербальные способности<input type="checkbox"/> Скорость восприятия<input type="checkbox"/> Память<input type="checkbox"/> Практические умения (способности)/ловкость<input type="checkbox"/> Свойства личности<input type="checkbox"/> Тип личности<input type="checkbox"/> Состояние личности<input type="checkbox"/> Когнитивные стили<input type="checkbox"/> Мотивация<input type="checkbox"/> Ценности<input type="checkbox"/> Интересы<input type="checkbox"/> Убеждения<input type="checkbox"/> Нарушения (расстройства) и патология<input type="checkbox"/> Функционирование группы<input type="checkbox"/> Функционирование семьи<input type="checkbox"/> Функционирование организации, совокупные измерения, климат в организации и т.д.<input type="checkbox"/> Школа и образование<input type="checkbox"/> Другое: (опишите ниже)
1.10.2	<p>Предполагаемая или основная область (области) применения. <i>(пожалуйста, отметьте то, что подходит)</i></p>	<ul style="list-style-type: none"><input type="checkbox"/> Клиническая психология<input type="checkbox"/> Психоневрология

	<p>Вам следует идентифицировать предполагаемые области применения, специфицированные издателем. Там, где они не ясны, их необходимо указать. Выбирая наиболее подходящие ответы на пункт 1.10.2, следует исходить из информации, представленной в руководстве (выборки стандартизации, применение, валидизация и т.д.).</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Правосудие <input type="checkbox"/> Образование <input type="checkbox"/> Труд и профессия <input type="checkbox"/> Консультирование, руководство и выбор карьеры <input type="checkbox"/> Общее здоровье, жизнь и благополучие <input type="checkbox"/> Спорт и досуг <input type="checkbox"/> Другое (опишите ниже)
1.10.3	<p>Предполагаемый способ применения (условия, при которых инструмент был стандартизирован и валидизирован) (выберите одно из предложенных)</p> <p>Этот раздел важен в силу того, что он призван установить, разработан ли инструмент с намерением его применения в неконтролируемых условиях (т.е. в условиях, когда нет надзора со стороны специалиста).</p> <p>На этот пункт следует отвечать исходя из информации, предоставленной издателем и проверенной на точность.</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Безнадзорное применение без контроля над личностью пользователя теста и без полного контроля над условиями применения (например, тесты с открытым доступом в интернете и тесты, продаваемые в книжных магазинах) <input type="checkbox"/> Контролируемый несупервизируемой администрацией. Контроль над условиями (время выполнения и т.д.) и некоторый контроль над личностью пользователя (например, тесты, внедряемые (распространяемые) посредством интернета, но имеющие ограниченный доступ. Тест доступен известным людям. Доступ обеспечивается знанием пароля). <input type="checkbox"/> Контролируемое применение. Тест применяется под контролем квалифицированного администратора и наблюдателя. <input type="checkbox"/> Управляемое применение. Применение теста только в специализированных тестовых центрах (например, оценочные программы на получение лицензии или сертификата)
1.10.4	<p>Описание групп людей, для которых тест предназначен:</p> <p>На этот пункт следует отвечать исходя из информации, предоставленной издателем.</p> <p>Для некоторых тестов эти группы могут быть очень обширными (например, взрослые), для</p>	

	<p>других тестов они могут быть более узкими (например, работники ручного труда, или мальчики от 10 до 14 лет). Этот раздел должен включать известные (установленные, устоявшиеся) группы. Ситуацию, когда они могут показаться несоответствующими, следует прокомментировать в Оценочном разделе рецензии.</p>	
1.10.5	<p>Число шкал и краткое описание переменных, измеряемых инструментом</p> <p>На этот пункт следует отвечать исходя из информации, предоставленной издателем</p> <p>Пожалуйста, укажите число шкал (если их больше, чем одна), и представьте краткое описание каждой шкалы, если ее значение неясно из ее названия. Описание инструмента должно включать обсуждение других производных оценок там, где они обычно используются в тесте, и описываются в стандартной документации – например, оценки 16PF теста и оценки баллов Adjustment Specification Equation, но не шкал по типу "дополнительных пунктов", например, использование оценок 16 PF теста для образования измерения типа команды Белбина.</p>	
1.11	<p>Формат пунктов (выберите одно)</p> <p>На этот пункт следует отвечать исходя из информации, предоставленной издателем</p> <p>Обратите внимание на то, что важно не спутать одну и ту же шкалу с множественным выбором с различными форматами пунктов шкалы с множественным выбором. Ипсативные форматы требуют, чтобы выполняющие тест сделали выбор между наборами из двух и более пунктов, полученными из разных шкал. Для «Разнородных шкальных альтернатив с вынужденным выбором» выполняющий тест должен выбрать, какая из двух шкал больше похожа на них, или с каким из двух утверждений они больше всего согласны. Для версии с множественным выбором могут существовать три или более утверждений, полученных из эквивалентного числа различных шкал. Как правило эти утверждения должно построить в ряд (классифицировать) или сделать из них выбор по принципу «что больше всего – меньше всего мне подходит».</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Открытый <input type="checkbox"/> Множественный выбор, альтернативы одной шкалы <input type="checkbox"/> Биполярные прилагательные <input type="checkbox"/> Ранжирование по Лайкерту <input type="checkbox"/> Вынужденный выбор, альтернативы смешанных шкал (ипсативных) – смотрите замечания для объяснения. <input type="checkbox"/> Множественный выбор, альтернативы смешанных шкал (ипсативных) – смотрите замечания для объяснения. <input type="checkbox"/> Дополнительная пара комплектов (наборов), смешанные шкалы (ипсативных) <input type="checkbox"/> Другое (опишите ниже)

1.12	<p>Отсутствие тестовых пунктов (вопросов):</p> <p>На этот пункт следует отвечать исходя из информации, предоставленной издателем</p> <p>Если инструмент имеет несколько шкал, проясните, указываете вы общее число пунктов или число пунктов для каждой шкалы. Там, где пункты нагружены на одну шкалу (одного масштаба), следует это подтвердить.</p>	
1.13	<p>Способ(ы) проведения теста:</p> <p>На этот пункт следует отвечать исходя из информации, предоставленной издателем</p> <p>Если для проведения теста необходимо какое-либо специальное оборудование (помимо того, что указано в списке, например, кассетный видеомэгаффон), оно должно быть описано здесь. Дополнительно должны быть описаны какие-либо специальные условия тестирования. Стандартные условия для проведения теста принимаются как приемлемые (доступные) для оценивания под наблюдением. Они включают тихую, хорошо-освещаемую и хорошо вентилируемую комнату с достаточным количеством парт и столов для ведущих и кандидатов (испытуемых).</p>	<ul style="list-style-type: none"> <input type="checkbox"/> интерактивное индивидуальное проведение теста <input type="checkbox"/> Групповое проведение под наблюдением <input type="checkbox"/> Компьютерное локальное применение - под наблюдением <input type="checkbox"/> Компьютерное веб-приложение - под наблюдением <input type="checkbox"/> Компьютерное локальное применение - без наблюдения/самооценивание <input type="checkbox"/> Компьютерное веб-приложение - без наблюдения/самооценивание <input type="checkbox"/> Другое (укажите)
1.14	<p>Способ отвечать:</p> <p>На этот пункт следует отвечать исходя из информации, предоставленной издателем.</p> <p>Если для проведения теста необходимо какое-либо специальное оборудование (помимо того, что указано в списке, например, кассетный видеомэгаффон, мэгаффон), оно должно быть описано здесь. Дополнительно должны быть описаны какие-либо специальные условия тестирования. Стандартные условия для проведения теста принимаются как приемлемые (доступные) для оценивания под наблюдением. Они включают тихую, хорошо-освещаемую и хорошо вентилируемую комнату с достаточным количеством парт и столов для ведущих и кандидатов (испытуемых).</p>	<p>Устное интервью</p> <ul style="list-style-type: none"> <input type="checkbox"/> Карандаш и бумага <input type="checkbox"/> Ручные операции <input type="checkbox"/> Компьютерное <input type="checkbox"/> Другое (укажите)
1.15	<p>Время, требуемое для проведения теста:</p> <p>На этот пункт следует отвечать исходя из информации, предоставленной издателем.</p>	<p>Подготовка: <input style="width: 100px; height: 20px;" type="text"/></p>

	<p>Ответ на этот вопрос может быть разбит на несколько составляющих его компонентов. В большинстве случаев скорее возможно произвести только некоторые общие подсчеты, чем представить точные цифры. Задача заключается в том, чтобы дать потенциальному пользователю хорошее представление о временных затратах, связанных с использованием этого инструмента. Не включайте сюда время, необходимое на знакомство с самим инструментом. Предположите, что пользователь опытен и квалифицирован. Время для подготовки (время, необходимое для ведущего подготовить и выложить (разложить) материалы для оценочной сессии).</p> <ol style="list-style-type: none"> 1. <i>Время, необходимое для одной сессии:</i> оно включает время, необходимое для завершения всех пунктов и время, необходимое для дачи инструкций, работы с примерами и кратких комментариев в конце сессии. 2. <i>Подсчеты:</i> время, потраченное на подсчет сырых данных. 3. <i>Анализ:</i> время, потраченное на то, чтобы произвести дальнейшую работу с сырыми данными, вывести другие оценки и произвести разумно исчерпывающую интерпретацию (предполагая, что вы знакомы с инструментом). 4. <i>Обратная связь:</i> время, требуемое для того, чтобы подготовить и предоставить обратную связь испытуемому. <p>Признается, что время для двух последних компонентов очень варьирует в зависимости от контекста, в котором используется инструмент. Тем не менее, какое-либо указание на время или комментарий по этому поводу будут полезны.</p>	<p>Проведение: <input type="text"/></p> <p>Подсчеты: <input type="text"/></p> <p>Анализ: <input type="text"/></p> <p>Обратная связь: <input type="text"/></p>
1.16	<p>Укажите, действительно ли различные формы инструмента доступны (истинные и псевдо-параллельные формы, короткие версии, компьютерные версии, и т.д.). Если компьютерные версии действительно существуют, опишите кратко требования к программному обеспечению и аппаратным средствам:</p> <p>Для раздела 1.16 сообщите, существуют или нет доступные альтернативные версии (формы) инструмента и опишите применимость (пригодность) каждой формы для различных групп людей. В некоторых случаях подразумевается, что различные формы инструмента являются эквивалентными друг другу – т.е. альтернативными формами. В других случаях различные формы могут существовать для совершенно различных групп (например, детская форма и взрослая форма). Там, где существует больше одной формы, укажите,</p>	

	<p>являются ли данные формы эквивалентными/альтернативными или они разработаны для разных целей – например, короткая и длинная версии; ипсативные и нормативная версии.</p> <p>Некоторые инструменты могут быть частично или полностью компьютеризированы или доступны в компьютерных версиях. Для каждой из четырех ступеней оценочного процесса укажите опции, которые можно получить от поставщика. Обратите внимание на то, что пакеты Компьютеризированной обработки данных, если они доступны, должны быть указаны.</p>	
--	---	--

.....

Раздел 3:

Измерения и подсчеты

.....

<p>1.17</p>	<p>Процедура подсчета для теста:</p> <p>Этот пункт следует выполнить ссылаясь на информацию от издателя, на руководства и документацию.</p> <p>Бюро услуг – услуги по подсчету и интерпретации, предоставляемые поставщиком или агентом поставщика. В основном это дополнительные услуги. В случае, если подсчет или/и интерпретация могут быть выполнены только через бюро услуг, сообщите об этом в описании теста, вместе с этим сообщите и о ценах на эти услуги в разделе цен.</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Компьютерный подсчет с непосредственным вводом испытуемым ответов в компьютер <input type="checkbox"/> Компьютерный подсчет при помощи ручного ввода в компьютер ответов, проставленных на бумажных формах (бланках) <input type="checkbox"/> Компьютерный подсчет ответов, полученных на бумажных формах, вводимых в компьютер оптическим считывающим устройством <input type="checkbox"/> Простой подсчет вручную – требуются только канцелярские навыки <input type="checkbox"/> Сложный подсчет вручную – требуется опыт в подсчете результатов теста <input type="checkbox"/> Бюро услуг - например, подсчет компанией, продающей инструмент <input type="checkbox"/> Другое (опишите ниже)
<p>1.18</p>	<p>Обработка:</p> <p>Этот пункт следует выполнить ссылаясь на информацию от издателя, на руководства и документацию.</p>	

	Краткое описание системы подсчета для получения итоговых и промежуточных оценок, коррекция выводов (догадок), качественные средства интерпретации и т.д.)	
1.19	Перевод сырых оценок в стандартные:	<input type="checkbox"/> Нормализованный – оценки получают с помощью справочной таблицы <input type="checkbox"/> Не нормализованный – оценки получают путем линейной трансформации (перевода)
1.20	Применяемые шкалы (отметьте то, что подходит) Этот пункт следует выполнить ссылаясь на информацию от издателя, на руководства и документацию.	Оценки, основанные на процентилях <input type="checkbox"/> Центили <input type="checkbox"/> 5 - ступенчатая классификация: 10:20:40:20:10 центильные деления <input type="checkbox"/> Децили и другие эквивалентные процентильным классификации Стандартные оценки <input type="checkbox"/> Z-баллы <input type="checkbox"/> Коэффициенты, отклоняющиеся от IQ (например, среднего 100, SD=15 для Векслера или 16 для Станфорд-Бине) <input type="checkbox"/> Комиссия по вступительным экзаменам в колледж (например, среднее SAT (Экзамен на определение академических способностей) = 500, отклонение =100) <input type="checkbox"/> Стены, Стенайны, С баллы <input type="checkbox"/> Т- баллы <input type="checkbox"/> Другое (пожалуйста, опишите)

.....

Раздел 4:

Компьютерные отчеты

.....

Обратите внимание на то, что это является чисто описательным. Оценка отчетов будет частью Оценочного раздела рецензии (описания)

К примеру, там, где возможно сделать несколько отчетов, пожалуйста, заполните пункты 1.21 – 1.22 для каждого отчета (пожалуйста, скопируйте страницы в случае необходимости). Эта система классификации могла бы быть использована для описания двух видов отчета. Например, Отчет 1 мог бы предназначаться для испытуемого или другого неопытного пользователя, а Отчет 2 для опытного пользователя, который компетентен в использовании инструмента и понимает, как его интерпретировать.

1.21	<p>Доступны ли компьютерные отчеты для инструмента?</p> <p>Если ответ на пункт 1.21 – «да», то следует применить следующую классификацию, чтобы классифицировать типы доступных (приемлемых) отчетов. Для многих инструментов существует диапазон доступных (возможных) отчетов. В разделе 1.21 следует сделать попытку охарактеризовать этот диапазон и разнообразие отчетов. На практике многие системы позволяют получить расширенные результаты, которые представляют собой комбинации типов, описанных ниже.</p>	<p><input type="checkbox"/> Да (выполните разделы ниже)</p> <p><input type="checkbox"/> Нет (перейдите к пункту 1.23)</p>
1.21.0	<p>Название или описание отчета: (Смотрите введение к этому разделу)</p>	
1.21.1	<p>Средства передачи информации (выберите одно)</p> <p>Отчеты могут состоять только из текста или содержать текст с графическим или табличным представлением результатов (например, стеновыми профилями). Там, где представлены и текст, и цифровые данные, они могут быть представлены последовательно или могут быть увязаны друг с другом таким образом, чтобы показать связь между положениями текста и оценками.</p>	<p><input type="checkbox"/> Только текст</p> <p><input type="checkbox"/> Независимый текст и графические данные</p> <p><input type="checkbox"/> Взаимосвязанные текст и графические данные</p>
1.21.2	<p>Сложность (выберите одно):</p> <p>Некоторые отчеты очень просты, например такие, которые заменяют единицу (часть, кусок) текста на стен (стенную оценку) в последовательном описании шкал. Другие являются более сложными, включающими единицы текста, которые соотносятся с образцами и конфигурациями шкальных оценок, и которые учитывают эффект взаимодействия шкал.</p>	<p><input type="checkbox"/> Простой (например, список параграфов, дающих описание шкал)</p> <p><input type="checkbox"/> Средний (смесь простых описаний и описаний некоторых конфигураций))</p> <p><input type="checkbox"/> Сложный (содержит описания образцов и конфигураций шкальных оценок и взаимодействий шкал)</p>

<p>1.21.3</p>	<p>Структура отчета: <i>(выберите одно)</i></p> <p>Структура связана со сложностью</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Основанный на шкале (когда отчет строится вокруг индивидуальных шкал) <input type="checkbox"/> Основанный на факторе (когда отчет конструируется вокруг факторов более высокого порядка, таких как «Большая пятерка» для личностных измерений) <input type="checkbox"/> Основанный на конструктах – когда отчет строится вокруг одного или более наборов конструктов (например, в наборах связанных с работой, могут быть такие конструкты как типы команд, стили лидерства, стрессоустойчивость и др.), которые связаны с оценками оригинальной шкалы. <input type="checkbox"/> Основанный на критерии, когда отчет фокусируется на связях с эмпирическими результатами (например, наращиванием потенциала, выполнением работы, прогулами и др.) <input type="checkbox"/> Другой
<p>1.21.4</p>	<p>Чувствительность к контексту <i>(выберите одно)</i></p> <p>Когда люди пишут отчеты, они приспосабливают язык, форму и содержание отчета к человеку, который будет читать его, и берут в расчет цели оценивания и контекст, в котором оно происходит. Отчет, созданный для целей отбора, будет отличаться от того, что написан для молодого человека, начинающего карьеру.</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Одна версия для всех контекстов <input type="checkbox"/> Версии, связанные с predetermined контекстом <input type="checkbox"/> Отчет, контекст для которого определяет пользователь, он же и редактирует отчет.
<p>1.21.5</p>	<p>Клинико-страховой <i>(выберите одно)</i></p> <p>Большинство систем отчетов основываются на клинических суждениях (выводах). Т.е. один или более человек, которые являются экспертами в использовании инструмента, о котором идет речь, напишут единицы текста. Отчеты поэтому будут олицетворять их конкретные интерпретации шкал. Некоторые системы включают страховые отчеты, когда утверждения основываются на эмпирических исследованиях валидности, связывающих шкальные оценки, например, с измерениями выполнения работы.</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Основанный на клинических суждениях одного эксперта <input type="checkbox"/> Основанный на эмпирических /страховых отношениях <input type="checkbox"/> Основанный на клинических суждениях группы экспертов

1.21.6	<p>Модифицируемость <i>(выберите одно)</i></p> <p>Отчетный результат является часто фиксированным. Тем не менее, некоторые системы имеют результат в форме файла, который может быть обработан пользователем.</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Не модифицируемый (фиксированный только напечатанный результат) <input type="checkbox"/> Ограниченно модифицируемый (ограниченный до определенных областей, например, до области биографии) <input type="checkbox"/> Неограниченно модифицируемый (например, посредством доступа к информационным массивам Word)
1.21.7	<p>Степень завершенности <i>(выберите одно)</i></p> <p>Родственным результатом является та степень, до которой система разработана, чтобы создавать интегрированный текст – в форме готового для чтения отчета или в форме набора заметок, комментариев, гипотез и т.п. Последнее используется гораздо чаще, когда текст доступен пользователю в модифицируемой форме и может образовывать основу для собственного отчета пользователя. Во многих случаях отчеты разрабатываются как отчеты с очень высокими стандартами презентации и имеют внешний вид и качество публикации.</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Качество публикации <input type="checkbox"/> Качество чернового документа
1.21.8	<p>Прозрачность <i>(выберите одно)</i></p> <p>Системы отличаются по своей открытости и прозрачности для пользователя. Открытая система – система, в которой связь между оценкой по шкале и текстом ясна и недвусмысленна. Такая открытость возможна только если представлены и текст и оценки и связь между ними определена (ясна). Другие системы действуют как «черные ящики», создавая трудности для пользователя в соотнесении шкальных оценок с текстом.</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Ясная связь между конструктами, оценками и текстом <input type="checkbox"/> Скрытая связь между конструктами, оценками и текстом <input type="checkbox"/> Смесь ясных/скрытых связей между конструктами, оценками и текстом
1.21.9	<p>Стиль и тон <i>(выберите одно)</i></p> <p>Системы также отличаются тем, в какой мере они могут предложить читателю отчета руководство или управление. Некоторые являются обусловленными «Мистер X очень застенчив и из него не получится хорошего коммерсанта». Другие разработаны так, что предлагают гипотезы или поднимают вопросы: «Судя по его оценкам по шкале У, мистер X очень застенчив. Если это именно так, он будет испытывать трудности,</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Директивная <input type="checkbox"/> Руководящая <input type="checkbox"/> Другое

	<p>работая в коммерческом окружении. Это требует дальнейшего исследования».</p>	
1.21.10	<p>Предполагаемые реципиенты (выберите все, что подходит)</p> <p>Отчеты обычно разрабатываются, чтобы адресоваться нуждам одной или более категории пользователей. Пользователи могут быть поделены на четыре основных группы:</p> <p>a) <i>Квалифицированные пользователи.</i> Это люди, которые обладают достаточными знаниями и опытом, чтобы самостоятельно составлять отчеты, основанные на шкальных оценках. Они должны быть способны использовать отчеты, которые содержат техническую психометрическую терминологию, и выявлять связь между шкалами и описаниями. Они должны также уметь модифицировать отчет в соответствии с требованием клиента (по заказу клиента).</p> <p>b) <i>Квалифицированные пользователи системы.</i> Пока некомпетентны, чтобы составлять самостоятельно отчеты, основываясь на шкальных оценках. Люди этой группы способны использовать результаты, генерируемые системой. Уровень подготовки, необходимый для достижения такой компетентности, будет значительно варьировать в зависимости от характера компьютерных отчетов (например, основанных на черте (признаке характера) в сравнении с теми, что основаны на компетентности, простыми и сложными) и от того, где будут применяться эти отчеты (как много будет поставлено на эти отчеты).</p> <p>c) <i>Те, кто выполняет тест.</i> Человек, который выполняет тест, обычно не имеет предварительных знаний ни о тесте, ни о типе отчета, производимого системой. Отчеты для них должны быть написаны на языке, понимание которого не требует знания психометрии и тестов.</p> <p>d) <i>Третьи роли.</i> Они включают людей, отличных от кандидатов, тех, кто будут посвящены в информацию, предоставляемую отчетом, и кто может получить копию отчета. Сюда могут входить потенциальные работодатели, менеджер по персоналу, супервизор или родитель молодого человека, получающего совет в отношении будущей карьеры. Уровень языка, требуемый для людей этой категории, будет таким же, какой требуется для отчетов, предназначенных для тех, кто выполняет тест.</p>	<p><input type="checkbox"/> Квалифицированные пользователи тестов</p> <p><input type="checkbox"/> Квалифицированные пользователи системы</p> <p><input type="checkbox"/> Те, кто выполняет тест</p> <p><input type="checkbox"/> Третьи роли</p>
1.22	<p>Предлагают ли дистрибьюторы услугу коррекции и/или разработки компьютерных отчетов? (выберите одно)</p>	<p><input type="checkbox"/> Да</p> <p><input type="checkbox"/> Нет</p>

Раздел 5:

Поставка, условия и цены

.....

Это означает, что издатель должен предоставить, кому, на каких условиях и по каким ценам. Это означает условия, предлагаемые поставщиком, у которого можно или нельзя получить тестовые материалы. Если одна из опций не соответствует (не подходит) условиям поставки, дайте (заранее заготовьте) описание подходящих условий.

1.23	Документация, предоставляемая дистрибьютором, как часть тестового пакета (выберите все то, что подходит)	<ul style="list-style-type: none"><input type="checkbox"/> Руководство для пользователя<input type="checkbox"/> Техническое (психометрическое) руководство<input type="checkbox"/> Дополнительная техническая информация и последние новости (информация о последней модернизации или т.п., например, локальные нормы, локальные исследования валидности и т.д.)<input type="checkbox"/> Книги и статьи на смежные темы<input type="checkbox"/> Комбинации из того, что перечислено выше (укажите точно, специфицируйте)
1.24	Способы публикации (выберите все, что подходит) К примеру, технические руководства могут сохраняться на высоте современных требований и быть доступными для скачивания из Интернета, в то время как руководства для пользователей предоставляются на бумажных носителях или на CD	<ul style="list-style-type: none"><input type="checkbox"/> Бумага<input type="checkbox"/> PC – дискеты<input type="checkbox"/> PC – CD-ROM<input type="checkbox"/> Загрузка из Интернета<input type="checkbox"/> Живой Интернет (тест есть в веб-браузере)<input type="checkbox"/> Другое (укажите точно)
<p>В разделах 1.25-1.27 говорится о ценах. Вероятно, это тот раздел, который быстрее всех устаревает. Рекомендуется входить в контакт с поставщиком или издателем как можно ближе ко времени публикации рецензии, чтобы получить текущую информацию для этого раздела.</p>		
1.25.1	Стартовые цены. Цена полного комплекта материалов (все руководства и другие материалы, достаточные, по крайней мере, для однократного проведения теста). Укажите точно, сколько кандидатов можно оценить с помощью материалов, приобретенных по стартовым ценам, в каком случае эти цены включают периодическое оценивание. В этом пункте должна быть сделана попытка идентифицировать «стартовую» цену. Это цена,	

	<p>включенная в цену на полный набор справочных материалов, ключи и т.д. (Это часть цены на полный пакет справочных материалов) Это стоимость подготовки к проведению теста, если тест носит «закрытый» характер, когда будет неизбежной оплата специальной подготовки к нему, независимой от уровня предшествующей квалификации пользователя. В таких случаях стоимость подготовки должна быть заявлена в общей цене. Начальные цены не включают цен на оборудование общего назначения (такое как компьютеры, кассетные магнитофоны и др.) Тем не менее, потребность в них должна быть упомянута. В целом, установите: цены на любую специальную подготовку; цены на руководство для ведущего, техническое руководство, комплект образцов или комплект справочных материалов; начальные цены на программное обеспечение и т.д.</p>	
1.25.2	<p>Периодические цены: Уточните, где необходимо, периодические цены на проведение и подсчет данных отдельно от цен на интерпретацию (см. пп 1.26 – 1.27).</p> <p>Этот пункт имеет дело с ценами на постоянное использование инструмента. Следует сообщить цены на материалы теста (бланки ответов, буклеты с вопросами однократного и многократного использования, бланки с профилями, компьютерные ключи и т.д. на одного человека для одного проведения теста. Обратите внимание на то, что в большинстве случаев проведения теста на бумажной основе такие материалы недоступны поштучно, а имеют тенденцию быть поставляемыми пакетами по 10, 25 или 50 штук.</p> <p>Перечислите по пунктам любые ежегодные или подушевые (на одного человека) лицензионные платежи (включая программное обеспечение ключей, где это значимо), цены на покупку или аренду материалов многократного использования и цены на материалы однократного использования на одного кандидата.</p>	
1.26.1	<p>Цены на отчеты, генерируемые программным обеспечением, установленным пользователем:</p>	
1.26.2	<p>Цены на отчеты, выполненные бюро услуг и пересылаемые по почте или факсу:</p>	

1.26.3	1.26.3. Цены на отчеты, получаемые от бюро услуг через Интернет:	
1.27	1.27. Цены на другие услуги бюро: коррекция или разработка автоматических отчетов:	
1.28	<p>Имеющая отношение к тесту квалификация, требуемая поставщиком теста <i>(выберите все, что подходит)</i></p> <p>Пункт 1.28 касается квалификации пользователя, требуемой поставщиком теста. Для этого раздела там, где издатель предоставил информацию и квалификации пользователя, следует сделать отметку против данных категорий. Там, где требования к квалификации не ясны, следует сделать отметку под «Другое», не под «Никакая». «Никакая» означает, что существует явное утверждение относительно отсутствия необходимости какой-либо квалификации.</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Никакая <input type="checkbox"/> Специальная тестовая аккредитация <input type="checkbox"/> Аккредитация в области тестирования общих достижений: измерения максимальных достижений <input type="checkbox"/> Аккредитация в области тестирования общих способностей: измерения максимальных потенциальных достижений <input type="checkbox"/> Аккредитация в области общей оценки личности: измерение типичного поведения, отношений и предпочтений. <input type="checkbox"/> Другое (уточните)
1.29	<p>Профессиональная квалификация, необходимая для использования инструмента <i>(выберите все, что подходит)</i></p> <p>Пункт 1.28 касается квалификации пользователя, требуемой поставщиком теста. Для этого раздела там, где издатель предоставил информацию и квалификации пользователя, следует сделать отметку против данных категорий. Там, где требования к квалификации не ясны, следует сделать отметку под «Другое», не под «Никакая». «Никакая» означает, что существует явное утверждение относительно отсутствия необходимости какой-либо квалификации.</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Никакая <input type="checkbox"/> Практический психолог с квалификацией релевантной области применения <input type="checkbox"/> Практический психолог <input type="checkbox"/> Психолог – исследователь <input type="checkbox"/> Ученый – исследователь, не психолог <input type="checkbox"/> Практик в смежных профессиях (терапия, медицина, консультирование, образование, человеческие ресурсы и др.) <input type="checkbox"/> Владелец Сертификата BPS, подтверждающего Компетентность в Профессиональном Тестировании Уровня А. <input type="checkbox"/> Владелец сертификата BPS, подтверждающего Компетентность в Тестировании в области Образования Уровня А. <input type="checkbox"/> Владелец Сертификата BPS, подтверждающего Компетентность в Профессиональном Тестировании

		Уровня В. <input type="checkbox"/> Другое (укажите)

Раздел 6:

Оценивание тестовых материалов

В потенциале существует четыре источника информации, с которой можно соотноситься при проведении этого оценивания.

- Руководство и/или отчеты для пользователя, поставляемые издателем/дистрибьютором:

а) Они всегда поставляются издателем/дистрибьютором прежде чем инструмент будет принят (утвержден) офисом и образуют сердцевину рецензии (описания).

2. Открытая информация, которая доступна в академической и другой литературе:

а) Она обычно берется за исходный источник рецензентом и рецензент может использовать эту информацию в рецензии (описании) и инструмент может быть оценен как соответствующий или не соответствующий информации в его руководстве.

3. Отчеты, которыми владеет издатель, которые формально не публикуются и не распространяются:

а) Дистрибьютор/издатель может сделать их доступными в начале или может отослать их, когда рецензия присылается назад издателю для проверки на фактическую точность. Рецензенту следует использовать эту информацию, но указать очень ясно в начале комментариев по технической информации, что «отмеченная звездочкой оценка в этой рецензии относится (отсылает) к материалам, которыми владеет издатель/дистрибьютор, которые обычно не поставляются пользователям». Если они содержат ценную информацию, в общей оценке (то есть в конце) следует порекомендовать, чтобы издатель публиковал эти отчеты и/или сделал их доступными для покупателей тестов.

4. Отчеты, которые являются коммерческой тайной:

а) В некоторых случаях издатели могут иметь технически важный материал, который они не хотят публиковать по коммерческим соображениям. На практике существует очень мало способов защиты интеллектуальной собственности (для) разработчиков тестов (закон о праве собственности – единственная надежда). Такие отчеты могут содержать способы разработки конкретных счетных алгоритмов, процедуры создания тестов или отдельных пунктов и технологию создания отчета. Там, где содержание подобных отчетов могло бы быть важным для формирования суждений о тесте, BPS могло бы предложить заключить тайное соглашение с издателем. Это соглашение связывало бы рецензентов и редактора (определенными обязательствами). Рецензент мог бы тогда оценить информацию и прокомментировать технические аспекты и все оценивание (всю процедуру оценивания, весь тест) до такого следствия, что «отмеченное звездочкой положение в этом описании относится к материалам, которыми владеет издатель/дистрибьютор, и которые были изучены рецензентом на условиях сохранения коммерческой тайны. Они не поставляются вместе с руководством».

Объяснение оценок со звездочкой

Все разделы обсчитываются по оценочной системе, где оценка обозначается знаком, состоящим из квадратных скобок и звездочек. Предоставляются подробные описания, дающие опорные точки для каждой оценки.

Там, где оценка типа [] дается на качество, которое рассматривается как критическое для безопасного использования инструмента, рецензент порекомендует не использовать инструмент за исключением случаев его использования высококвалифицированными экспертами или в исследовании.

В рецензии (описании) инструмента необходимо указать, что, обусловленное природой инструмента и его предполагаемым использованием, является критическими качествами. Предлагается согласиться с тем, что оценки этих критических качеств тогда должны показываться жирным шрифтом.

В последующих разделах общие оценки соответствия требованиям информации, имеющей отношение к валидности, надежности и нормам, показываются по умолчанию жирным шрифтом.

Любой инструмент с одной или более оценкой типа [] or [*] относительно качества, которые рассматриваются как критические в отношении безопасного использования инструмента, не будет считаться отвечающими минимальным стандартам.

Начало по оценочной форме	EFPA стандартная оценка	UK: изображение в рецензии	Объяснение
[n/a]	[n/a]	[n/a]	Это свойство не применимо (не подходит, не пригодно) к этому инструменту
0	[-]	[None]	Невозможно оценить как нет, или предоставлено недостаточно информации
1	[-1]	[*]	Не соответствующий требованиям
		[**]	СЕЙЧАС НЕ ИСПОЛЬЗУЕТСЯ
3	[0]	[***]	Соответствует требованиям
4	[1]	[****]	Хороший
5	[2]	[*****]	Отличный
		[N.r.i.o.r]* (Только для последних модификаций)	Пункт не был оценен в оригинальном описании (рецензии)

В этом разделе необходимо дать ряд оценок различным аспектам или свойствам документации, поставляемой вместе с инструментом (или пакетом). Термин «документация» взят для обозначения всех тех материалов, которые с готовностью предоставляются квалифицированному пользователю: например, руководство для ведущего, технические руководства, буклеты с нормами, справочные приложения, последние новости от издателей/поставщиков и т. д.).

Поставщиков просят предоставить полный набор таких материалов каждому рецензенту. Если вы думаете, что существует что-то, чем снабжаются пользователи и чего не содержится в информации, присылаемой вам для рецензии, пожалуйста, свяжитесь с редактором-консультантом.

Пункты, которые должны быть оценены как n/a или 0 до 5 (оценки с половинками допускаются)

Оценка

Качество объяснения теоретических основ, презентация и качество представленной информации: ((Эта общая оценка достигается путем использования суждений, на которых основаны оценки в пунктах 2.1 – 2.8)	
--	--

2.1	Общая оценка качества объяснения теоретических основ (Эта общая оценка достигается с опорой на оценки, данные в пунктах 2.1.1 – 2.1.5)	
2.1.1	i) Теоретические основы конструкторов:	
2.1.2	II) Процедура разработки теста:	
2.1.3	III) Тщательность анализа пунктов и модель анализа пунктов:	
2.1.4	IV) Объяснение содержательной валидности:	
2.1.5	V) Итоговые результаты релевантных исследований:	
2.2	Адекватность документации, доступной для пользователя (руководства для пользователя и технические руководства, приложения с нормами и др.): (Эта общая оценка достигается путем использований суждений, на которых основаны оценки в пунктах 2.2.1 – 2.2.6) <i>Для раздела 2.2 следующие отметки уровня предоставляются для отличной (****) оценки.</i> Внимание здесь концентрируется на качестве охвата информации, представленной в документации, доступной для квалифицированных пользователей. Обратите внимание на то, что раздел 2.2 касается полноты и ясности документации, доступной для пользователя (руководство для пользователя, техническое руководство, приложения с нормами и т.д.) в терминах ее охвата и объяснения. О качестве инструмента, свидетельствуемом в документации, речь идет в пунктах 2.1, 2.3, 2.9, 2.10 и 2.11.	
2.2.1	Теоретические основы: [см. оценку 2.1] Хорошо аргументированное и четко представленное описание того, что разработано для измерения, и почему оно было сконструировано так, как оно сконструировано.	
2.2.2	Процесс разработки инструмента: Все подробности об источниках пунктов, пилотирование (?), анализ пунктов, сравнительные исследования и изменения, сделанные во время испытаний.	
2.2.3	Стандартизация: Ясная и подробная информация о размерах и источниках выборки и процедуры стандартизации	
2.2.4	Нормы: Ясная и подробная информация о размерах и источниках групповых норм, условиях оценивания и др.	
2.2.5	Надежность: Хорошее объяснение надежности и полный диапазон измерений внутренней согласованности и ретестовой надежности, предоставленных с объяснениями их релевантности и обобщающей способности оценивающего инструмента.	
2.2.6	Валидность: Хорошее объяснение валидности с широким рядом исследований, ясно и честно описанных.	

2.3	Качество процедурных инструкций, предоставляемых пользователю: (Эта общая оценка достигается путем использования суждений, на которых основаны оценки в пунктах 2.3.1. – 2.3.7)	
2.3.1	О проведении теста: Предоставляются ясные и подробные объяснения и пошаговые процедурные инструкции с хорошими подробными советами относительно того, как поступать с вопросами кандидатов и проблемными ситуациями.	
2.3.2	О подсчетах, нормировании и др.: Предоставляется ясная и подробная информация со способами проверки, описывающими как поступать с возможными ошибками в подсчетах.	
2.3.3	Об интерпретации и составлении отчетов: Подробные советы по интерпретации различных результатов, пониманию нормативных измерений и тому, как учитывать взаимодействие различных шкал, с большим количеством иллюстративного материала и исследований конкретных случаев.	
2.3.4	О предоставлении обратной связи и опроса тестируемых и других: Подробные советы о том, как предоставить обратную связь кандидатам	
2.3.5	Об обеспечении хороших практических результатов по поводу честности и предубеждений: <i>Сообщается подробная информация об исследованиях в области пола и этнических предубеждений с релевантными предупреждениями относительно использования и обобщения валидностей.</i>	
2.3.6	Ограничения применения: Ясные описания тех, кого следует и тех, кого не следует оценивать, с хорошими аргументами для ограничений (например, требуется указать типы неспособности, уровни грамотности и др.)	
2.3.7	Ссылки и вспомогательные материалы: Подробные ссылки на релевантную вспомогательную академическую литературу и ссылки на другие материалы, имеющие отношение к инструменту оценивания	
Качество материалов: (Эта общая оценка достигается путем использования суждений, на которых основаны оценки в пунктах 2.4 – 2.8)		
2.4	Общее качество тестовых материалов: (тестовых буклетов, бланков ответов, тестовых объектов, программного обеспечения и др.)	
2.5	Качество локальной адаптации теста (если тест был переведен на другой язык и адаптирован):	
2.6	Легкость, с какой выполняющий тест может понять свою задачу:	
2.7	Легкость, с какой выполняющий тест может давать ответы или реагировать:	
2.8	Качество пунктов:	

Комментарии рецензентов по поводу документации: (комментарии к теоретическим основам, дизайну, разработке теста и приемлемости)



.....

Раздел 7: Оценивание норм, надежности и валидности

.....

- Общее руководство по вынесению оценок для этих разделов

Почти невозможно установить четкие критерии для оценки технических качеств инструмента. При некоторых условиях надежность в пределах 0,70 хороша, в других условиях она может быть недостаточной. В некоторых ситуациях критериальная валидность при 0,20 может иметь значительную пользу, тогда как в других ситуациях она же при 0,40 может представлять малую ценность. По этим причинам ваши итоговые оценки должны быть основаны на ваших рассуждениях и знаниях рецензента, а не просто на вычислении среднеарифметических результатов.

Эти заметки дают некоторое руководство касательно достоинств (свойств) инструмента, заслуживающих оценок «не отвечающий требованиям», «отвечающий требованиям», «хороший» или «отличный». Тем не менее предполагается, что они будут выступать только как советчики. Характер инструмента, область его применения, качество данных, на которых основаны оценки надежности и валидности, и типы решений, для вынесения которых будет использоваться инструмент, все должно повлиять на способ присуждения оценок.

Чтобы дать некоторое представление о ряде и распределении достоинств, связанных с различными шкалами, образующими инструмент, запишите (внесите) *число шкал* в каждом разделе. Например, если инструмент, используемый для определения группового уровня, имел 15 шкал, из которых пять шкал имели ретестовую надежность ниже, чем 0,6, шесть шкал между 0,6 и 0,7 и остальные четыре - в пределах от 0,7 до 0,8, то это следовало бы записать так:

Стабильность:

Средняя стабильность:

- | | |
|-----|---|
| [] | Нет информации. |
| [5] | Не отвечает требованиям(например, $r < 0.6$). |
| [6] | Отвечает требованиям (например, $0.6 < r < 0.70$). |
| [4] | Хорошо (например, $0.70 < r < 0.80$). |
| [] | Отлично (например, $r > 0.80$) |

Понятно, что во многих случаях невозможно подсчитать точные действительные средние цифры. Что требуется от вас, так это произвести как можно лучше расчеты, учитывая информацию, предоставляемую в документации. Есть место, чтобы добавить комментарии. Здесь вы можете изложить (обратить внимание на) любые ваши проблемы относительно точности ваших расчетов. Например, в некоторых случаях очень высокий уровень внутренней согласованности мог бы быть прокомментирован как означающий «раздутую особенность».

Для каждой из возможных оценок даются образцы (примеры) достоинств с целью помочь увидеть различия между оценками, особенно различия между оценками «отвечает требованиям», «хорошо» и «отлично». Признается, что свойство, «хорошее» для одного инструмента, может быть нереальным для другого. (Например, при прочих равных условиях для длинных шкал мы ожидаем более высокой внутренней согласованности, чем для коротких).

Что касается норм, то правила относительно размеров выборки должны брать в расчет тип норм, который используется. Если они претендуют быть общими для целой популяции, то тогда размер выборки

должен быть достаточно большим даже для того, чтобы быть оцененным как «отвечающий требованиям». Если это нормы - специфичные для какой-либо профессиональной группы, меньшие по размерам выборки могут быть «отвечающими требованиям».

Внимательного рассмотрения заслуживает пригодность (приемлемость) международных (тот же язык) норм. Если они тщательно выведены на выборках, скомплектованных из группы стран, они должны оцениваться на той же основе, что и национальная (один язык) нормативная группа.

Для большинства целей выборки меньше 150 человек будут слишком малы, поскольку разрешающая способность на концах распределения будет очень маленькой. Значение средних с z-баллами в $n=150$ равно 0.082 от SD – это лучше чем для T-баллов.

Для валидности правила, касающиеся размеров выборки, основываются на анализе мощности размеров выборки, необходимых для отыскания валидностей выборки средних размеров, если они существуют. Конкурентная и предсказательная валидность имеют отношение к исследованиям, в которых измерения реальных критериев (т.е. не других тестовых результатов, баллов, очков) были скоррелированы со шкалами. Предсказательные исследования обычно имеют отношение к ситуациям, когда оценивание осуществлялось на «качественно» отличной точке во времени по отношению к измерению критерия – например, для измерения, связанного с отбором для работы, имеющего целью предсказать успех на работе, тест был бы проведен скорее во время отбора, чем просто оставался вопросом того, сколь продолжителен был временной интервал между тестом и измерением критерия.

Конструктивная валидность включает корреляции шкал из подобных (похожих) инструментов. Правила выведения коэффициентов конструктивной валидности необходимо интерпретировать гибко. Там, где два похожих инструмента скоррелированы (с данными, полученными конкурентно), мы бы ожидали обнаружить корреляции, равные 0,6 или более для оценки «отвечающий требованиям». Там, где инструменты меньше похожи, или проведение тестов разделено некоторым временным интервалом, более низкие коэффициенты могут быть «отвечающими требованиям». При оценке конструктивной валидности следует обратить внимание на интерпретацию очень высоких корреляций. Там, где корреляции выше 0,90, есть вероятность, что изучаемые шкалы измеряют одно и то же. Это не является проблемой, если шкалы, о которых идет речь, представляют новую шкалу и установленный маркер. Хотя было бы проблемой, если бы подразумевалось, что шкала, о которой идет речь, привносит полезное разнообразие в то, что уже измеряет другая шкала.

Когда мы судим об общей валидности, важно держать в уме значимость конструктивной валидности как лучшего индикатора того, измеряет ли тест то, что он претендует измерять. В некоторых случаях главное свидетельство этого могло бы принять форму исследований, связанных с критериями. Такой тест мог бы иметь оценку «отвечающий требованиям» или лучше за критериальную валидность и оценку меньше чем «отвечающий требованиям» за конструктивную валидность. В целом, если доказательство критериальной валидности и доказательство конструктивной валидности являются, по крайней мере, отвечающими требованиям, подразумевается, что и общая валидность должна быть, по крайней мере, отвечающей требованиям. Ее не следует рассматривать как средний или самый низкий общий знаменатель.

В отношении надежности правила основываются на необходимости иметь маленькую стандартную ошибку для оценок надежности. Путеводные критерии для надежности даются в отношении двух разных контекстов: использование инструментов для принятия решений касательно групп людей (например, классификация людей на категории) и их использование для индивидуального оценивания. Требования к надежности выше в отношении последнего. Другие факторы также могут повлиять на требования к надежности. Например, такой фактор как интерпретируются ли шкалы сами по себе или объединяются с другими шкалами в сложную шкалу. В последнем случае оценивается надежность не компонентов, а всей составной шкалы в целом.

Пункты, которые надо оценить как n/a или 0 до 5 (половинчатые оценки допускаются)

Оценка

Оценивание технической информации – Общее соответствие

требованиям: (Эта общая оценка достигается путем использования суждений, на которых основаны оценки в пунктах 2.9 – 2.11)

Лучше всего выполнить эту оценку после того, как вы завершили разделы 2.9 – 2.11

Этот раздел касается природы и качества технической информации, которая представлена в доступной документации. Но это не том, как ясно или как хорошо представлена информация. *Пожалуйста, будьте уверены (убедитесь), что вы основываете свои оценки на информации, которая доступна обычному пользователю оценочного инструмента, а не академику или специалисту.* Если имеется другая доступная информация о качествах инструмента (например, американские руководства, неопубликованные в этой стране, статьи в психологической литературе), на нее следует сделать ссылку в оценочном разделе рецензии.



Нормы или информация о референтных группах

2.9	<p>Общее соответствие требованиям: (Эта общая оценка достигается путем использования суждений, на которых основаны оценки в пунктах 2.9.1 – 2.9.5. Не стремитесь к среднеарифметическим цифрам для получения общей оценки)</p> <p>Оценки могут быть получены путем регистрации числа шкал, которые отвечают следующим критериям, и последующего анализа характера распределения.</p> <p>Там, где инструмент разработан для применения без обращения к нормам или референтным группам, лучше использовать категорию «не применим», чем «информация не дана».</p>	[]
2.9.1	<p>Пригодность для локального использования локальных или международных норм:</p> <p>[n/a] Не применим</p> <p>[] Информация не дана</p> <p>[*] Не подходит для локальных условий (например, неподходящие иностранные выборки)</p> <p>[**] Локальная выборка населения в целом или не локальные нормы, которые могли бы быть использованы с осторожностью</p> <p>[****] Локальные выборки или релевантные (уместные) международные выборки с хорошей релевантностью (уместностью) для предполагаемого применения</p> <p>[*****] Локальные выборки или релевантные (уместные) международные выборки, составленные из хорошо характеризующихся выборок с территории, релевантной (уместной) для применения</p>	[]
2.9.2	<p>Пригодность для предполагаемого применения:</p> <p>[n/a] Не применим</p> <p>[] Информация не дана</p> <p>[*] Норма или нормы не отвечают требованиям предполагаемого применения</p> <p>[**] Адекватные нормы для общего населения и/или упорядоченный ряд таблиц с нормами</p> <p>[****] Хороший диапазон таблиц с нормами</p> <p>[*****] Отличный диапазон (ряд) релевантных выборок, нормы, связанные с возрастом и полом, с информацией о других различиях между группами (например, группа, смешанная по этническому признаку)</p>	[]

2.9.3	<p>Размеры выборки:</p> <p>[n/a] Не применим</p> <p>[] Информация не дана</p> <p>[*] Не отвечающие требованиям выборки (например, менее 150 человек)</p> <p>[**] Отвечающие требованиям выборки (например, 150-300 человек)</p> <p>[****] Большие выборки (например, 300-1000 человек)</p> <p>[*****] Очень большие выборки (например, 1000 +)</p>	[]
2.9.4	<p>Процедуры, применяемые при отборе выборки: (выберите одно)</p> <ul style="list-style-type: none"> - Информация не представлена - Репрезентативны для популяции [суммарные критерии] - Случайные, несущественные - Выбранные наугад, беспорядочные 	
2.9.5	<p>Качество представленной информации о различиях в групповых меньшинствах, эффектах возраста, пола и т. д.:</p> <p>[n/a] Не применим</p> <p>[] Информация не дана</p> <p>[*] Не отвечающая требованиям информация</p> <p>[**] Отвечающая требованиям общая информация с минимальным анализом</p> <p>[****] Хорошие описания и анализ групп и различий</p> <p>[*****] Отличный диапазон анализа и обсуждение релевантных результатов, относящихся к использованию и интерпретации</p>	[]

2.9.6 Комментарии рецензентов по поводу норм: Короткое сообщение о нормах и их истории, включая информацию о мерах, предусмотренных издателем/автором по поводу модернизации норм на регулярной основе.



Валидность

2.10	Общее соответствие требованиям: (Эта общая оценка достигается путем использования суждений, на которых основаны оценки в пунктах 2.10.1 – 2.10.2.4. Не стремитесь к среднеарифметическим цифрам для получения общей оценки. Обычно она бывает равна либо конструктивной валидности, либо критериальной валидности, неважно какая больше).	
2.10.1	Конструктивная валидность – Общее соответствие требованиям (Эта общая оценка достигается путем использования суждений, на которых основаны оценки в пунктах 2.10.1.2 – 2.10.1.6. Не стремитесь к среднеарифметическим цифрам для получения общей оценки)	[]
2.10.1.1	<p>Реализуемые проекты: <i>(отметьте столько, сколько подходит)</i></p> <p>Информация не представлена Корреляции с другими инструментами и критерии выполнения Внутришкальный (корреляции пунктов) Различия между группами Матрицы множественных признаков и методов Исследовательский факторный анализ Подтверждающий (подкрепляющий) факторный анализ Экспериментальные проекты Другое (укажите)</p>	
2.10.1.2	<p>Размеры выборки:</p> <p>[] Информация не дана</p> <p>[*] Одно не отвечающее требованиям исследование (например, размер выборки меньше чем 100 человек)</p> <p>[**] Одно отвечающее требованиям исследование (например, размер выборки 100-200 человек)</p> <p>[****] Более чем одно отвечающее требованиям или больших размеров исследование</p> <p>[*****] Хороший диапазон больших исследований</p>	[]
2.10.1.3	<p>Процедура отбора выборки <i>(выберите одно)</i></p> <p>Информация не представлена Репрезентативная для популяции (суммированные критерии) Случайная, несущественная Выбранная наугад</p>	

2.10.1.4	<p>Медиана и диапазон корреляций между тестом и другими подобными тестами:</p> <p>[] Информация не дана</p> <p>[*] Не отвечающий требованиям ($r < 0.55$).</p> <p>[**] Отвечающий требованиям ($0.55 < r < 0.65$).</p> <p>[****] Хорошо ($0.65 < r < 0.75$).</p> <p>[*****] Отлично ($r > 0.75$)</p>	[]
2.10.1.5	<p>Качество инструментов как критериев или маркеров:</p> <p>[] Информация не дана</p> <p>[*] Дана не отвечающая требованиям информация</p> <p>[**] Отвечающее требованиям качество</p> <p>[****] Хорошее качество</p> <p>[*****] Отличное качество с широким диапазоном для конвергентной и дивергентной валидации</p>	[]
2.10.1.6	<p>Дифференциальный анализ функционирования пунктов</p> <p>[n/a] Не применим</p>	[]
2.10.2	<p>Критериальная валидность – Общее соответствие требованиям</p> <p>(Эта общая оценка достигается путем использования суждений, на которых основаны оценки в пунктах 2.11.1 – 2.10.4.2. Не стремитесь к среднеарифметическим цифрам для получения общей оценки)</p>	[]
2.10.2.1	<p>Описание использованных критериев и характеристик популяций:</p> <p><i>(выберите столько, сколько применимо)</i></p> <p>Конкурентный Предсказательный Предсказывающий возврат</p>	
2.10.2.2	<p>Размеры выборки:</p> <p>[] Информация не дана</p> <p>[*] Одно не отвечающее требованиям исследование (например, размер выборки меньше чем 100 человек)</p> <p>[**] Одно отвечающее требованиям исследование (например, размер выборки 100-200 человек)</p> <p>[****] Одно большое или более чем одно отвечающее требованиям размера исследование</p> <p>[*****] Хороший диапазон больших исследований</p>	[]

2.10.2.3	Процедура отбора выборки: <i>(выберите одно)</i> Информация не представлена Целевая или репрезентативная Случайная, несущественная Выбранная наугад	
2.10.2.4	Медиана и диапазон корреляций между тестом и критериями: <input type="checkbox"/> Информация не дана <input type="checkbox"/> Не отвечающий требованиям ($r < 0.2$) <input type="checkbox"/> Отвечающий требованиям ($0.2 < r < 0.35$). <input type="checkbox"/> Хорошо ($0.35 < r < 0.50$) <input type="checkbox"/> Отлично ($r > 0.50$)	<input type="checkbox"/>
2.10.3 Комментарии рецензентов по поводу валидности:		





Надежность

2.11	<p>Общее соответствие требованиям: <i>(Эта общая оценка достигается путем использования суждений, на которых основаны оценки в пунктах 2.11.1 – 2.11.3.2. Не стремитесь к среднеарифметическим цифрам для получения общей оценки)</i> <i>Для некоторых инструментов внутренняя согласованность может быть несоответствующей (расплывчатые свойства или объединенные шкалы). В этом случае сделайте акцент на ретестовых данных. В других случаях (измерения состояний) ретестовая надежность может вводить в заблуждение, поэтому акцент следует сделать на внутренней согласованности.</i></p> <p>В отношении критериев надежности два типа критериев берутся во внимание. Инструмент, который сконструирован для индивидуального оценивания, требует более высоких уровней надежности для практической эффективности, чем те, что используются для вынесения решений относительно групп людей. Из предлагаемых достоинств, данных ниже, первое относится к инструментам, предназначенным для принятия решения о группах (например, отборочные инструменты), в то время как второй набор достоинств относится к тем, что предназначены для поэтапного (последовательного) индивидуального оценивания.</p>	[]
2.11.1.	<p>Представленные данные о надежности: <i>(выберите одно)</i></p> <ul style="list-style-type: none"> - Дан только один коэффициент надежности - Дан только один расчет стандартной ошибки измерения Коэффициенты надежности для ряда разных групп - Дана стандартная ошибка измерения для ряда различных групп 	
2.11.1	Внутренняя согласованность::	
2.11.1.1	<p>Размер выборки:</p> <ul style="list-style-type: none"> [] Информация не дана [*] Одно не отвечающее требованиям исследование (например, размер выборки меньше чем 100 человек) [**] Одно отвечающее требованиям исследование (например, размер выборки 100-200 человек) [****] Одно большое или более чем одно отвечающее требованиям размера исследование [*****] Хороший диапазон больших исследований 	[]
2.11.1.2	<p>Медиана коэффициентов:</p> <ul style="list-style-type: none"> [] Информация не дана [*] Не отвечающий требованиям (например, $r < 0.7$) [**] Отвечающий требованиям (например, $r = 0.7$ to 0.79) [****] Хорошо (например, $r = 0.8$ to 0.89) [*****] Отлично (например, $r > 0.9$) 	[]
2.11.2	Ретестовая стабильность теста:	

2.11.2.1	<p>Размер выборки:</p> <p>[] Информация не дана</p> <p>[*] Одно не отвечающее требованиям исследование (например, размер выборки меньше чем 100 человек)</p> <p>[**] Одно отвечающее требованиям исследование (например, размер выборки 100-200 человек)</p> <p>[****] Одно большое или более чем одно отвечающее требованиям размера исследование</p> <p>[*****] Хороший диапазон больших исследований</p>	[]
2.11.2.2	<p>Медиана коэффициентов:</p> <p>[] Информация не дана</p> <p>[*] Не отвечающий требованиям (например, $r < 0.6$)</p> <p>[**] Отвечающий требованиям (например, $r = 0.6$ to 0.69)</p> <p>[****] Хорошо (например, $r = 0.7$ to 0.79)</p> <p>[*****] Отлично (например, $r > 0.8$)</p>	[]
2.11.3	<i>Эквивалентная надежность:</i>	
2.11.3.1	<p>Размер выборки:</p> <p>[] Информация не дана</p> <p>[*] Одно не отвечающее требованиям исследование (например, размер выборки меньше чем 100 человек)</p> <p>[**] Одно отвечающее требованиям исследование (например, размер выборки 100-200 человек)</p> <p>[****] Одно большое или более чем одно отвечающее требованиям размера исследование</p> <p>[*****] Хороший диапазон больших исследований</p>	[]

2.11.3.2	<p>Медиана коэффициентов:</p> <p>[] Информация не дана</p> <p>[*] Не отвечающий требованиям (например, $r < 0.6$)</p> <p>[**] Отвечающий требованиям (например, $r = 0.6$ to 0.69)</p> <p>[****] Хорошо(например, $r = 0.7$ to 0.79)</p> <p>[*****] Отлично (например, $r > 0.8$)</p> <p>[n/a] Не применим</p>	
----------	---	--

Комментарии рецензентов по поводу надежности:

- Комментарий по поводу доверительных интервалов для коэффициента надежности
- Представьте эквиваленты Спирмена-Брауна





.....

2.12 Раздел 8: Качество компьютерных отчетов:

.....

По поводу каждого из следующих свойств поставлено несколько вопросов, которые должны помочь вам вынести суждение, и представлено определение 'отличной' [****] оценки.

2.12	<p>Общее соответствие требованиям: (Эта общая оценка достигается путем использования суждений, на которых основаны оценки в пунктах 2.12.1 – 2.12.7. Не стремитесь к среднеарифметическим цифрам для получения общей оценки)</p>	[]
2.12.1	<p>Диапазон охвата</p> <p>Как видно, отчеты могут варьировать как по широте, так и по специфике. Отчеты могут также варьировать по диапазону людей, для которых они подходят. В некоторых случаях может быть такое, что отдельно сделанные отчеты подходят для разных групп людей.</p> <ul style="list-style-type: none"> • Охватывает ли отчет весь диапазон свойств, которые измеряет инструмент? • Делает ли он это на уровне специфичности, справедливой в терминах уровня детальности, достижимого из результатов инструмента? (Отражает ли отчет в полной мере специфику результатов?) • Может ли дробность отчета (т.е. число отличных групп результатов по шкале, которые используются в отчете, чтобы превратиться в различные текстовые единицы, используемые в отчете) быть справедливой в терминах шкальных ошибок измерения? • Используется ли отчет для тех же самых групп людей, для которых был разработан инструмент? (например, группы, для которых релевантны групповые нормы, или для которых существуют релевантные критериальные данные, и т.д.) <p>[****] Отличное соответствие между диапазоном инструмента и диапазоном отчета с уровнем специфичности в отчете, гармонирующим с уровнем детальности, измеряемой шкалами. Хорошее использование всех данных, которые получены.</p>	[]
2.12.2	<p>Надежность</p> <ul style="list-style-type: none"> • Насколько последовательны (согласованы) отчеты в своей интерпретации похожих групп данных? • Если содержание отчета варьирует (например, при случайном выборе из эквивалентных текстовых единиц), сделано ли это удовлетворительно? • Является ли интерпретация данных и различий между данными справедливой в терминах шкальных ошибок измерения? <p>[****] Отличная согласованность в интерпретации и соответствующих предупреждениях, предусмотренных для утверждений, интерпретации и рекомендаций относительно кроющихся в них ошибок измерения.</p>	[]
2.12.3	<p>Релевантность или валидность</p> <p>Связь между инструментом и содержанием отчета может быть объяснена как в отчете, так и подтверждена отдельно. Там, где отчеты основываются на клинических суждениях, процесс, в результате которого эксперты продуцировали содержание, и правила, устанавливающие связь между данными и содержанием, должны быть подтверждены.</p> <ul style="list-style-type: none"> • Как сильна связь между содержанием отчета и данными инструмента? До какой степени отчет выходит за рамки или расходится с информацией, предоставляемой данными инструмента? • Связано ли четко содержание отчета с характеристиками, измеряемыми инструментом? 	[]

	<ul style="list-style-type: none"> • <i>Предоставляет ли он приемлемые выводы о критериях, с которыми, как можно было бы ожидать, эти характеристики связаны?</i> • <i>Какие эмпирические доказательства предоставляются, чтобы показать, что эти связи действительно существуют?</i> <p>Уместно рассматривать как конструктивную валидность отчета (т.е. уровень интерпретации, который согласуется с определением лежащих в основе конструктов), так и критериальную валидность (т.е. где сделаны утверждения, которые могут быть связаны обратной связью с эмпирическими данными).</p> <p>[****] Отличная связь между шкалами и содержанием отчета с ясными подтверждениями.</p>	[]
2.12.4	<p>Честность или свобода от предубеждения (пристрастности)</p> <ul style="list-style-type: none"> - <i>Создается ли впечатление, что содержание отчета и язык не соответствуют определенным группам?</i> - <i>Проясняет ли отчет какие-либо области возможных предубеждений в результатах инструмента?</i> - <i>Используются ли альтернативные языковые формы? Если да, то сделаны ли адекватные шаги, чтобы гарантировать их эквивалентность?</i> <p>[****] Отличные, ясные предостережения и объяснения относительно возможного предубеждения, имеющиеся на всех уместных языках пользователя.</p>	[]
2.12.5	<p>Доступность</p> <p>Она будет зависеть в большой мере от сложности языка, используемого в отчете, от сложности описываемых конструктов и цели, для которой он предназначен.</p> <ul style="list-style-type: none"> • <i>Существует ли вероятность того, что форма и содержание отчета приемлемы (доступны, понятны) для предполагаемых реципиентов?</i> • <i>Написан ли отчет языком, который соответствует вероятным уровням умения считать и грамотности предполагаемого читателя?</i> <p>[****] Очень высокая доступность, хорошо сконструирован и хорошо подходит предполагаемой аудитории.</p>	[]
2.12.6	<p>Практичность</p> <p>Практические вопросы также влияют на доступность. Главным практическим преимуществом компьютерных отчетов является то, что они экономят время для человека, который должен сделать отчет иным способом. Когда этот человек не конечный пользователь, могут быть сложности при вынесении аргументов в пользу практичности могут быть.</p> <ul style="list-style-type: none"> • <i>Сколько времени экономит для пользователя каждый отчет?</i> • <i>Сколько времени требует прочтение и применение каждого отчета?</i> <p>[****] Отличный в терминах эффективности и ценности.</p>	[]
2.12.7	<p>Продолжительность</p> <p>Она является аспектом практичности и должна быть отражена в ее (практичности)</p>	

	<p>общей оценке. Специфика этой стороны практичности состоит в том, что она обеспечивает индекс соотношения количества отдачи к количеству вложений (вложенного). Число шкал, на которых базируется содержание отчета, рассматривается как вложение, а число страниц отчета (исключая титульные страницы, авторские заметки и т.д.) рассматриваются как отдача.</p> <p><i>Чтобы подсчитать этот индекс, подсчитайте количество шкал, включая производные шкалы и сложные шкалы, (например, для личностных измерений шкалы факторов более высокого порядка, шкалы типов команды, стилей лидерства и т.д. могут быть производны от базовых шкал).</i></p> <ol style="list-style-type: none"> 1. Разделите общее число страниц на число шкал 2. Умножьте это соотношение на 10 и округлите результат до ближайшего целого числа. <p>Обычно оценки больше 10 должны, вероятно, указывать на отчеты, которые являются сверхдлинными и имеют сверхдлинную интерпретацию.</p> <p>Например: Отчет о развитии - $8/7 \times 10 = 11.42$.</p>	
--	--	--



2.12.8	<p>Комментарии рецензентов по поводу компьютерных отчетов:</p> <p>Оценивание может учитывать дополнительные обстоятельства, такие как придается ли в отчетах значение проверке согласованности (последовательности) ответов, измерению предубежденности в ответах (например, измерению центральной тенденции в оценках) и другим индикаторам степени доверия, с которой данные испытуемого интерпретируются.</p> <p>Могут быть включены комментарии по поводу сложности алгоритмов. Например, рассматриваются ли составные (множественные) шкалы одновременно, как рассматриваются профили шкал и т.д. Такая сложность должна быть, конечно, подкреплена разумным объяснением в руководстве.</p> <p>Оценивание компьютерных отчетов осложняется тем фактом, что многие поставщики захотят, понятно, защитить свою интеллектуальную собственность на алгоритмы и правила подсчета. На практике информация, достаточная для целей описания и оценки теста, должна быть доступна из технического руководства, описывающего процесс создания отчета и его теоретические основы, и посредством просмотра выборки примеров (случаев) конфигураций тестовых данных</p> <p>В идеале документация должна также описывать процедуры, которые были применены для проверки подготовки отчета на точность, согласованность и релевантность.</p>
--------	---





.....

Раздел 9: Окончательное оценивание:

.....

3.4	<p>Отчет с оценкой теста:</p> <p>Этот раздел должен содержать точное, ясно аргументированное мнение (оценочное заключение, оценку) об инструменте/продукте. Оно должно описывать все «за» и «против» и давать некоторые общие рекомендации касательно того, как и когда он мог бы быть использован ,вместе с предостережениями (где необходимо) относительно того, когда он не должен быть использован.</p> <p>Оценивание должно охватывать такие темы как соответствие инструмента различным оценочным функциям или областям приложения, необходимость какой-либо специальной подготовки или наличие(я) специальных умений, существование какой-либо важной информации, которая не дана пользователю.</p> <p>Включите комментарии по поводу какого-либо исследования, о котором известно, что оно ведется, и планах поставщика на развитие и усовершенствование инструмента в будущем.</p>
-----	--





Заключение(я):



<p>4.0</p>	<p>Рекомендации (выберите одно)</p> <p>Следует указать релевантную рекомендацию из данного списка. В норме это потребует некоторого комментария, подтверждения или квалификации. К краткому утверждению должны быть добавлены информация о связи с ситуациями и способами, в которых и с помощью которых инструмент мог бы быть использован, и предостережения относительно возможных областей неправильного использования.</p> <p>Все характеристики, перечисленные ниже, должны иметь одну из оценок, если инструмент должен быть рекомендован для общего использования (квадраты 5 и 6).</p> <p>[2.9] Нормы и референтные группы.</p> <p>[2.10.1] Конструктивная валидность</p> <p>[2.10.2] Критериальная валидность</p> <p>[2.11] Надежность – общая</p> <p>[2.12] Компьютерные отчеты</p> <p>Если какая-нибудь из этих оценок будет [] или [*], инструмент в норме будет классифицирован под рекомендацией 1,2,3 или 4 или он будет классифицирован как «Другие» с соответствующим объяснением.</p>	<ul style="list-style-type: none"> <input type="checkbox"/> 1. Только исследовательский инструмент. Не для использования на практике. <input type="checkbox"/> 2. Подходит только для использования экспертом при тщательно контролируемых условиях или в очень ограниченных областях применения. <input type="checkbox"/> 3. Подходит для использования под контролем супервизора в областях применения, определенных дистрибьютером, любым пользователем с общей компетенцией в области использования и проведения тестов. <input type="checkbox"/> 4. Требуется дальнейшего развития. Подходит только для использования в исследовании. <input type="checkbox"/> 5. Подходит для использования в области (областях) применения, определенных дистрибьютером, пользователями тестов, которые отвечают специфическим квалификационным требованиям дистрибьютера. <input type="checkbox"/> 6. Подходит для неконтролируемого самооценки в областях применения, определенных дистрибьютером. <input type="checkbox"/> 7. Другое
------------	--	---

5.	Заметки, ссылки и библиография
<p>Вам следует контролировать стандартные источники для описания (рецензий) для каждого инструмента (например, «Бюро» и Тестовые рецензии (критика) Test Critiques). Вам следует добавить детали любых ссылок, цитированных в вашем Оценивании и перечислить ссылки на любые другие рецензии, о которых вы знаете. Там, где уместно, вы можете добавить краткую аннотацию (не более 50 слов или около того) к каждой из них, касаясь заключений, сделанных в рецензии. Укажите любые другие источники информации, которые могли бы помочь пользователю.</p> 	
See 1.1 0.5	Измеряемые конструкты: