

## Валидность психологического теста

*Х.-В. Гессманн*

*доктор психологических наук, профессор Костромского государственного университета им. Н.А. Некрасова*

*Е.А. Шеронов*

*аспирант Костромской государственного университета им. Н.А. Некрасова*

В данной работе представлен теоретический обзор современных взглядов на проблему валидности психологического теста. В статье аргументирована значимость разработки психологических тестов и необходимость исследования их валидности, описаны различные виды валидности и возможные пути измерения и определения коэффициента валидности. Статья рекомендована для исследователей, чья работа посвящена разработке, модификации или адаптации психологического теста.

**Ключевые слова:** психологический тест; валидность; внутренняя валидность; внешняя валидность; очевидная валидность; содержательная валидность; конструктивная валидность; конвергентная валидность; дискриминантная валидность; прогностическая валидность; текущая валидность; параллельная валидность.

Дифференциальная психология связана с развитием методов измерения межличностных и индивидуальных различий. Здесь с высокой точностью проверяются наиболее важные аспекты.

«В Российском психологическом обществе (РПО) наблюдается переориентация психологических взглядов как в университетах, так и в практике навстречу использованию валидных и научно обоснованных методов и процедур исследования» [14].

Используемый метод психодиагностики должен соответствовать критериям, которые позволяют заявить о качестве теста.

Обязательно должны соблюдаться три критерия: тест должен иметь высокую объективность, надёжность и валид-

ность. Данная статья посвящена изучению валидности.

Психологический тест — это инструмент измерения. Любой инструмент измерения, имеет погрешность. Другими словами, любой психологический тест имеет ошибки измерения, от степени которых зависит качество измерительного инструмента. Цель разработчика тестового испытания — свести к минимуму данную погрешность, повысив тем самым качество теста. Данная цель достигается посредством конструирования качественных тестовых заданий, которые должны соответствовать необходимым критериям валидности.

Линерт и Раатц (Lienert, Raatz) определяют психологический тест как «... научный метод рутинного исследования

одного или более эмпирически определяемых черт личности с целью возможного количественного отчета об относительной степени выражения индивидуальной особенности» [9].

Шмидт-Атцерт и Амеланг (Schmidt-Atzert, Amelang) определяют психологический тест как метод исследования, обладающий следующими свойствами:

- Это метод измерения;
- Применяется для измерения одной или нескольких психологических черт;
- Стандартизированная процедура;
- Включает в себя сбор эмпирических данных;
- Процедура проводится в определённых условиях;
- Целью измерения является количественное выражение качественной характеристики или качественное утверждение о наличии той или иной функции [15].

Любой научный метод измерения должен соответствовать определенным критериям качества. Существуют основные и второстепенные критерии качества. Основными критериями качества являются:

1. Объективность: Независимость теста от условий, в которых он применялся, от человека, проводящего испытание, анализирующего и интерпретирующего полученные результаты [4].

2. Надёжность: согласованность результатов, полученных при каждом повторном выполнении теста одним и тем же испытуемым, с результатами его первого тестирования [1].

3. Валидность: соответствие результатов теста той характеристике, для измерения которой он предназначен [4].

Второстепенными критериями качества являются:

- Полезность: Полезен ли метод и насколько удобен для изучения конкретной характеристики? Иногда это называют внешней валидностью.

- Репрезентативность: Могут ли проходить измерение люди с разным полом, возрастом, местом жительства и т. д.?

- Экономичность: Являются ли затраты на исследование адекватны полученным данным?

- Прозрачность: Является ли инструкция и порядок выполнения исследования понятным для испытуемого? Предусматривается ли процедура объяснения инструкции?

- Фальсификация: Устроен ли процесс исследования таким образом, что испытуемый имеет возможность умышленно исказить результат?

- Рациональность: Насколько рационально предусмотрено время исследования, учтены умственные способности испытуемых и окружающие физические условия?

- Нормализация: Существует ли нормальный показатель по используемому методу? Каким образом можно сравнить полученные от разных людей результаты? [10]

*Валидность* является общим оценочным суждением об адекватности и степени пригодности полученных эмпирическим путём выводов [12]. Важно отметить, что валидность может измеряться только у объективного и надёжного теста. Другими словами, объективность, надёжность и валидность находятся в иерархических отношениях друг к другу. Объективность является необходимым, но не достаточным условием надёжности; надёжность является необходимым, но не достаточным условием валидности; а целью является создание валидного

инструмента измерения. То есть, валидность является заключительным звеном проверки основных критериев качества теста.

Как в психологии, так и в любой другой науке валидное измерение — это такое измерение, которое измеряет то, что оно должно измерять. Другими словами, при измерении математических способностей измерительный инструмент должен быть направлен на изучение именно математических способностей, а ничего другого. Безупречный эксперимент, возможный только в теории, будет обладать идеальной валидностью, то есть его результаты будут на 100% соответствовать реальности, а полученные закономерности будут возможным обобщить без ограничений. Несмотря на то, что валидность не измеряется к каким-либо условных единицах, говорят о степени валидности, подразумевая то, насколько результаты исследования соответствуют поставленной цели.

В отличие от базовой фальсифицируемости (опровержимости теории) и проверяемости начальное утверждение является критерием достоверного описания определённого явления. В контексте эмпирической достоверности исследований, валидность также относится к исследованию отдельных факторов, описанных в причинно-следственной модели. Валидность, следовательно, с одной стороны — это способность определённого действия («Насколько точно измерительный прибор измеряет объект?»), а с другой стороны — это способность производить определённые выводы на основе измерений («В какой степени верно полагать, что X и Y находятся в зависимости?»).

При использовании хороших измерительных приборов, измеренные зна-

чения не зависят от крайних показателей, при условии объективности исследователя. Также, хорошие измерительные приборы обеспечивают получение одинаковых результатов при повторном тестировании, что является критерием надёжности. Третьим критерием валидности является мера того, что данные, полученные при измерении, как и предполагалось, представляют величину, которая должна быть измерена. Только в случае выполнения всех критериев данные могут быть интерпретированы правильно. Срок действия определяется экспертной оценкой. Все критерии качества должны соблюдаться одновременно [3].

Применимо ко всем психологическим тестам выделяют *внешнюю (или практическую) и внутреннюю (или теоретическую) валидность*.

*Внешняя валидность* — это соответствие реального исследования изучаемому явлению или объекту. Основной идеей здесь является вопрос обобщения (индукция). Согласно классической точки зрения, выводы, полученные в результате эмпирического исследования и имеющие высокую степень внешней валидности могут быть обобщены на всю популяцию, для которых исследование было разработано.

Внешняя валидность показывает соотношение между результатами измерительного прибора и эмпирического критерия [16]. Например, исследователь рассматривает отношение своего нового теста лингвистических способностей с учебными оценками субъектов по предмету иностранный язык, проверяя правильность измерений. При условии использования объективной оценки в качестве критерия или экспертной оценки,

говорят о внешней валидности. Внешняя валидность достигается при условии, что исследуемый конструкт диагностируется одновременно при помощи нескольких тестов. Процедура определения конвергентной и дискриминантной валидности являются частными случаями этой категории.

Наиболее распространенные ошибки, влияющие на внешнюю валидность, совершаются при наборе группы испытуемых. Добровольным ли является участие? Как испытуемые узнали о возможности участия (через объявление в газете, сайт и т. д.)? Что мотивирует испытуемых к участию (заинтересованность в теме, в деньгах, и т. д.)? Эти фильтры могут повлиять на качество подбора группы испытуемых. Наиболее распространенная опасность нарушения внешней валидности заключается в искусственности лабораторных экспериментов [5].

Внешняя валидность увеличивается с каждой успешной ревизией результатов, так как повторение с другими испытуемыми (другая возрастная группа, пол, культура и т. д.) или вариации экспериментальных условий накладывают ограничения на применимость выводов [3]. Для статистического анализа результатов повторного исследования используется метод мета-анализа.

Выводы, сделанные на основе эмпирического исследования обладают высокой степенью *внутренней валидности* тогда, когда альтернативные объяснения явления или количества наблюдаемых эффектов могут быть полностью исключены. Внутренняя валидность определяется тогда (при прочих равных условиях), когда изменение зависимой переменной можно четко отнести к измене-

нию независимой переменной (т. е. нет альтернативного объяснения). Чтобы это обеспечить, должны контролироваться все переменные. Таким образом, эффекты не могут быть отнесены к характеристикам субъектов, они должны быть случайным образом распределены в экспериментальных условиях [3].

Факторы, при которых внутренняя валидность находится под угрозой [8]:

- **История.** Любое незапланированное событие между двумя измерениями может иметь нежелательный эффект. Например, при первом измерении, для тестирования нового антидепрессанта погода холодная и дождливая, а при втором измерении погода была тёплой и солнечной.

- **Созревание.** При длительном промежутке между измерениями возможно изменение испытуемых (взросление, развитие).

- **Реактивность.** Испытуемые могут реагировать на сам процесс измерения, например, может возникать привыкание или раздражительность, особенно тогда, когда процесс измерения неприятный.

- **Изменения в измерительном приборе.** Во время одного исследования, характеристики измерительных приборов могут быть изменены. Это может быть, например, путем нарастания скуки у испытуемых во время измерения. Зависимой переменной, может выступать конструкт, измерение которого производится менее точно. Например, для измерения влияния обучения на развитие интеллекта у детей используется тест интеллекта. Обучение является настолько успешным, что дети достигают высоких результатов при следующем тестировании.

- **Регрессия к среднему.** Этим свойством могут перекрываться крайние значе-

ния выборки. Например, наличие большого количества показателей низкого и высокого интеллекта порождает среднее значение, которое нельзя применить ни к одному из испытуемых.

- Выбор неадекватной рандомизации. При рандомизации испытуемых в экспериментальные и контрольные группы могут быть допущены ошибки, таким образом измерение эффекта будут искажённым. Кроме того, вышеописанные эффекты (История, созревание, регрессия к среднему и др.) могут влиять на группы по-разному.

- Недостаточность испытуемых. Использование небольших экспериментальных групп не позволяет считать результаты исследования достоверными.

- Направление причинного вывода. Причинная связь между независимой и зависимой переменной может подвергаться сомнению при наличии третьей переменной.

- Обмен информацией. Когда испытуемые взаимодействуют между тестированиями (например, «я думаю, что я принадлежу к группе плацебо»), возможно искажение результатов последующего тестирования. Испытуемые могут быть не довольны, например, тем, что у испытуемых из другой группы более комфортные (по их мнению) условия тестирования, несмотря на то, что это может быть замыслом экспериментатора. Этот эффект может вызвать демотивацию испытуемых [8].

С классической точки зрения внутренняя и внешняя валидность находится в конфликте: высокая степень внутренней валидности лучше всего достигается в жестко контролируемых и поэтому довольно искусственных (лабораторных) условиях. Однако, в естествен-

ном эксперименте возникает риск неконтролируемого вмешательства, которое может быть просто неучтено исследователем. С другой точки зрения это только кажущееся противоречие. Так как критерии были разработаны на основе исследований индуктивной логики из обобщения эмпирических данных (например, из эксперимента), они находятся на переднем плане. Здесь возникает вопрос о воспроизводимости результатов в различных условиях на различной выборке. Тем не менее, исследования дедуктивной логики имеют другую цель. В целом эта теория пытается подвергнуть фальсификации полученные выводы, а не проверять логику эмпирического исследования. Согласно этой логике, возникшее противоречие сфальсифицировано. Доказанная эмпирическим путём теория должна подвергнуться дополнительному испытанию. Возражения, которые ставят под сомнение достоверность результатов эксперимента заключаются в вопросе о внутренней валидности эксперимента [3]. Обоснованность эмпирических выводов о причинно-следственных связях всегда подвергается тщательной проверке в экспериментальных и квази-экспериментальных исследованиях.

В 1954 году Американской психологической ассоциацией были предложены рекомендации по конструированию психологических тестов и других диагностических методов. Они предложили четыре типа валидности: *очевидная валидность*, *содержательная валидность*, *конструктивная валидность* и *прогностическая валидность* [10]. То есть, кроме оценки внутренней и внешней валидности психологический тест должен быть оценен и по данным типам валидности.

*Очевидная валидность* — это представление о тесте, сфере его применения, результативности и прогностической ценности, которое возникает у испытуемого или другого лица, не располагающего специальными сведениями о характере использования и целях методики [1]. Другими словами, очевидная валидность — это валидность с точки зрения испытуемого, то есть насколько по его мнению соответствует стимульный материал определённой ситуации исследования.

- Высокая очевидная валидность является необходимым показателем. Она выступает в качестве мотиватора, побуждающего испытуемых к выполнению заданий теста. Очевидная валидность способствует более ответственному отношению к исследованию, к выполнению заданий теста и к заключениям, формулируемым психологом на основе проведённого исследования.

- Представление пользователей методики об очевидной валидности в небольшой степени зависит от её названия, инструкции, стимульного материала [3]. Для формирования адекватного мнения испытуемых следует избегать размытых формулировок, которые могут быть истолкованы неоднозначно. Очевидная валидность повышается путём краткого описания методики и цели исследования. При наполнении содержания методики важно обратить внимание на соответствие терминологии стимульного материала лексикону выборки. Необходимо обладать информацией о возрасте испытуемых, половом составе, социальном статусе, уровне образования, профессиональной специфике испытуемых и т. д.

*Содержательная валидность* отражает степень репрезентативности содержания

заданий теста измеряемой области психических свойств. Содержательная валидность предполагает, что способ измерения конкретного объекта или характеристики напрямую связан с деятельностью этого объекта. То есть, например, чтобы измерить способность водить автомобиль, нужно непосредственно провести экзамен по вождению. Поэтому содержательную валидность часто называют логической или тривиальной валидностью.

Содержательная валидность психологического теста закладывается при составлении стимульного материала [3]. Первым этапом является изучение вида деятельности испытуемого и его собственных свойств. Далее разрабатывается модель тестовой деятельности на основе наиболее важных элементов реальной деятельности. И, наконец, на последнем этапе проводится анализ соответствия разработанной модели деятельности реальной деятельности испытуемого.

Использование большого количества экспертных оценок существенно повышает уровень содержательной валидности психологического теста [3].

*Конструктивная валидность* заключается в степени адекватности метода интерпретации экспериментальных данных теории, которая определяется правильностью употребления терминов той или иной теории. Конструктивная валидность характеризуется способностью теста к измерению такой черты, которая была обоснована теоретически (как теоретический конструкт). Под конструктом понимается теоретический размер функции (латентной переменной). Конструктивная валидность характеризуется способностью теста к измерению функции, которая была обоснована теорети-

чески, как теоретический конструкт. Это возможно, когда диапазон значений конструкции показан полным, точным и понятным. Проверка валидности здесь заключается в том, что тест не должен измерять никакой черты, с которой он теоретически и не должен быть связан. Эмпирическим индикатором конструктивной валидности является конвергентная и дискриминантная (или дивергентная) валидность.

*Конвергентная валидность* отражает уровень ожидаемой связи тестовых оценок с другими показателями (при условии независимости переменных). Основной целью конвергентной валидации является определение того, будут ли оценки теста варьироваться в соответствии с ожиданиями. Для обеспечения полного доказательства валидности необходимо произвести конвергентную валидацию. Эмпирический подход к конвергентной валидации является частным случаем. Например, если для измерения определённого явления существует несколько методик, то целесообразно было бы провести хотя бы две и сравнить полученные результаты. Если данные, полученные от разных методик, будут иметь высокую корреляцию, то можно говорить о конструктивной валидности теста.

В психологии, проверка дискриминантной валидности заключается в том, что тест не должен измерять никакой черты, с которой он теоретически и не должен быть связан.

Кэмпбелл и Фиск [6] ввели понятие дискриминантной валидности в их обсуждении вопроса оценки валидности психологического теста. Они подчеркнули важность использования как конвергентной, так и дискриминантной ва-

лидности при оценке новых психологических тестов. Успешная оценка дискриминантной валидности показывает, что результаты теста не имеют значимой корреляции с результатами других тестов предназначенных для измерения теоретически разных понятий.

Факторы снижения конструктивной валидности [8]:

- Расплывчатое определение измеряемого конструкта;
- Моно-операция: рассматривается только один аспект измеряемого конструкта;
- Моно-метод: при изучении конструкта используется только один метод;
- Переоценка (эффект Хоторна);
- Эффект социальной желательности;
- Ожидания экспериментатора (эффект Розенталя);
- Пренебрежение некоторыми факторами;
- Несколько независимых переменных;
- Взаимодействие между измерением и лечением;
- Ограниченные обобщения аналогичных переменных.

Конструктивная валидность является наиболее сложным видом валидности психологического теста. Валидность существенно понижается при неточном описании измеряемого конструкта с использованием абстрактных терминов.

Чтобы данные, полученные в результате измерения, возможно было использовать для дальнейших исследований психологический тест должен обладать *прогностической валидностью*. Прогностическая валидность отражает насколько точно методика может диагностировать психологическое качество спустя некоторое время после измерения, то

есть является индикатором временного интервала актуальности результатов исследования [3]. Разумеется, прогностическая валидность зависит не только от качества методики, но и от характера измеряемого признака.

Уровень прогностической валидности может использоваться как в широком смысле, означая прогноз на основе данных тестирования поведения испытуемого в реальной (критериальной) ситуации, так и в более узком смысле, указывая лишь на временной интервал [1].

Процедура определения прогностической валидности опирается на анализ корреляции между оценками по тесту и предсказанными результатами деятельности, формирующимися свойствами личности, исходом лечения и т. п. [1].

В психометрии, прогностическая валидность определяется как степень соответствия прогнозов, полученным с помощью исследуемого теста, прогнозируемой мере [7].

Например, валидность когнитивного теста для выявления профессиональной пригодности заключается в корреляции между тестовым баллом и профессиональным успехом. То есть, при приёме на работу абитуриенты проходят исследуемый тест, а спустя какое-то время, например, через 1 год результаты теста сравниваются с объективными показателями успеха принятых работников. Такой когнитивный тест будет иметь прогностическую валидность в случае, если наблюдаемые корреляции окажутся статистически значимыми.

Как и во многих других аспектах социальных наук, величина корреляции, полученные из исследования прогностической валидности, как правило, не высока. Типичная прогностическая валид-

ность для тестов на профессиональную пригодность имеет корреляционный показатель в окрестности  $r = 0,35$ . Тем не менее, полезность использования такого теста для прогноза профессиональной успешности весьма значительна [13].

Точность прогноза и, следовательно, суждения о прогностической валидности находится в обратной зависимости от времени, заданного для такого прогнозирования. Обоснование отдаленной экстраполяции данных теста требует учета большего количества факторов, чем оценка диагностической значимости теста [1]. Поэтому, перед оценкой прогностической валидности целесообразно оценить *текущую валидность*. Текущая валидность отражает точность оценки настоящего положения измеряемого признака.

Тогда, когда результат теста хорошо коррелирует с мерой, которая ранее уже была подтверждена, проявляется *параллельная валидность*. Оба измерения могут отражать один и тот же, либо два разных, но тесно связанных конструкта. В этом заключается отличие параллельной валидности от прогностической валидности, где одно измерение происходит ранее с целью спрогнозировать результат второго измерения [11]. Параллельная валидность используется для проверки исследований, в которых две меры измеряются приблизительно в одно и то же время. Например, результаты проведённого при приёме на работу теста могут быть соотнесены с результатами уже работающих специалистов в тот же день или на той же неделе. Полученное соотношение будет одновременно являться коэффициентом достоверности. Прогностическая валидность отличается только тем, что время между тестирова-

ниями составляет несколько месяцев или даже лет. В приведенном выше примере, прогностическая валидность будет лучшим выбором для проверки профессиональной пригодности, поскольку такие тесты предназначены для прогнозирования производительности на рабочем месте.

«Современный тест — это не только надежный, но и валидный тест, однако не на все случаи жизни, а разработанный для конкретной цели; а современный психологический тест — это теоретически и эмпирически обоснованная система высказываний (заданий), позволяющая получить измерения соответствующих психологических свойств» [2]. Вы-

воды, сделанные на основе эмпирического исследования (обычно о причинно-следственной связи) с использованием психологических тестов обладают высокой степенью статистической достоверности тогда, когда надежность измерительных приборов и выбранного статистического метода высока и, как правило, когда допустима низкая вероятность ошибки, то есть математические расчёты (например, корреляционные матрицы) не искажены. Показатель внутренней и внешней валидности даёт основания полагать о верности этих выводов. Заключение о валидности обсуждаемо и всегда может быть подвержено сомнениям и проверке.

## ЛИТЕРАТУРА

1. Анастаси А., Урбина С. Психологическое тестирование. СПб: Питер, 2007. 688 с.
2. Бодалев А.А., Столин А.В. Общая психодиагностика. СПб: Речь, 2000. 440 с.
3. Гессманн Х.-В. Конструирование психологических тестов. Дуйсбург: Издательство Психотерапевтического института Бергерхаузен. 2013. 576 с.
4. Клайн П. Справочное руководство по конструированию тестов. Киев: ПАН Лтд, 1994. 288 с.
5. Aronson E., Wilson T.D., Akert R.M. Sozialpsychologie. Pearson Studium. 6. Auflage. Weinheim: Beltz, 2008. 688 s.
6. Campbell D.T., Fiske D.W. Convergent and discriminant validation by the multitrait-multimethod matrix // Psychological Bulletin. 1959. № 56. P. 81—105.
7. Cronbach L.J., Meehl P.E. Construct validity for psychological tests // Psychological Bulletin. 1955. № 52. S. 281—302.
8. Krauth J. Experimental Design. Amsterdam: Elsevier Science; Saunders, 2000. 296 s.
9. Lienert G.A., Raatz U. Testaufbau und Testanalyse. 5., vollig neubearbeitete und erweiterte Auflage. Weinheim: Beltz, 1994. 423 s.
10. Lienert G.A., Raatz U. Testaufbau und Testanalyse. 6. Auflage. Weinheim: Psychologie Verlags Union, 1998. 432 s.
11. McIntire S.A., Miller L.A. Foundations of Psychological Testing. 2nd edition. London: Sage Publishing Co., 2005. 613 p.
12. Messick S. Validity // Educational measurement / R. L. Linn (Ed.). 3rd ed. New York: Macmillan, 1989. P. 13—103.
13. Messick S. Validity of Psychological Assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning // American Psychologist. 1995. Vol. 50, Iss. 9, September. P. 741—749.
14. Psychology in Russia: State of the Art. Scientific Yearbook / Ed. By Yu.P. Zinchenko, V.F. Petrenko. Moscow: Lomonosov State University; Russian Psychological Society, 2011. 519 p.
15. Schmidt-Atzert L., Amelang M. Psychologische Diagnostik und Intervention. 5 Auflage. Marburg: Springer Verlag, 2006. 624 s.
16. Schnell R., Hill P.B., Esser E. Methoden der Empirischen Sozialforschung. 8., unveränderte Auflage. Munchen: Oldenbourg Verlag, 2008. 589 s.

## Psychological Test Validity

*H.-W. Gessmann*

*Ph.D in psychology, professor of Kostroma State University named after N.A. Nekrasov*

*E.A. Sheronov*

*Post-graduate student, Kostroma State University named after N.A. Nekrasov*

This paper presents a theoretical overview of current aspects about the validity problems of psychological tests. The article demonstrates the importance of the development of psychological tests and the scientific studies of their validity, describes the different types of validity and the possible ways of measurement and determination of the validity coefficients. The paper is recommended for researchers, whose work is dedicated to the development, modification or adaptation of the psychological test.

**Keywords:** psychological test; validity; internal validity; external validity; face validity; content validity; construct validity; convergent validity; discriminant validity; predictive validity; the current validity; concurrent validity.

### REFERENCES

1. *Anastazi A., Urbina S.* Psikhologicheskoe testirovanie. SPb: Piter, 2007. 688 s.
2. *Bodalev A.A., Stolin A.V.* Obshchaia psikhodiagnostika. SPb: Rech', 2000. 440 s.
3. *Gessmann H.-W.* Konstruirovaniye psikhologicheskikh testov. Duisburg: Izdatel'stvo Psikhoterapevticheskogo instituta Bergerkhauzen. 2013. 576 s.
4. *Klain P.* Spravochnoe rukovodstvo po konstruirovaniyu testov. Kiev: PAN Ltd, 1994. 288 s.
5. *Aronson E., Wilson T.D., Akert R.M.* Sozialpsychologie. Pearson Studium. 6. Auflage. Weinheim: Beltz, 2008. 688 s.
6. *Campbell D.T., Fiske D.W.* Convergent and discriminant validation by the multitrait-multimethod matrix // Psychological Bulletin. 1959. № 56. P. 81—105.
7. *Cronbach L.J., Meehl P.E.* Construct validity for psychological tests // Psychological Bulletin. 1955. № 52. S. 281—302.
8. *Krauth J.* Experimental Design. Amsterdam: Elsevier Science; Saunders, 2000. 296 s.
9. *Lienert G.A., Raatz U.* Testaufbau und Testanalyse. 5., vollig neubearbeitete und erweiterte Auflage. Weinheim: Beltz, 1994. 423 s.
10. *Lienert G.A., Raatz U.* Testaufbau und Testanalyse. 6. Auflage. Weinheim: Psychologie Verlags Union, 1998. 432 s.
11. *McIntire S.A., Miller L.A.* Foundations of Psychological Testing. 2nd edition. London: Sage Publishing Co., 2005. 613 p.
12. *Messick S.* Validity // Educational measurement / R. L. Linn (Ed.). 3rd ed. New York: Macmillan, 1989. P. 13—103.

13. *Messick S.* Validity of Psychological Assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning // *American Psychologist*. 1995. Vol. 50, Iss. 9, September. P. 741—749.
14. *Psychology in Russia: State of the Art. Scientific Yearbook / Ed. By Yu.P. Zinchenko, V.F. Petrenko.* Moscow: Lomonosov State University; Russian Psychological Society, 2011. 519 p.
15. *Schmidt-Atzert L., Amelang M.* Psychologische Diagnostik und Intervention. 5 Auflage. Marburg: Springer Verlag, 2006. 624 s.
16. *Schnell R., Hill P.B., Esser E.* Methoden der Empirischen Sozialforschung. 8., unveränderte Auflage. Munchen: Oldenbourg Verlag, 2008. 589 s.