ФЕДЕРАЛЬНОЕ АГЕНСТВО ПО ОБРАЗОВАНИЮ

ВЛАДИВОСТОКСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ЭКОНОМИКИ И СЕРВИСА

ИНСТИТУТ ЗАОЧНОГО И ДИСТАНЦИОННОГО ОБУЧЕНИЯ

КАФЕДРА ПСИХОЛОГИИ

КОНТРОЛЬНАЯ РАБОТА

по дисциплине Психодиагностика

Валидность теста

Студент

Т.А. Карпова

Владивосток 2007

**Содержание**

Введение

1. Понятие теста

2. Показатели качества тестов (методик)

3. Виды валидности и её нахождение

4. Проблемы валидности

Заключение

Список использованных источников

**Введение**

Термин «тест», получивший чрезвычайно широкое распространение в различных областях знания в смысле испытания, проверки, имеет давнюю историю. По Р. Пэнто и М. Гравитц (1972), слово «тест» происходит из старофранцузского языка и является синонимом слова «чашка» (лат. testa – ваза из глины). Этим словом обозначали небольшие сосуды из обожжённой глины, использовавшие алхимиками для проведения опытов. В русском языке слово «тест» долгое время имело два значения:

1) испытательная присяга, религиозная английская клятва, которую каждый вступающий в общественную должность должен давать, чтобы доказать, что он не тайный католик;

2) плоский плавильный сосуд или сосуд из выщелоченной золы для выделения олова из золота или серебра (Русский энциклопедический словарь. 1877).

Близкое современному содержание термин «тест» как термин психологический получает в конце XIX в.

Тестами в психологии называют стандартизированные методики психодиагностики, позволяющие получать сопоставимые количественные и качественные показатели степени развитости изучаемых свойств.

Под стандартизированностью таких методик имеется ввиду то, что они всегда и везде должны применяться одинаковым образом, начиная от ситуации и инструкции, получаемой испытуемым, кончая способами вычисления и интерпретации получаемых показателей. Сопоставимость означает, что оценки, получаемые при помощи теста, можно сравнивать друг с другом независимо от того, где, когда и кем они были получены, если, разумеется, тест применялся правильно. Из всех возможных психодиагностических методик к тестам предъявляются самые строгие требования, касающиеся валидности, надежности, точности и однозначности.

Существует много разновидностей тестов, которые разделяются на группы по нескольким основаниям: по предмету тестирования (тому качеству, которое оценивается с помощью данного теста); по особенностям используемых в тесте задач; по материалу, предъявляемому испытуемым; по объекту оценивания.

**1. Понятие теста**

В психодиагностике известны разнообразные классификации тестов. Они могут подразделяться по особенностям используемых тестовых заданий на тесты вербальные и тесты практические, по форме процедуры обследования – на тесты групповые и индивидуальные, по направленности – на тесты способностей, тесты личности и тесты отдельных психических функций, а в зависимости от наличия или отсутствия временных ограничений – на тесты скорости и тесты результативности. Также тесты могут различаться по принципам их конструирования. За последние десятилетия многие известные тесты были приспособлены к среде компьютера (предъявление, обработка данных и др.), их можно обозначить как компьютеризированные тесты. Активно разрабатываются тесты компьютерные, изначально конструируемые с учётом возможностей современной вычислительной техники. Тесту, как и любому другому инструменту познания, присущи особенности, которые в конкретных обстоятельствах исследования могут рассматриваться в качестве его достоинств и недостатков. Эффективное использование тестов зависит от учета многих факторов, из которых к важнейшим относятся: теоретическая концепция, на которой базируется тот или иной тест; область применения; весь комплекс ведений, обусловленных стандартными требованиями к психологическим тестам, их психометрическим характеристикам. Распространённые представления о «простоте» и доступности тестов не соответствуют действительности. Являясь средством исследования сложнейших психических явлений, тест не может толковаться упрощённо как предложение задания и регистрация его решения. Научное использование тестов возможно лишь при условии опоры на общепсихологические знания, компетентность в области теории и практики соответствующих психодиагностических исследований. Не менее существенно следование этическим нормам психодиагностики.

**2. Показатели качества тестов (методик)**

Особое значение имеет выбор критерия оценки получаемых результатов. Сами по себе, без сопоставления с точкой отсчёта, результаты психологических испытаний ничего не значат. Чаще всего они сопоставляются с нормами выполнения в выборке стандартизации (репрезентативная выборка того типа, для которого данная методика предназначена, например, дети определённого возраста, студенты высших учебных заведений сходного профиля, представители той или иной профессии и др.). Эти нормы, таким образом, устанавливаются опытным путём. Затем результаты каждого испытуемого сопоставляются с нормой и квалифицируются как средние, высокие и низкие. Это позволяет определить положение конкретного испытуемого относительно нормативной выборки или выборки стандартизации. Для этого «сырые» баллы (т.е. количество правильных ответов) с помощью специальных таблиц переводятся в стандартные показатели, сопоставимые друг с другом и с нормой. В данном случае используется статистическая норма – средний диапазон значений на шкале измеряемого свойства – это тот уровень, который характеризует статистически среднего человека. Однако в последние годы такой способ оценки результатов тестирования подвергается критике.

Полезнее ориентироваться на систему требований, которую общество на определённом этапе своего развития предъявляет каждому из его членов. Это могут быть требования к умственному, эстетическому, нравственному развитию человека. Чтобы жить в обществе, человек должен отвечать данным требованиям. Содержание таких социально-критериальных норм (социально-культурных нормативов) вполне реально, оно присутствует в образовательных программах, в квалифицированных профессиональных характеристиках, общественном мнении, мнении педагогов, воспитателей. В качестве точек отсчёта могут быть использованы и другие критерии, например, возрастные нормы, критерии выполнения деятельности и др.

Важнейшими показателями качества психодиагностических методик являются надёжность и валидность. Наша задача – рассмотреть, что же такое валидность теста.

**3 Виды валидности и её нахождение**

По определению американского текстолога А.Анастази, «валидность теста – понятие, указывающее нам, что тест измеряет и насколько хорошо он это делает». Валидность свидетельствует о том, пригодна ли методика для измерения определённых качеств, особенностей и насколько эффективно она это делает. Наиболее распространённым способом нахождения теоретической валидности теста (методики) является конвергентная валидность, то есть сопоставление данной методики с авторитетными родственными методиками и доказательство значимых связей с ними.

Сопоставление с методиками, имеющими другое теоретическое основание, и констанция отсутствия значимых связей с ними называется дискриминантной валидностью. Другой вид валидности – прагматическая валидность – проверка методики с точки зрения её практической значимости, эффективности, полезности. Для проведения такой проверки, как правило, используются так называемые независимые внешние критерии, то есть используется независимый от теста, внешний источник информации о проявлении в реальной жизни и деятельности людей измеряемого психического свойства. Среди таких внешних критериев могут быть успеваемость, профессиональные достижения, успехи в разных видах деятельности, субъективные оценки (или самооценки). Если например, методика измеряет особенности развития профессионально важных качеств, то для критерия необходимо найти такую деятельность или отдельные операции, где именно эти качества реализуются.

Для проверки валидности теса можно использовать метод известных групп, когда приглашаются люди, про которых известно, к какой группе по критерию они относятся (например, группа «высокоуспешных, дисциплинированных студентов» - высокий критерий и группа «неуспевающих, недисциплинированных студентов» - низкий критерий, а студенты со средними значениями не участвуют в тестировании), проводят тестирование и находят корреляцию между результатами теста и критерием.

|  |  |  |
| --- | --- | --- |
|  | высокий критерий | низкий критерий |
| высокий результат теста | a | b |
| низкий результат теста | c | d |

Здесь a – число испытуемых, попавших в высокую группу по тесту и по критерию, c – число испытуемых, попавших в высокую группу по критерию и имеющих низкие результаты теста. При полной валидности теста элементы b и c должны быть равны нулю. Меру совпадения, корреляции между крайними группами по тесту и критерию оценивают с помощью фи-коэффициента Гилфорда. Существует достаточно много разных способов доказательства валидности теста. Тест называется валидным, если он измеряет то, для измерения чего предназначен. Валидность внешняя – применительно психодиагностических методик означает соответствие результатов психодиагностики, проводимой посредством данной методики, независимым от методики внешним признакам, относимым к субъекту обследования. Означает примерно то же самое, что валидность эмпирическая, с той разницей, что здесь речь идёт о связи между показателями методики и самыми важными, ключевыми внешними признаками, относимыми к поведению испытуемого. Психодиагностическая методика считается внешне валидной, если, например, её посредством оцениваются черты характера индивида и его внешне наблюдаемое поведение согласуется с результатами проведённого тестирования.

Валидность внутренняя – применительно психодиагностических методик означает соответствие содержащих в ней заданий, субтестов; соответствие результатов психодиагностики, проводимой посредством данной методики, определению оцениваемого психологического свойства, использованному в самой методике. Методика считается внутренне не валидной или недостаточно валидной, когда все или часть включённых в неё вопросов, заданий и субтестов измеряют не то, что требуется данной методикой. Очевидная валидность – описывает представление о тесте, сложившиеся у испытуемого. Тест должен восприниматься обследуемым как серьёзный инструмент познания его личности. Очевидная валидность приобретает особое значение в современных условиях, когда представление о тестах в общественном сознании формируется многочисленными публикациями в популярных газетах и журналах того, что можно назвать квазитестами, с помощью которых читателю предлагается определить всё, что угодно: от интеллекта до совместимости с будущим супругом.

Конкурентная валидность оценивается по корреляции разработанного теста с другими, валидность которых относительно измеряемого параметра установлена. П. Клайн отмечает, что данные о конкурентной валидности полезны тогда, когда есть неудовлетворительно работающие тесты для измерения некоторых переменных, а новые создаются для того, чтобы улучшить качество измерения. В самом деле, если уже существует эффективный тест, то для чего нужен такой же новый? Прогностическая валидность устанавливается с помощью корреляции между показателями теста и некоторым критерием, характеризующим измеряемое свойство, но в более позднее время. Например, прогностическая валидность какого-нибудь теста интеллекта может быть показана корреляцией его показателей, полученных у испытуемого в возрасте 10 лет, с академической успеваемостью в период окончания средней школы. Л. Кронбах считает прогностическую валидность наиболее убедительным доказательством того, что тест измеряет именно то, для чего он был предназначен. Основная проблема, с которой сталкивается исследователь, пытающийся установить прогностическую валидность своего теста, состоит в выборе внешнего критерия. В особенной степени чаще всего это касается измерения личностных переменных, где подбор внешнего критерия – чрезвычайно сложная задача, решение которой требует немалой изобретательности. Несколько проще обстоит дело при определении внешнего критерия для когнитивных тестов, однако и в этом случае исследователю приходится «закрывать глаза» на многие проблемы. Так, академическая успеваемость традиционно используется в качестве внешнего критерия при валидизации тестов интеллекта, но в то же время хорошо известно, что успехи в обучении далеко не единственное свидетельство высокого интеллекта. Инкрементная валидность имеет ограниченное значение и относится к случаю, когда один тест из батареи тестов может иметь низкую корреляцию с критерием, но не перекрываться другими тестами из этой батареи. В этом случае данный тест обладает инкрементной валидностью. Это может быть полезно при проведении профотбора с помощью психологических тестов. Дифференциальная валидность может быть проиллюстрирована на примере тестов интересов. Тесты интересов обычно коррелируют с академической успеваемостью, но по-разному для разных дисциплин. Значение дифференциальной валидности, так же как и инкрементной, ограничено.

Содержательная валидность определяется через подтверждение того, что задания теста отражают все аспекты изучаемой области поведения. Обычно она определяется у тестов достижений (смысл измеряемого параметра полностью ясен), которые как еже указывалось, тестами, собственно психологическими не являются. На практике для определения содержательной валидности подбираются эксперты, которые указывают, какая область поведения наиболее важна, например, для музыкальных способностей, а затем, исходя из этого, генерируются задания теста, которые вновь оценивают эксперты. Конструктная валидность теста демонстрируется полным, насколько это возможно, описанием переменной, для измерения которой предназначается тест. По сути дела, конструктная валидность включает в себя все подходы к определению валидности, которые были перечислены выше. Кронбах и Мил, которые ввели в психодиагностику понятие конструктной валидности, пытались решить проблему отбора критериев при валидизации теста. Они подчёркивали, что во многих случаях ни один отдельно взятый критерий не может служить для валидизации отдельного теста. Можно считать, что решение вопроса о конструктной валидности теста представляет собой поиск ответа на два вопроса:1) существует ли реально некоторое свойство; 2) надёжно ли измеряет данный тест индивидуальные различия по этому свойству. Вполне понятно, что с конструктной валидностью связана проблема объективности в интерпретации результатов по изучению конструктной валидности, однако эта проблема общепсихологическая и выходит за рамки валидности.

**4. Проблемы валидности**

Из вышесказанного следует, что не существует какого-либо единичного показателя, с помощью которого устанавливается валидность психологического теста. В отличие от показателей надёжности и дискриминативности, нельзя осуществить точные статистические расчёты, подтверждающие валидность методики. Тем не менее, разработчик должен представить весомые доказательства в пользу валидности теста, что потребует от него психологических знаний и интуиции. Несмотря на то, что проблеме валидности в классической теории теста уделяется много внимания, теоретически она никак не решается. Приоритет отдан надёжности, что и выражено в правиле: валидность теста не может быть больше его надежности. Валидность означает пригодность теста измерять то свойство, для измерения которого он предназначен. Следовательно, чем больше на результат выполнения теста или отдельного задания влияет измеряемое свойство и чем меньше – другие переменные (в том числе внешние), тем тест валидней. Валидность – одна из важнейших характеристик психодиагностических методик и тестов, один из основных критериев их качества. Это понятие близко к понятию достоверности, но не вполне тождественно. Проблема валидности возникает в ходе разработки и практического применения теста или методики, когда требуется установить соответствие между степенью выраженности интересующего свойства личности и методом его измерения. Валидность указывает, что именно тест или методика измеряет и насколько хорошо это делает; чем они валиднее, тем лучше отображается в них то качество, ради измерения которого они создавались. Количественно валидность может выражаться через корреляции результатов, полученных с помощью теста или методики, с другими показателями, - например, с успешностью выполнения соответственной деятельности. Валидность можно обосновать разными путями, чаще всего – комплексно. Применяются также дополнительные понятия валидности концептуальной, критериальной, конструктивной, и прочие виды валидности – со своими способами установления их уровня. Требование валидности весьма важно, и многие нарекания в адрес тестов или иных психодиагностических методик связаны с сомнительностью их валидности. Например, валидность измерения интеллекта зависит:

1) от определения понятия интеллекта, являющего ту или иную концепцию этого феномена;

2) от состава тестовых заданий, разрабатываемых согласно этой концепции;

3) от эмпирических критериев.

Разные концепции требуют различного состава заданий, поэтому важен вопрос валидности концептуальной. Чем больше задания соответствуют данной авторской концепции интеллекта, тем уверенней можно говорить о концептуальной валидности теста. Корреляция теста в эмперическим критерием указывает на его возможную валидность по отношению к данному критерию. Определение валидности теста всегда требует постановки дополнительных вопросов: валидность для чего? Для какой цели? По какому критерию?

**Заключение**

Понятие валидности отражает прагматическую идею обязательной практической проверки качества измерения психологических свойств человека. Если раньше понятие валидности чаще других ассоциировалось с понятием «качество теста», то в последние годы всё больше осознаётся роль интерпретации тестовых результатов. Таким образом, валидность – это прежде всего, адекватность интерпретации тестовых результатов по отношению к цели тестирования, содержанию основного понятия (содержанию диагносцируемого психического свойства) и применению результатов тестирования (в профотборе и других прикладных психологических обседованиях). В контексте прикладной задачи валидность представляет собой не столько критерий качества теста, сколько показатель меры соответствия теста цели тестирования. Итак, понятие валидности относится не только к тексту, но и к критерию оценки его качества. Чем выше коэффициент корреляции теста с критерием, тем выше валидность. Развитие факторного анализа позволило создавать тесты, валидные по отношению к идентифицируемому фактору. Только проверенные на валидность тесты можно использовать в профессиональной ориентации, профессиональном отборе и в научных исследованиях.

**Список использованных источников**

1. Анастази А. Психологическое тестирование. – СПб, 2001.

2. Бурлачук Л.Ф. Психодиагностика: Учебник для вузов.- СПб.: Питер, 2005.

3. Столяренко Л.Д. Основы психологии. Практикум. – Ростов-на-Дону, 2005.

4.Общая психология / Под ред. А.В. Петровского. – Москва, 2005.