

## ВАЛИДНОСТЬ КАК ОДИН ИЗ КРИТЕРИЕВ КАЧЕСТВА ТЕСТА

ЖУНУСАКУНОВА А.Д.

[bibliosu@rambler.ru](mailto:bibliosu@rambler.ru)

*В статье рассматриваются понятия о валидности, виды валидности применяемых в педагогических измерениях.*

Валидность является наиболее важным аспектом качества теста и определяется для образовательного и психологического тестирования как «приемлемость, осмысленность, полезность выводов, которые могут быть сделаны по результатам тестирования». Валидность можно рассматривать как интегрированное оценивающее суждение относительно того, в какой мере эмпирические доказательства и теоретические объяснения подтверждают адекватность и уместность выводов и следствий, которые базируются на тестовых баллах или других результатах оценивания.

Валидность – это понятие указывающее нам, что тест измерят и насколько он хорошо это делает [1,с.133]. Она зависит от качества знаний, их числа, от степени полноты и глубины охвата содержания учебной дисциплины в заданиях теста. От баланса и распределения заданий по трудности, от метода отбора заданий из общего бланка, от интерпретации результатов тестирования, от организации сбора данных, от подбора выборочной совокупности испытуемых

Валидность является целезависимой специфической характеристикой, а именно тест может быть валидным для использования в одних ситуациях, целях, условиях и невалидным для других. Тест может быть валидным по одной своей характеристике (например, оценивать правильную область домена, который диагностируется) и невалидным по другой (например, оценивания знания на не соответствующем таксономическом уровне). Доказательства валидности теста должны быть представлены в контексте, в котором они были получены. Поэтому желательна разносторонняя проверка валидности.

Чтобы установить, действительно ли измерил то, что хотели, разумеется, требуется критерии. В практике педагогических измерений выделяются следующие виды валидности [2,с.32-33;4,с.41;6,с.335]:

1. очевидная валидность;
2. содержательная валидность;
3. конструктурная (концептуальная) валидность;
4. валидность по критерию (критериальная или эмпирическая валидность);
5. сравнительная или конкурентная валидность;
6. валидность прогноза и др.

### 1. Очевидная валидность.

Очевидная валидность с точки зрения испытуемого, она определяет, насколько адекватно впечатление о предмете измерения, которое формируется у испытуемого при знакомстве с инструкцией и материалом теста тому, что данный инструмент в действительности выявляет. Очевидная валидность (иногда называется доверительной валидностью) играет важную роль в процессе тестирования, поскольку именно она определяет отношение испытуемого к обследованию. Так, тест может не вызвать доверия, если в нем отсутствуют знакомые, традиционно используемые испытуемым термины, понятия, слова, определения.

### 2. Содержательная валидность.

В отличие от очевидной валидности, содержательная валидность всегда связана с оценкой отношения содержания тестового материала к истинной, а не мнимой цели измерения. Содержательная валидность определяется на основе экспертных методов. Эксперты оценивают задачи теста по шкале близости к реальным задачам, встающим перед человеком в процессе обучения или работы, целям предполагаемого тестирования.

Первый этап экспертного оценивания может быть связан с отбором существенно важного материала, на основе которого должно строиться задания. На первом этапе подбираются эксперты в данной области. Окончательный вывод о содержательной валидности делается из обобщения суждений отдельных экспертов, которые они дают по реально составленным заданиям теста. В простейшем случае это может быть процент заданий, признанных группой экспертов соответствующими данному виду деятельности.

Валидность содержания, в сущности, означает систематическую проверку содержания теста с тем, чтобы установить, соответствует ли оно репрезентативной выборке исследуемой области знаний. Валидность содержания часто иллюстрирует примерами педагогических измерений уровня усвоения. Таким образом, валидность – это соответствие требований, которые предъявляется к измерению, содержание и сущности объекта измерения.

Как отмечает А. Анастази "... валидность по содержанию должна закладываться в тест при отборе соответствующих заданий". В этом случае содержание определяется учебным планом и методами обучения, а требования – совокупностью знаний, умений, навыков выпускников. Нарушение необходимого соответствия между требованиями (целями) и содержанием обучения (и контроля) приводит к недостоверности результатов измерения. Модель обеспечения содержательной валидности измерения можно представить следующим образом *рис. 1*:

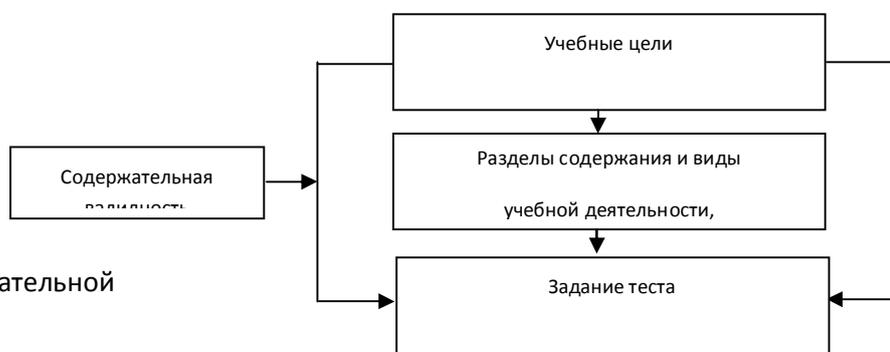


Рис 1. Модель содержательной валидности

Для повышения содержательной валидности теста необходимы [3,с.165]:

- подбор оптимальной трудности задания;
- экспертиза качества содержания;
- расчет оптимального времени выполнения задания;
- подбор валидных заданий с высокой дискриминативностью.

Дискриминативность обозначает различительную способность задачи, т.е. способность отделять испытуемых с высоким общим баллом по тесту от тех, кто получили низкий балл.

### 3. Конструктивная (концептуальная) валидность.

Этот вид валидности определяется в тех случаях, когда представление об измеряемом психическом феномене существует только в сознании исследователя. Разработчик теста может лишь строить гипотезу о существовании данного феномена, о его формах и характере проявления. Устанавливается концептуальная валидность путем доказательств правильности теоретических концепций, положенных в основу теста, это особенно необходимо в тех, случаях, когда результат тестовых измерений используется не просто для предсказания поведения, а как основа для выводов о том, в какой степени испытуемые обладают некоторой характеристикой.

4. Критериальная валидность (соответствия) – совпадение результатов измерения одного признака различными методами. Она помогает установить в какой мере результаты, полученные посредством использования исследовательского инструментария, соответствуют данным полученным другими способами. Для получения валидности соответствия на начальной стадии сравнивают полученные с его помощью результаты с данными эталонных способов (если эталон отсутствует, то валидность соответствия оценивается с уже существующими методами, которые удовлетворяют критериям качества).

Основным видом валидности является валидность по критерию или эмпирическая валидность. Суть ее заключается в определении способности теста служить индикатором или предсказателем строго определенной психической особенности, формы поведения человека и др.

Валидализация теста по критерию состоит в сравнении результатов, полученных

испытуемыми за решение теста с данными по критерию и вычислении коэффициента корреляции тестового результата с внешним критерием. Вопрос о выборе критерия является в этом виде валидности основным. От выбора критерия зависят качественная и количественная оценка валидности. Критерий валидности – это признак его обоснованности, правомерности его применения как измерительного инструмента.

При составлении и апробации тестов используются разнообразные критерии, классифицировать которые можно по следующим основаниям [6]:

- по объективности: субъективные и объективные;
- по метричности: качественные и количественные;
- по составу: одиночные и составные;
- по широте охвата явления: узкие и широкие.

Природа изучаемого явления определяет и широту используемых критериев. Более узкие критерии применяются при обосновании тестов на изменение уровня развития конкретных знаний, умений и навыков. Чем сложнее измеряемое явление, тем шире и обобщеннее должен быть критерий валидности.

#### 5. Сравнительная или конкурентная валидность.

Этот вид валидности рассчитывается по корреляции результатов выполнения данного теста с результатами других тестов, которые доказательно предназначены для измерения того же качества или свойства.

Для эффективного изучения конкурентной валидности существует несколько правил, представленным П. Клайном:

- 1) Убедитесь, что выборка испытуемых отражает те категорию лиц (популяцию), для которой данный тест предназначен, особенно по отношению к полу, возрасту, уровню образования и социальному положению.
- 2) Убедитесь, что выборки достаточно велики для получения статистически значимых корреляций, могущих быть затем использованными факторном анализе, минимальное количество испытуемых – 200.
- 3) Используйте настолько широкое разнообразие других факторов данной переменной, насколько возможно, чтобы убедиться, что корреляция получена благодаря близости групповых, а не специфических факторов.
- 4) Исследования конкурентной валидности, удовлетворяющие этим, критериям, должны дать недвусмысленное свидетельство валидности, которое не может быть опровергнуто.

6. Валидность прогноза позволяет вычислить взаимосвязь между данными, полученными в ходе проведенного исследования, и более поздними результатами той или иной деятельности. В прогностической валидности в качестве внешнего критерия выступает прогноз успеваемости учащегося, например, прогнозирование его успеваемости на основании результатов вступительных экзаменов [5].

Таким образом, анализ валидности демонстрирует всю сложность и тонкость конструирования тестов, а также их использования в учебном процессе.

#### Использованная литература:

1. Анастази А., Урбанина С. Психологическое тестирование.-7-ое изд. / -СПб: Питер, 2003.-688с.
2. Гулидов И.Н., Шатун А.Н. Методика конструирования тестов. -М.: ФОРУМ:ИНФРА-М. 2003.,112с.
3. Звонников В.И., Чельшкова М.Б. Современные средства оценивание результатов обучения: учеб.пособие. для студ. высш. учебн. завед. / -2-е изд., стер.-М.:/Академия, 2008.-224с.
4. Калдыбаев С.К. Педагогические измерения: Становление и развитие: - Б., 2008.-208с.
5. Талызина Н.Ф. Психодиагностика. Теория и практика. – М.: Прогресс, 1986.-297с.
6. Чельшкова М.Б. Теория и практика конструирования педагогических тестов: Учебное пособие. - М.: Логос, 2002., 432с.