

КУДРИНА Е.В.

г. Хабаровск

*Филиал сибирского государственного университета
телекоммуникаций и информатики в г. Хабаровске*

МЕТОДИЧЕСКИЕ КРИТЕРИИ КАЧЕСТВА ИЗМЕРЕНИЙ В ТЕСТОЛОГИИ

Тест как метод изучения индивидуальных особенностей человека возник сравнительно недавно. Временем его возникновения считается конец 19 – начало 20 веков, когда для определения физических, физиологических и психических особенностей человека психологи пытались применить различные способы измерения. В настоящее время тест как измерительный инструмент используется в большинстве стран мира.

Тестология как теория и практика тестирования за время своего существования накопила громадный опыт использования тестов в различных сферах человеческой деятельности, включая образование. Применительно к педагогике тест – это инструмент, состоящий из квалитетически выведенной системы тестовых заданий, стандартизированной процедуры проведения и заранее спроектированной технологии обработки и анализа результатов, предназначенных для измерения качеств и свойств личности, измерение которых возможно в процессе систематического обучения [1]. А само тестирование – это метод педагогической диагностики, с помощью которого выборка поведения, репрезентирующая предпосылки и результаты учебного процесса, должна максимально отвечать принципам сопоставимости, объективности, надежности и валидности измерений, должна пройти обработку и интерпретацию и быть готовой к использованию в педагогической деятельности [2].

Приведенное выше понятие «тестирования» было дано Ингенкампом и указывает на то, что важнейшими методическими критериями качества измерения в тестологии являются – объективность, надежность, валидность и сопоставимость. Кратко рассмотрим перечисленные критерии качества измерения.

Объективность применительно к методике измерения означает устранение воздействия субъективных факторов со стороны лиц, проводящих исследование. Другими словами измерение объективно, если результаты измерения не зависят от исследователя и различные исследователи при измерении одного и того же признака приходят к одному и тому же результату. Для достижения объективности:

1. результаты учебного процесса должны изучаться различными исследователями, проводящими тестирование в максимально унифицированных условиях;
2. разные лица, обрабатывающие полученные данные, при установленном тождестве фактического материала должны прийти к одинаковым результатам;
3. интерпретация результатов должна быть свободна от посторонних субъективных воздействий.

Под *надежностью* измерения понимается степень точности, с которой может быть измерен тот или иной конкретный признак. Проверка надежности измерения затрагивает, прежде всего, вопрос о том, насколько можно доверять результатам, полученным в ходе однократного измерения. Для определения надежности, как правило, проводится специальная апробация теста, которая выполняется после окончания анализа и составления заключительной формы теста.

Надежность характеризуется точностью теста как измерительного инструмента, устойчивостью его к действию помех (состоянию испытуемых, их отношения к процедуре тестирования, случайным ошибками и т.п.). Качественный тест не может быть создан без тщательного изучения этого важного аспекта измерения. Использование ненадежных тестов, допуск большого количества ошибок в таком ответственном деле, как тестирование людей, может стать причиной педагогических и административных ошибок, последствия которых трудно исправить.

Валидность метода необходима для того, чтобы определить, действительно ли измеряется то, что требуется измерить или что-то другое. Чтобы установить, действительно ли измерили то, что хотели, требуется критерий. В зависимости от вида критерия в методике тестирования различают четыре вида валидности:

1. валидность содержания – определяет насколько адекватно вопросы теста отражают смысл измеряемого явления и насколько полно они охватывают материал, подлежащий усвоению;
2. валидность соответствия – устанавливает, в какой мере результаты, полученные посредством тестирования, соответствуют данным, полученным другими измерительными способами;
3. валидность прогноза – позволяет установить взаимосвязь между данными, полученными в ходе ранее проведенного исследования (тестирования), и поведением, наблюдаемым позднее;
4. валидность конструктора – определяется в тех случаях, когда представление об измеряемом феномене-конструкторе (типичными конструкторами являются, например, умственное развитие,

интровертность, чувство страха перед каким-либо событием) существует только в сознании исследователя, и разработчик теста может лишь строить гипотезу о существовании данного конструктора, о его формах и характеристиках; валидность конструктора устанавливается путем доказательства правильности теоретических концепций, положенных в основу теста.

Применительно к тестам школьных достижений особое значение приобретает валидность содержания, валидность соответствия и валидность прогноза, в то время как валидность конструктора важна для тестов психологических методов.

Сопоставимость теста, возможна благодаря параллельным формам или благодаря тестам с одинаковой валидностью. В таком случае тест можно сравнивать с его параллельной формой. К тому же при групповых исследованиях параллельные формы затрудняют списывание.

Наряду с приведенными выше методическими критериями качества измерения в литературе называются и другие критерии. Линер [2], например, выделял следующие критерии:

1. *нормирование* – это процесс получения количественных и качественных характеристик, на основании сравнения с которыми можно дать оценочные суждения о результатах, показанных испытуемыми, что позволяет включить результаты индивидуального тестирования в систему соотносительных понятий;
2. *экономичность* – имеет место в том случае, если проведение теста и обработка данных не требует больших затрат времени и средств, если тест не сложен и может быть использован в группах;
3. *полезность* – имеет место и является максимальной, если существует большая практическая потребность в исследовании данного поведения и если для этой цели еще не разработано или разработано слишком мало исследовательский методов.

Даже если предположить, что тесты удовлетворяют требованиям всех критериев качества измерений, все равно тесты не будут являться универсальными измерительными средствами, т.к. границы использования тестирования достаточно хорошо известны. Однако качественно подготовленный тестовый инструментарий дает качественную и надежную информацию, соответствующую реальному положению дел. А само тестирование позволяет снизить количество апелляций более чем в три раза [1], сделать процедуру оценивания одинаковой для всех учащихся вне зависимости от места

проживания, типа и вида образовательного учреждения, в котором занимаются учащиеся.

Литература

1. **Майоров А.Н.** Теория и практика создания тестов для системы образования (Как выбирать, создавать и использовать тесты для целей образования). – М: «Интеллект-центр», 2004. – 296с.
2. **Ингенкамп К.** Педагогическая диагностика: Пер с нем. – М.: Педагогика, 1991. – 240с. – (Зарубежная школа и педагогика).