НАДЕЖНОСТЬ ТЕСТА

Автор: словарь
16.12.2008 20:50 -

Надежность теста — один из важнейших психометрических критериев качества теста, указывающий на степень устойчивости теста к искажающему воздействию случайных шумовых факторов (англоязычный термин — reliability). Н. т. связана с определением точности. Чем выше Н. т., тем свободнее он от погрешностей измерения, тем в меньшей степени тестовый балл отклоняется от истинного значения измеряемого психического свойства (фактора). Уменьшению погрешностей способствует стандартизация условий проведения тестирования: всем испытуемым даются одинаковые инструкция, система заданий, время, одинаково приписываются баллы за одинаковые результаты и т. п. Операциональное понятие Н. т. основано на идее арифметического расщепления общей дисперсии тестового балла испытуемых на отдельные компоненты измерения — истинные и ошибочные. Соответственно, за Н. т. принимается мера отношения дисперсии истинных компонентов измерения к общей дисперсии, включающей в себя и истинную, и ошибочные части. Поскольку всякое измерение всегда имеет некоторую степень погрешности и значения любого компонента дисперсии всегда больше нуля, то получаемая мера Н. т. всегда меньше единицы и больше нуля. Именно в этих пределах и определяется эмпирическое значение надежности тестовых результатов.

Следовательно, сам факт проверки тестовых результатов на надежность является первым необходимым (но еще недостаточным) признаком научной обоснованности каждого теста. При определении Н. т. обычно известно лишь эмпирическое значение общей дисперсии тестовых баллов испытуемых, получаемой как сумма квадратов отклонений баллов испытуемых от среднего арифметического балла (SS), деленная на число испытуемых. Наибольшую трудность при определении Н. т. представляет выявление значения дисперсии ошибочных компонентов измерения. Из-за влияния множества трудноучитываемых и неучитываемых факторов, вызывающих различную погрешность данных, непосредственное определение точного значения ошибочных компонентов измерения становится невозможным.

Однако в практике выработано немало способов примерного оценивания дисперсии ошибочных компонентов. Каждый из них основан на тех или иных предположениях относительно источника возникновения ошибок измерения.

Например, если имеются две параллельные формы одного и того же теста, то в качестве основного источника погрешностей рассматриваются различия в этих формах. Коррелирование результатов тестирования испытуемых по одному варианту теста и результатов тех же испытуемых по другому варианту дает значение коэффициента Н. т. в смысле эквивалентности вариантов одного и того же теста. Если же испытуемые дважды тестируются по одному и тому же тесту (или варианту) через некоторый промежуток времени, то коррелирование получаемых при этом результатов дает представление о Н. т. в смысле устойчивости данных по отношению к фактору времени.

Определение Н. т. может быть связано и с т.н. понятием внутренней однородности (согласованности) теста, что проверяется расчленением теста на части с последующим коррелированием результатов частей или суммированием дисперсий этих же частей. Н. т. определяется также методами дисперсионного и факторного анализов. В последние годы, в дополнение к методам т.н. классической теории тестов, для определения погрешности измерения используется вычисление значения информационной функции теста, если тест разрабатывается на основе современной математической теории тестов, называемой по-английски Item Response Theory.

В.С. Аванесов, А.Г. Шмелев