

СТАНДАРТНИ ИЗИСКВАНИЯ ЗА ОПИСАНИЕ, ИЗПОЛЗВАНЕ И ПРЕДСТАВЯНЕ НА ТЕСТОВЕ

СЪДЪРЖАНИЕ НА “ПАСПОРТА” НА ТЕСТА

1. Име на теста (година на издаване/разработка)
2. Автор
3. Издател; източник от който може да бъде доставен; права върху теста
4. Категория на теста (А, В, С). Условия на достъп и използване (изисквана квалификация на потребителя)
5. Цена
6. Описание на теста. Предназначение на теста: възрастов диапазон; какво се оценява; брой и наименование на скалите и брой айтеми (обща скала, интегрираща частните показатели, подскали, скали за валидизиране на индивидуалния резултат; формиране на скалите – чрез независими или припокриващи се айтеми).
7. Процедура на конструиране
8. Надеждност:
 - а. тест-ретест надеждност – обем на извадката, времеви интервал;
 - б. съгласуваност (надеждност) на взаимозаменяеми форми;
 - в. съгласуваност в оценките на независимите експертите;
 - г. вътрешна съгласуваност между айтемите.
9. Конструкт валидност:
 - а. съдържателна;
 - б. критериална;
 - в. конвергентна и дискриминантна.
10. Стандартизационна извадка (обем, селекция, възрастови и други характеристики; данни от клинична извадка). Норми:
 - а. в стандартни скали;
 - б. в персентилни скали.Видове норми: по възраст, по пол, по други признаци.
11. Процедура за тестиране и оценяване на резултата. Време за изследване. Форма на провеждане (индивидуална или групова). Формат на скалата за отговори, изчисляване на суровия бал, трансформация в стандартна скала/персентилно разпределение. Наличие/отсъствие на ръководство към теста.

Наличие/отсъствие на компютъризиран вариант на провеждане и/или оценка на резултатите. Характер на заключението – диагностична категория, емпирично формиране променлива, класификация на изследваното лице, описание, друго.

12. Отзиви/рецензии на инструмента от независими експерти (публикувани или непубликувани мнения).

КАТЕГОРИЯ НА ТЕСТА – А, В, С – И КОМПЕТЕНТНОСТИ, НУЖНИ ЗА ПРОФЕСИОНАЛНА РАБОТА С ПСИХОЛОГИЧЕСКИТЕ ТЕСТОВЕ

Равнище А	Равнище В	Равнище С
Не се изисква специална квалификация.	<p>В1. Лица с бакалавърска или магистърска степен по психология + дисциплини по статистически анализ, психометрия, психологическа диагностика и оценяване и сродни курсове със задължителен минимален хорариум 400 часа*;</p> <p>В2. Лица с бакалавърска или магистърска степен по психология, без необходимия хорариум, но преминали тренинг или подготовка за използване на конкретен инструмент от категория “В”**.</p>	Равнище В1 + специализирана подготовка за работа с конкретен тест от категория “С”**.

* Дисциплини по статистически анализ, психометрия, психологическа диагностика и оценяване и сродни курсове със задължителен минимален хорариум 400 часа:

1. Методология на психологическото изследване;
2. Статистика;
3. Психометрия;
4. Конструирание на тестове;
5. Психологическа диагностика и оценяване;
6. Други сходни дисциплини, включващи конструирание и използване на психологически тестове.

Подобно на потребителите, психологическите тестове се категоризират в същите категории – “А”, “В” и “С”. Като примери за тестове от:

- *категория “С”* могат да се посочат индивидуално провеждани тестове за интелигентност (от типа на Уекслър), многомерни личностни скали (от типа на Минесотския многопрофилен въпросник) и проективни тестове (от типа на Роршах). Интерпретацията на тестовете от посочения тип, предназначени за диференциална диагноза и взимане на решение относно изследваното лице, изискват задълбочени познания по психометрия, значителен практически опит и индивидуален тренинг за прилагането на конкретния инструмент, с акцент върху анализа не на формалните тестови показатели, а върху конфигурацията на анализирания съвкупност от параметри;
- *към категория “В”* могат да се отнесат личностни скали за оценка на интроверсия–екстраверсия (на Х. Айзенк), тревожност – състояние и черта (Ч. Спилбъргър) и др. Тестовете от тази категория обикновено са предназначени за оценка на едномерни или многомерни конструкти, с достатъчно уточнена психологична природа, като данните се интерпретират директно, въз основа на стандартния тестови бал. За компетентната оценка на резултатите се изисква задълбочена подготовка по психометрия, детайлно познаване на ръководството към теста и изследвания конструкт, без необходимост от специален тренинг за използване на конкретния инструмент;
- *категория “А”* включва инструменти като скала на Рокич за ценностите, скала на Мос за работната среда, въпросник планиране на кариерата – ръководство за самопомощ, скали за социална подкрепа и др. При ясно дефинирани инструкции, провеждането на изследването с помощта на посочената категория тестове не изисква специализирани психологически познания, но е препоръчително интерпретацията на резултатите да се извършва от лица с базова подготовка по психометрия и знания за оценявания конструкт.

Видове тестове:

1. За оценка на интелигентност, способности, когнитивни процеси;
2. Личностни скали;
3. Скали за нагласи, интереси, мотивация;
4. Проективни тестове;
5. Невропсихологични;
6. За професионален подбор.

НАДЕЖДНОСТ И ВАЛИДНОСТ НА РЕЗУЛТАТИТЕ ОТ ТЕСТА

След операционализацията на измервания конструкт изследователят *разработва собствен метод* за оценка и представяне параметрите на теста или *избира от наличните инструменти*.

Емпиричните индикатори (тестовите резултати) трябва в максимална степен да се свободни от грешки на измерването, т. е. необходимо е да бъдат едновременно *надеждни* и *валидни*. От гледна точка на двата възможни типа грешки на измерването – случайни и систематични, *надеждността* се отнася само

до случайните (несистематични, некорелирани с оценявания конструкт) грешки, докато *валидността* е повлияна както от случайни, така и систематични грешки на измерване.

Надеждността и *валидността* се отнасят до *результатите*, получени с помощта на даден инструмент при определени условия и с определени групи лица. Затова извадката, с която е работено, трябва да бъде описана и степента ѝ на представителност за генералната съвкупност статистически оценена.

Надеждност

Тестовият бал е надежден, ако е повлиян в минимална степен от случайни грешки на измерването. В зависимост от характера на теста и съобразно различни типове случайни грешки, се използват различни аспекти на измерване на надеждността (както и термини за тяхното означаване), обединени от по-общото понятие *съгласуваност*.

Съгласуваността на измерването следва да бъде оценена и представена статистически в “паспорта” в някои от следните категории съобразно спецификата на теста:

а. съгласуваност във времето – повторемост на резултатите при двукратно измерване: *тест-ретест надеждност*;

б. съгласуваност на две взаимно заменяеми (паралелни) форми на теста: *надеждност на взаимозаменяеми форми*;

в. съгласуваност в оценките на независимите експерти: *експертна надеждност*;

г. съгласуваност на теста със самия себе си – *на двете половини на теста и вътрешна съгласуваност между айтемите*.

Валидност

В широк смисъл валидността се отнася до това как резултатът трябва да се интерпретира – т. е. дали тестът измерва това, за което претендира и е предназначаен. Различните тестове допускат различни оценки за валидност: *съдържателна; критериална; конвергентна и дискриминантна; конструкт-валидност*.

Конструкт-валидността е най-общият тип валидност и включва останалите видове¹. Този термин се отнася най-общо до въпроса дали тестът наистина измерва конструкта, който изследователят предполага. Съдържателната,

¹ Традиционно се различават три типа валидност: съдържателна, критериална и конструкт-валидност (към която се отнасят конвергентната и дискриминантната валидност). От съвременна гледна точка е налице тенденция за по-широко тълкуване на конструкт-валидността, като термин включващ останалите видове.

критериалната, конвергентна и дискриминантна валидност включват по-частни аспекти на този въпрос и методи за неговата оценка

а. *Съдържателната валидност* следва да отговаря на въпроса доколко съвкупността от отделните емпирични индикатори (айтеми, задачи и др.) са репрезентативни за измерваната психологическа характеристика (черта, способност и т.н.), т. е. доколко избраните емпирични показатели представят цялостно полето на изследвания конструкт. Например математическите умения във втори клас, фактологическите знания, които кандидат-шофьорите трябва да притежават за правилата на уличното движение. Съдържанието и на двете посочени области може да бъде дефинирано с относително висока степен на съгласие от подходящи експерти (учители, специалисти по пътно движение).

Съдържанието на други конструкти (например “интелигентност”), което не може толкова ясно да бъде определено, налага предварително формулиране на *операционални дефиниции*, които да служат не само като ръководство за подбора на емпиричните индикатори, но да позволяват и на други специалисти да оценят доколко съдържанието на индикаторите е съгласувано с дефиницията, дадена от изследвателя.

б. *Критериалната валидност* се определя, когато е налице външен/независим стандарт на измерваната с теста психична характеристика, т. е. когато тестовият резултат (предиктор) е свързан с външен (независим) критерий за оценка на същия конструкт. Например валидността на входният изпит за постъпване в университета (тестов предиктор) да бъде оценена чрез успеха по време на следването (външен критерий). За проверка на критериалната валидност в повечето случаи, но не винаги, могат да се използват корелационни коефициенти.

От гледна точка на времето за оценка на критерия се различават и могат да се използват *текуща валидност* (когато оценката на предиктора и критерия се извършват по едно и също време), *прогностична (предиктивна) валидност* (когато критерият се измерва по-късно). В някои случаи можа да се използва и *постдиктивна валидност* (когато критерият се измерва преди предиктора).

в. *Конвергентната и дискриминантната валидност* се отнасят до релацията на тестовия бал с резултати от тестови оценки на други психологически конструкти.

Значимите връзки с аналогични или сходни конструкти, измерени с други тестове или инструменти, с които презентираният тест теоретично трябва да корелира, се означава като *конвергентна валидност*.

Докато въобще отсъстващите или относително по-слабите/незначими връзки с различни по-същество конструкти (с които тестът теоретично не трябва да корелира) се означава като *дискриминантна валидност* (използва се и терминът *дивергентна валидност*).

Тъй като силната корелация между два тестови резултата не означава с необходимост, че те измерват същия или аналогичен конструкт, препоръчително е да се намерят множество индикатори за оценка валидността на един и същ

конструкт, базирани на различни методи за оценка (самоописание, рейтингови скали, наблюдение, проективни методи и други).

Обратно, ниската надеждност на тестовите резултати може да доведе до слаба корелация, която погрешно да се интерпретира като потвърждаване на дискриминантната валидност

Като конструкт-валидност се обозначават всички данни в подкрепа на намерението на автора на теста той да измерва точно определени психологически феномени. Доказателствата за конструктната валидност се базират на факта, че научните понятия (и означаваните с тях феномени) са свързани по между си по точно определен начин в така наречените номотетични мрежи. Затова конструктната валидност на един тест се определя от степента, в която е установено, че измереният от теста конструкт се намира в същите отношения спрямо други релевантни конструкти, в които би трябвало да се намира съгласно неговата номотетична мрежа.

За оценка на конструкт-валидността могат да се използват методите на експлораторен и конфирматорен факторен анализ, структурно моделиране, регресионен и корелационен анализ, дисперсионен анализ и други, включително експериментални манипулации за оценка дали тестовите резултати ще се проявят в очакваната посока.

Само един отделен метод не може да осигури потвърждаването на валидността на тестовия резултат. Необходими са оценки на различни, достъпни за конкретния случай аспекти на конструкт-валидността.

В “паспорта” на теста следва да има приведени статистически и други доказателства в подкрепа на основните изводи, заради които се препоръчва използването му. Степента на генерализация на изводите в интерпретативната част на ръководството за неговото използване трябва да съответства на генерализируемостта на резултатите от проведените емпирични проверки на валидността.

НОРМИ

Индивидуалният тестов бал добива смисъл, когато се отнесе към тестовия бал на други изследвани лица, към баловете на определена група хора или на цяла популация. Тогава тестираният може да бъде разположен между останалите и оценен.

За целта следва да се познават баловете на другите или като се измерят за цялата съвкупност, или като за целта се използва *представителна случайна извадка*. На фона на тази картина всяко изследвано лице може да се ситуира между останалите – в норма, под нормата, или над нормата.

Авторите на теста следва да посочат степента на представителност на нормиращата извадка, начина на нейното сформирание (проста, стратифицирана, гнездова и пр.), обема ѝ, генералната съвкупност, за която е репрезентативна.

Съществуват различни методи за представяне *норми* на тестовите резултати, които в различна степен отразяват теоретични основания, потребности на ползвателите, използвани статистически показатели, предпочитания на изследователите.

Могат да се посочат различни разновидности норми: по пол, възраст и други съществени за изследвания конструкт признаци.

а. Основният начин за отнасяне резултатите на едно изследвано лице към резултатите на индивидите в стандартизацията (нормативната) извадка, а следователно и в генералната съвкупност, при условие, че разпределението е нормално или близко до нормалното, са *стандартните балове*. Те се изчисляват като отношение: разликата на индивидуалния бал със средната величина за извадката, разделена на стандартното отклонение. Получените стойности са интервално скалирани в единици стандартно отклонение.

На практика, за да осигури на потребителите на теста възможност за преминаване от суров към стандартен бал, неговият автор следва да представи оценки за *средната стойност* на изследваната характеристика в нормативната извадка и *стандартното отклонение*. За норма обикновено се използва интервалът около средната стойност на разстояние плюс-минус едно стандартно отклонение.

Горната процедура изисква представянето и на данни за *степенна на нормалност на разпределението*.

б. Друга форма, чрез която се придава значение на индивидуалния бал, е превръщането му в *персентили*, т. е. пресмятането какъв процент от индивидите в нормиращата извадка, имат по-нисък бал от изследваното лице. Резултатите от подобно измерване са рангово скалирани.

На практика, на базата на нормативната извадка се изчисляват и представят *кореспондиращите балове* – например на 3^{-ти}, 10^{-ти}, 25^{-ти}, 75^{-ти}, 90^{-ти} и 97^{-ми} персентил, които определят съответни зони.

ЗАКЛЮЧЕНИЕ

Представените “Стандартни изисквания за описание и представяне на тестове” дават необходимата насока и ориентири при публикуването на даден тест и използването му с диагностична или изследователска цел. Те обаче нито са изчерпателни, нито са задължителни в целия си обем.

Някои от параметрите могат да не са приложими към конкретен тест, а в други случаи авторът да ползва свои оригинални методи за верификация и доказване правилността на своите заключения относно изработения инструмент.

Възможно е също тестът да е на по-ранно или по-късно ниво на своята разработка и да не предлага всички възможни и желателни проверки и гаранции.

Във всички случаи обаче академичната и професионалната етика на психолога изисква да бъдат посочени методите, оценките, статистическите показатели, които в актуалния момент представят адекватно състоянието на теста.