

Лекція №3 з дисципліни «психодіагностика».

ПОНЯТТЯ ТЕСТУ. ШКАЛЮВАННЯ. РЕПРЕЗЕНТАТИВНІСТЬ. ЕТАПИ СТВОРЕННЯ ТЕСТУ

Автор-розробник: к.психол.н, доц. каф. прикладної психології О.Л.Луценко

Психологічний тест - стандартизоване, часто обмежене в часі випробування, призначене для встановлення кількісних (і якісних) індивідуально-психологічних особливостей через вербальні або невербальні зразки відповідей або інші види поведінки.

Тест складається із завдань – окремих пунктів методики, які призначені в сукупності для моделювання діяльності, в рамках якої, згідно теоретичним міркуванням або емпіричним спостереженням психолога, проявляється та або інша психологічна особливість індивідуума: його риси, темпераментальні властивості, настанови, мотиви, соціально-психологічні якості.

Психологічний тест може бути охарактеризований як ефективний, якщо він відповідає наступним вимогам: будові на шкалі інтервалів; обґрунтованості конструкту тесту; репрезентативності; дискримінативності, надійності, валідності окремих тестових завдань та тесту загалом; стандартизованості пред'явлення й обробки результатів; незалежності результатів від впливу експериментальної ситуації й особистості психодіагноста; наявності нормативних даних, а у випадку індивідуально-орієнтованих тестів - наявності критеріїв, установлених експертами; адаптованості тестів.

Дані психометричні вимоги до тестів або методичні принципи психодіагностики є «професійною мовою» психодіагноста, яку потрібно знати і розуміти бездоганно. Розглянемо їх по черзі.

Шкалювання. Види вимірювальних шкал

Шкала - форма фіксації сукупності ознак досліджуваного об'єкта з упорядкуванням їх у певну числову систему. Просто кажучи, шкала – це мірка, без якої вимірювання неможливе.

Наприклад, Цельсій для шкалювання температури взяв дві точки відліку - точку замерзання води й точку кипіння води. Розділив цей інтервал на 100 рівних відрізків і, таким чином, увів температурну шкалу, якою ми користуємося, говорячи, що на вулиці -5 градусів морозу або 25 градусів тепла. Фаренгейт використав інші точки відліку, тим самим створив іншу шкалу. Тому градус за шкалою Фаренгейта не дорівнює градусу за шкалою Цельсія.

С.Стівенс запропонував класифікацію з 4-х типів шкал вимірювання, ієрархічно впорядкованих за точністю. Тип шкали, що закладається до тесту, буде визначати всі подальші процедури з конструювання і перевірки тесту, а також методи обробки і аналізу даних, які будуть отримані за допомогою цього тесту. Ці шкали носять назви: номінативних, порядкових, інтервальних і шкал відношень.

1. Номінативні шкали (шкали найменувань) – класифікація, що встановлює відповідність ознаки тому або іншому класу. Найпростіший випадок номінативної шкали - дихотомічна шкала, що складається всього лише із двох пунктів: «має братів і сестер - єдина дитина в родині»; «іноземець - співвітчизник»; «проголосував «за» - проголосував «проти»» і т.п. Ознака, що вимірюється за дихотомічною шкалою найменувань, може приймати всього два значення. Тоді можна говорити про те, що та або інша ознака «виявилась» або «не виявилась». Наприклад «Ознака наявності досвіду праці виявилась у 8 досліджуваних з 20. У 12 досліджуваних не виявилось досвіду праці».

Складніший варіант номінативної шкали - класифікація із трьох і більше пунктів: «екстрапунитивні - інтрапунитивні - імпунитивні реакції»; «вибір кандидатури А - кандидатури Б - кандидатури В - кандидатури Г» і т.п. Допустимо, що кандидатуру А вибрали 7 досліджуваних, кандидатуру Б - 11, кандидатуру В - 28, а кандидатуру Г - усього 1. Тепер можна оперувати цими числами, що представляють собою частоти зустрічальності різних найменувань, тобто частоти прийняття знаком «вибір» кожного з 4 можливих значень.

Для обробки даних, що вимірюються за номінативними шкалами, вживають методи χ^2 , кутового перетворення Фішера ϕ^* , кореляції якісних ознак – тетрагорічний і полігорічний коефіцієнти зв'язку, дисперсійний аналіз Фішера.

Порядкові (рангові, ординальні) шкали – тут елементи шкали впорядковані (ранжирувані) за принципом «більше - менше», але в цій шкалі не враховуються відстані між ними. У шкалі найменувань не має значення, в якому порядку розташовані класифікаційні категорії, а в шкалі порядків вони утворюють послідовність від пункту «найменше значення» до пункту «найбільше значення» (або навпаки). У порядковій шкалі повинно бути не менше трьох рангів, наприклад в методиці Ч.Осгуда «Семантичний диференціал» 7 рангів: -3, -2, -1, 0, 1, 2, 3. Є багато опитувальників з варіантами відповідей, які утворюють порядкову шкалу, наприклад: «Буває, що я ніяк не можу прийняти якесь остаточне рішення»...а) Повністю згоден. б) Можу погодитися. в) Не впевнений. г) Скоріше не згоден. д) Зовсім не згодний.

Для обробки даних, які вимірюються ранговими шкалами, використовується непараметрична статистика – критерії Манна-Уїтні, Уїлкоксона, Колмогорова-Смірнова, Крускала-Уоллеса, ранговий кореляційний аналіз Спірмена та інші.

Шкали інтервалів (рівних інтервалів, кількісні). Вони класифікують за принципом «більше на певну кількість одиниць – менше на певну кількість одиниць». В них різниці між значеннями у всіх точках даної шкали рівні. Рівноінтервальними можна вважати лише шкали в одиницях стандартного відхилення шкали, і те лише за умови, що розподіл значень у вибірці стандартизації був нормальним. Принцип конструювання більшості інтервальних шкал використовує так зване правило «трьох сигм»: це емпірична закономірність, яка полягає в тому, що приблизно 97,7-99,7%

всіх значень ознаки при нормальному його розподілі укладаються в діапазоні $M \pm 3\sigma$, де M – середнє арифметичне, σ - середнє квадратичне відхилення. Можна побудувати шкалу в одиницях часток стандартного відхилення (наприклад, інтервал може складати $\frac{1}{2} \sigma$, як це було зроблено Р.Кеттеллом для стандартизації тесту «16 Особистісних факторів»), яка буде охоплювати весь можливий діапазон змін ознаки, якщо крайній ліворуч і крайній праворуч інтервали залишити відкритими. Еквівалентні інтервальні шкали можуть бути лінійно перетворені друг у друга, що дозволяє приводити результати тестування до загальних шкал і в такий спосіб здійснювати порівняння показників.

Коли показники виміряні за шкалою інтервалів для їх обробки підходять більшість процедур параметричної статистики – коефіцієнт лінійної кореляції Пірсона, t - критерій Ст'юдента тощо. До розряду шкал інтервалів відносяться шкали IQ-показника стандартного, Т-балів, стенив, стенайнів, с-шкала, Z-показники.

Шкала процентілей Р (окремий вид стандартних тестових показників, що відображають, скільки відсотків вибірки стандартизації мали тестові показники, які були нижче балів даного досліджуваного) ще відноситься до порядкових шкал і не потрапляє до рівноінтервальних.

Шкали відносин (рівних відносин, пропорційні). Шкали відносин крім усього, чим характеризуються інтервальні шкали, мають значиму нульову точку. Тобто на відміну від інтервальної шкали, де нульова точка довільна (умовна), в шкалі відносин вона справжня, і вказує на повну відсутність вимірюваної властивості. Наявність такого справжнього «нуля» - проблема для більшості психологічних змінних. Можливості людської психіки настільки великі, що важко уявити абсолютний нуль у будь-якої психологічної змінній (абсолютна дурість або абсолютна чесність - поняття скоріше життєвої психології). Хоча існують методи конструювання тестів, які дозволяють використати цю можливість. Вважається, що прикладами шкал рівних відносин у психології є шкали порогів абсолютної чутливості аналізаторів. До шкал відносин наближуються тести, сконструйовані також за моделлю Раша (теорією завдання-відповідь або item-response theory - IRT).

До шкал відношень можуть застосовуватися всі найбільш точні й потужні статистичні методи обробки даних.

В ідеалі розроблювач тестів повинен прагнути використовувати шкали відносин. Якщо це неможливо, бажано будувати шкали інтервалів. Проте, як можна бачити, дуже багато поширених у СНД тестів і методик будуються на ранговій шкалі.

Окрім поняття шкали, для коректного здійснення процедури вимірювання необхідно ознайомитися ще з деякими психометричними термінами – поняттями істинного показника і репрезентативності.

Репрезентативність завдань тесту. Істинний показник. Похибка вимірювання

Репрезентативність - властивість вибіркової сукупності представляти характеристики генеральної сукупності. Генеральна сукупність – це всі існуючі об'єкти (ситуації) даного типу. Вибіркова сукупність – частина генеральної сукупності, з якою працює дослідник, тому що зазвичай генеральна сукупність недоступна для вивчення і навіть не є необхідною за умов репрезентативності і достатньої кількості об'єктів у вибірковій сукупності. Тобто дослідження генеральної сукупності є надлишковим і зайвим.

З певною погрішністю (похибкою вимірювання) можна вважати, що представлений у вибірковій сукупності розподіл досліджуваних ознак психологічної змінної відповідає їхньому реальному розподілу.

Для будь-якої риси (наприклад, домінантності, доброзичливості, креативності) індивідуум має істинний, справжній показник. Його можна було б обчислити, якби було можливо пред'явити досліджуваному всі можливі завдання, в яких діагностується ця риса. Будь-який тестовий показник для будь-якого індивідуума в кожному окремому випадку відрізняється від його істинного показника на величину випадкової погрішності, що з'являється в результаті виконання досліджуваним не всіх можливих завдань, а тільки вибірових.

Тест буде працювати помилково, коли вибірка його завдань не відбиває задовільно генеральну сукупність завдань (більшість життєвих ситуацій, в яких проявляється вимірювана психологічна риса), тому що тоді похибка вимірювання буде надмірно великою, такою, яку не можна ігнорувати, адже вона буде серйозно відхиляти отриманий розподіл ознаки від його реального стану. В цьому полягає важливість утворення репрезентативної сукупності завдань тесту.

Для створення репрезентативності завдань тесту роблять так звану специфікацію тесту – розраховують, скільки завдань різних типів необхідно включити до тесту.

Похибка вимірювання виникає ще через випадкові коливання вимірюваної психологічної якості під впливом настрою, стану, обставин. Навіть, якщо той самий досліджуваний декілька днів поспіль буде виконувати той самий тест, його результати кожного разу будуть дещо відрізнятися. Усереднення цих відхилень дозволить обчислити похибку вимірювання тесту.

Репрезентативність вибірки досліджуваних осіб

Як і вибірка завдань, вибірка досліджуваних осіб повинна адекватно відбивати генеральну сукупність людей, до якої буде застосовуватися даний тест, тобто на кого його результати можна буде екстраполювати. Якщо тест буде застосовуватися для школярів, то у вибірці повинні бути у відповідній пропорції представлені школярі приватних і державних шкіл, шкіл математичного й гуманітарного профілів, сільських і міських шкіл і т.п.

Дотримання репрезентативності вибірки досліджуваних у психодіагностиці важливо під час визначення норм тесту, формування вибірок критеріальних і контрольних груп при конструюванні тесту за критеріально-ключевим принципом, під час апробації тесту, перевірки надійності й валідності тесту й інших психометричних процедурах.

Етапи створення тесту

Для конструювання ефективного тесту необхідно виконати наступні кроки:

- 1) розглянути й установити строгі обмеження на зміст завдань у відповідності з теоретичною концепцією досліджуваної психологічної властивості;
- 2) сформулювати так багато різноманітних завдань, релевантних цьому змісту, наскільки це можливо;
- 3) випробувати ці завдання на такій великій вибірці досліджуваних, наскільки можливо;
- 4) виконати аналіз завдань тесту, відбираючи ефективні. (Аналіз завдань - процедура для оцінки ефективності завдань у тесті; вона може включати визначення кореляції завдань з загальним тестовим показником, перевірку складності/дискримінативності та інші способи.);
- 5) крос-валідизувати завдання на новій вибірці. Треба перевірити, щоб завдання адекватно охоплювали відповідний контекст;
- 6) обчислити показники дискримінативності та надійності тесту;
- 7) валідизувати тест;
- 8) виконати факторний аналіз завдань і порівняти його із процедурою аналізу й відбору завдань;
- 9) визначити норми для цього тесту;
- 10) якщо тест виявляється ефективним і корисним, представити результати в посібнику з тесту.

Рекомендована та використана література:

- Анастаси А., Урбина С. Психологическое тестирование. - СПб.: Питер, 2003.
- Бурлачук Л.Ф. Психодиагностика: Учебник для вузов. - СПб.: Питер, 2006.
- Бурлачук Л.Ф., Морозов С.М. Словарь-справочник по психодиагностике. - СПб.: Питер, 1999.
- Клайн П. Справочное руководство по конструированию тестов. – Киев: ПАН Лтд., 1994.
- Общая психодиагностика / Под ред. А.А. Бодалева, В.В. Столина. СПб.: Изд-во «Речь», 2004.
- Практикум по общей, экспериментальной и прикладной психологии / В.Д.Балин, В.К. Гайда, В.К. Гербаческий и др. Под общей ред. А.А. Крылова, С.А. Маничева. – СПб.: Питер, 2006.
- Психологическая диагностика: Учебник для вузов / Под ред. М.К. Акимовой, К.М.Гуревича. – СПб.: Питер, 2008.

Сидоренко Е.В. Методы математической обработки в психологии. - Спб., 1996.