

Лекція №4 з дисципліни «психодіагностика».

КЛАСИФІКАЦІЯ ТЕСТІВ. КОНСТРУКТ ТЕСТУ. МЕТОДИ КОНСТРУЮВАННЯ ТЕСТІВ

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

Класифікація тестів

Для впорядкування, структурування та виявлення певних закономірностей серед великої кількості розроблених за всю історію психодіагностики тестів дослідники намагаються запропонувати різні способи класифікації тестів.

За процедурою проведення тести розділяють на індивідуальні й групові.

За конструкцією вони розпадаються на гомогенні й гетерогенні (багатофакторні).

За обладнанням, яке використовуються для тестування, їх розподіляють на бланкові (олівець - папір), предметні (картки, кубики, фігурки), апаратні (для діагностики уваги, темпераменту, психічної працездатності, сенсо-моторної координації, психічних станів тощо), комп'ютерні. Для цього використовують прилади контактної координаціометрії, критичної частоти злиття світлових мерехтінь, реакції вибору, стеження за рухомим об'єктом тощо, програмне забезпечення і комп'ютерну техніку.

За ступенем впливу експериментатора серед тестів виділяють: дуже залежні від цього впливу – діалогічні/клінічні методи (спостереження, бесіда, інтерв'ю), проєктивні методики, ситуативні тести; і слабо залежні – стандартизовані особистісні опитувальники, тести інтелекту і здібностей, об'єктивні тести особистості, психофізіологічні методики.

За психодіагностичним методом всі тести поділяють на: 1) об'єктивні (діагностика здійснюється на основі успішності/результативності й способу/особливостей виконання діяльності); 2) суб'єктивні (діагностика здійснюється на основі самоопису досліджуваного особливостей власного характеру, відчуттів, реакцій, поведінки в тих або інших ситуаціях) і 3) проєктивні (діагностика здійснюється на основі взаємодії досліджуваного з зовні нейтральним стимульним матеріалом, що стає в силу його невизначеності об'єктом проєкції).

За принципом оцінки виділяють номотетичні (оцінка основана на порівнянні результатів досліджуваного з нормативними даними – усередненими даними за репрезентативною вибіркою людей подібної статі, віку, статусу), ідеографічні (оцінка основана на аналізі індивідуальних особливостей кожного конкретного досліджуваного), критеріально-орієнтовані психодіагностичні методики (оцінка основана на критеріях, обґрунтованих експертами).

За досліджуваними змінними тести розділяють на дві великі групи - інтелекту й особистості. В свою чергу, інтелектуальні тести утворюють гілки тестів інтелекту в цілому, тести загальних і спеціальних здібностей, тести креативності і тести досягнень. Особистісні тести теж утворюють підгрупи особистісних опитувальників, проєктивних тестів, тестів мотивів, інтересів,

цінностей, шкал установок, психосемантичних методик і об'єктивних тестів. Далі певні підгрупи теж можуть утворювати власні ієрархії, наприклад, об'єктивні тести особистості діляться на об'єктивні тести дії і ситуативні тести, а проєктивні тести утворюють групи структурування, доповнення, конструювання, інтерпретації, катарсису, вивчення експресії і продуктів творчості.

Будь-яка класифікація є не абсолютною, а умовною формою організації інформації, тому що завжди є певні «змішані» категорії. Наприклад, серед психодіагностичних методик тести креативності є такою змішаною категорією, тому що по суті відносяться і до інтелектуальних і до особистісних тестів, так як містять особистісні (ризикованість, уяву, допитливість, складність) і когнітивні фактори (бігкість, гнучкість, оригінальність, розробленість), або певні медико-психологічні проби, наприклад, метод піктограм, інтерпретації прислів'їв або доповнення речень, які нагадують проєктивні тести особистості, але спрямовані на діагностику когнітивних порушень і т.п.

Конструкт тесту (концептуальна змінна) – змінна, існування і природа якої впливає з теорій, наукових уявлень, а не з прямих спостережень. Більшість психологічних рис і якостей, які психологи зазвичай згадують, коли описують психологію людини, є недоступними для безпосереднього спостереження, наприклад, інтелект, агресивність, тривожність або екстраверсія. Ми спостерігаємо не інтелект, а те, що одна людина вчиться, вирішує завдання, знаходить вихід із складних ситуацій краще ніж інша; ми спостерігаємо не агресивність, а те, що певна людина ображає, б'є інших людей, розповсюджує плітки про них і т.п. Тести вимірюють ці концептуальні гіпотетичні конструкти через завдання, в яких йдеться про об'єктивні спостережувані ознаки поведінки. Теоретично припускається, що група корелюючих ознак поведінки не випадково ведуть себе узгоджено – має бути прихована (латентна) неспостережувана причина, над-змінна, яка об'єднує всі ці більш дрібні ознаки, «примушує» їх діяти узгоджено.

Розробка будь-якого тесту починається з обґрунтування теоретичного конструкту - психологічної змінної, котру треба вивчити. Таке обґрунтування повинне включати різнобічний опис змінної: опис її передбачуваних взаємозв'язків з іншими властивостями особистості й гіпотези про те, у яких завданнях (вербальних або невербальних) ця змінна може виявити себе. У якості джерел для обґрунтування теоретичного конструкта використовують літературу з досліджуваної проблеми, словники, думки експертів. Ретельна ідентифікація конструкту важлива ще й через співіснування в психологічній науці різних шкіл та напрямків, в рамках яких по різному трактуються основні рушійні сили психіки, структура особистості і т.д.

Методи конструювання тестів

Перший метод або спосіб конструювання тесту полягає в розробці тестових завдань відповідно до обраної моделі тесту на основі сформульованої теоретичної гіпотези (теоретичного конструкту). Тобто це такий прямий безпосередній спосіб – він включає обґрунтування конструкту, розробку завдань, перевірку психометричних властивостей.

Спочатку обирається форма завдань тесту, серед яких є:

- завдання із запропонованими відповідями (закриті);
- завдання з вільними відповідями (відкриті).

Завдання із запропонованими відповідями у свою чергу діляться на: завдання з альтернативними відповідями; завдання із множинним вибором; завдання з вираженням відповіді у вигляді відновлення частин речень, фігур і т.д.; завдання з переструктуруванням даних, що вимагають перегрупування елементів.

Завдання з вільними відповідями включають: завдання доповнення символу, малюнка, висловлення і т.п.; завдання вільного конструювання відповіді, які у свою чергу діляться на завдання з малоструктурованим стимулом (проективні) і завдання конструювання за чітко заданою схемою.

Далі розробляється низка завдань, яка спочатку має бути як мінімум у два рази більшою, ніж потрібна кількість.

Дотримання певних правил розробки тесту дозволить заздалегідь підвищити надійність і валідність майбутнього тесту. Ці правила стосуються створення легких для розуміння завдань (запобігання вкладання більш ніж однієї думки або ситуації в одне завдання, використання простої загальноприйнятої мови, чітких, однозначних, лаконічних, оптимальних за довжиною формулювань), відносно нових за способом рішення і змістом завдань (без банальностей, повторень одних і тих же формулювань з широковідомих тестів, сформульованих так, щоб не слугувати підказкою до інших завдань), завдань економічних з точки зору виконання та обробки, таких, що допускають можливість формалізації остаточних результатів, чіткої стандартизації процедури й інтерпретації, завдань адекватних (реальних, сучасних, не надто екстравагантних та таких, що ставлять досліджуваного в незручне, дурне становище, породжують роздратування, відчуття безглуздості всієї процедури тестування).

Відібрані завдання розміщуються за ступенем труднощі (тести інтелекту, здатностей, досягнень), у випадковому порядку (особистісні і інтелектуальні тести) або в певному, запланованому згідно із замислом тесту, порядку. Вони також можуть поєднуватися в блоки для вимірювання певних факторів в черзі (в багатофакторних тестах, наприклад, тесті «Тривога і тривожність» Спілбергера-Ханіна).

Завданням передує назва тесту, яка має бути загально-нейтрального характеру для того щоб, з одного боку не розкрити істинну валідність тесту, а з іншого – створити необхідну

очевидну валідність. Ми рекомендуємо розробляти дві назви – справжню, для фахівців, та маскувальну – для досліджуваних, вона має розміщуватися на бланку тесту, або повідомлятися усно, якщо тест не бланковий, а предметний. Після назви розміщується інструкція, яка пояснює спосіб виконання тесту, часові обмеження і приклади розв’язання типових завдань, якщо це тести інтелекту, здібностей, досягнень.

Другий спосіб конструювання тестів – критеріально-ключовий принцип – це коли емпірично підшукуються психологічні ознаки, які дозволяють диференціювати релевантні критеріальні групи (де концентруються носії психологічних якостей, які вимірюються) від контрольних (що складаються з носіїв інших ознак або недиференційовані - звичайні, нормальні люди). Класичними прикладами методик, створених за цим принципом, є Мінесотський багатоаспектний особистісний опитувальник (ММРІ) та Бланк інтересів Стронга: перший – у клінічній психології, другий – в області профорієнтації. Так, при розробці ММРІ з первісного банку тверджень в основні клінічні шкали включалися тільки ті твердження, які добре диференціювали досліджуваних з тим або іншим клінічним діагнозом від контрольної групи здорових людей. В якості критерію, тут використовувався психіатричний діагноз досліджуваних, встановлений традиційним способом. До шкали “Бланка інтересів” Стронга ввійшли ті твердження з первісного набору, які реально розрізняли осіб, що були носіями окремих професійних інтересів.

До речі, для тестів, побудованих за критеріально-ключовим методом і за факторно-аналітичним методом (дивися нижче), завдання, з яких буде складатися тест, розробляються як вказувалося вище – відповідно до обраної форми завдань тесту, враховуючи правила розробки завдань тесту. Далі тест перевіряється за всіма психометричними вимогами – дискримінативністю, надійністю, валідністю і стандартизується.

Третій спосіб – факторно-аналітичний принцип – це конструювання тестів за допомогою факторизації великої кількості завдань, які попередньо сформульовані як релевантні конструкту або конструктам (цей спосіб дозволяє відразу сконструювати багатофакторні тести або декілька однофакторних тестів). Розрізняють так званий розвідковий (експлораторний) факторний аналіз, за яким Р. Кеттелл побудував тест “Шістнадцять особистісних факторів”, і перевірочний (конфірматорний) факторний аналіз, за яким Г. Айзенк побудував тест ЕРІ (Інтроверсія-нейротизм).

Крім того факторному аналізу можуть піддаватися тести, створені іншими способами, завдяки чому перевіряється факторна структура тесту – однорідність, наявність субшкал, виявляються неякісні завдання. Далі тест валідизується і стандартизується.

Існує спосіб розробки тестів за так званою моделлю Раша. Модель Раша (Rasch, 1966) є математично набагато складнішою, ніж класична модель похибки вимірювання, за якою

будуються тести вже обговореними способами. Для її обчислення необхідна відповідна комп'ютерна програма. Вважається, що модель Раша є особливим випадком моделі латентних рис Бірнбаума (Birnbaum, 1968), відповідно до якої імовірність правильної відповіді – це функція від складності завдання, здібності тестувальника і параметра угадування.

Для шкалювання за Рашем необхідно розуміти так звану теорію «завдання-відповідь», за якою вважається, що для дихотомічних завдань існують криві їхньої здібності виявити латентну рису особистості. Наприклад, завдання «а» може бути вирішено досліджуваними з низькими інтелектуальними здібностями з імовірністю 0,015, а завдання «в» з більшою імовірністю 0,15. Теж саме завдання «а» може бути вирішено досліджуваними з середніми здібностями з імовірністю 0,8, а завдання «в» з меншою імовірністю 0,6. І нарешті, завдання «а» може бути вирішено досліджуваними з високими інтелектуальними здібностями з імовірністю 0,95, а завдання «в» з більшою імовірністю 1,00. Якщо нанести на графік ці дані, то криві завдань «а» і «в» будуть різними – одна більш пологою, інша – більш крутою, а в цілому вони нагадують криві нормального розподілу (див. рис. 1).

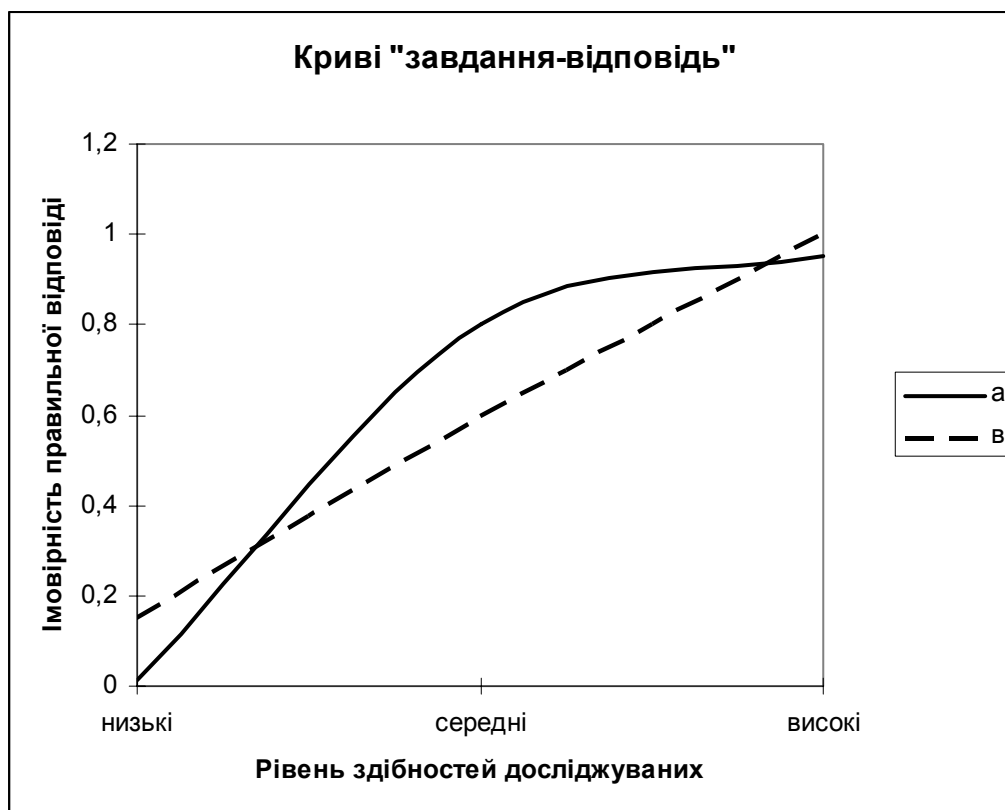


Рис. 1 Криві «завдання-відповідь»

Таким чином, імовірність правильної відповіді залежить від сполучення ступеню виразності даної риси у досліджуваного і здібності завдання виявити дану рису. Тобто є завдання, які краще або гірше виявляють певну рису. Це їхня незалежна характеристика, яка не так прямо прив'язана до відсотку людей, які впоралися з завданням в нормативній вибірці, як це робиться для визначення трудності завдань в класичній моделі.

Для створення шкали Раша спочатку завдання апробуються (калібруються, градууються) на репрезентативній вибірці не менш ніж 1000 досліджуваних. За її допомогою та з використанням відповідної комп'ютерної програми виявляють ці незалежні властивості завдань щодо вимірювання певної психологічної риси, тобто будуються криві «завдання-відповідь». Подальше застосування програми призводить до того, що завдання та групи завдань мають точно визначену трудність і дискримінативність і можуть застосовуватися для утворення якісних (однорідних і дискримінативних) тестів, де вони будуть розташовані зі зростаючою трудністю, а також будуть відрізнятися одне від одного на строго однакову величину психологічної ознаки, тим самим утворюючи пропорційну шкалу (шкалу відносин).

Ще за допомогою завдань, прошкальованих за Рашем, можна вирішувати проблему створення дійсно еквівалентних (паралельних) форм тестів для перевірки надійності паралельних форм тесту або повторного тестування.

Шкалювання за Рашем надає можливість створення коротких, проте надійних тестів, що в традиційній моделі досягається за рахунок збільшення кількості завдань.

Прошкальовані за Рашем завдання ефективно використовувати для адаптивного тестування. Адаптивне тестування – комп'ютерне тестування за допомогою певного алгоритму, коли, якщо людина не може впоратися з певним завданням, їй далі пред'являються не всі завдання, в тому числі більш складні, а ті, які будуть легше. І навпаки, якщо людина може впоратися з певним завданням, їй пропонуються більш складні завдання допоки вона зможе їх вирішувати. Таким чином, швидко визначається рівень здібностей досліджуваного, тестування стає більш коротким і менш фруструючим для досліджуваного.

Передбачалося також, що шкали Раша будуть придатні для створення крос-культурних тестів, але ця можливість піддається критиці, принаймні для дуже «далеких» за типом культури популяцій, наприклад, ескімосів і китайців.

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.

Мельников В.М., Ямпольский Л.Т. Введение в экспериментальную психологию личности. – М.: Просвещение, 1985.

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.

Практикум по психодиагностике. Дифференциальная психометрика / Под ред. В.В. Столина, А.Г. Шмелёва. - М.: Изд-во Моск. ун-та, 1984.

Психологическая диагностика: Учебник для вузов / Под ред. М.К. Акимовой, К.М.Гуревича. – СПб.: Питер, 2008.

Шмелёв А.Г. Психодиагностика личностных черт. – СПб.: Речь, 2002.