

## Лекція №6 з дисципліни «психодіагностика».

### ВАЛІДНІСТЬ ТЕСТУ. ФАКТОРИЗАЦІЯ ТЕСТУ

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

#### Валідність тесту

Валідність - комплексна характеристика тесту, що включає відомості про область досліджуваних явищ і репрезентативність використовуваної стосовно неї діагностичної процедури (Л.Ф.Бурлачук). Інакше кажучи, тест називається валідним, якщо він вимірює те, для вимірювання чого він призначений (П. Клайн). А. Анастасі пише, що валідність тесту це поняття, яке вказує нам, що тест вимірює і наскільки гарно він це робить.

У поняття валідності входить велика кількість інформації про тест. Різні категорії такої інформації і способи її одержання утворюють типи валідності.

Очевидна (зовнішня, лицьова) валідність – уявлення про тест, сферу його застосування, результативність й прогностичну цінність, що виникає у досліджуваного або іншої особи, яка не володіє спеціальними знаннями про справжню валідність. Очевидна валідність зазвичай демонструється через назву й інструкцію до тесту, і допомагає налагодити співробітництво з досліджуваними.

Основними видами справжньої валідності є: змістовна, конструктна і критеріальна валідності.

#### Змістовна валідність

Змістовна валідність (внутрішня, логічна) – комплекс відомостей про репрезентативність змісту завдань тесту відносно вимірюваних психологічних властивостей і особливостей. Найбільше значення цей вид валідності має для тестів досягнень (тестів успішності, екзаменаційних тестів). Для конструювання завдань тесту залучаються експерти з досліджуваної якості, література, словники.

Завдання – це вибірка поведінки, що відображена в твердженнях (малюнках, предметах і т.п.) тесту, за якою стоїть психологічна якість. Важливо, щоб завдання відбивали всі боки психологічної якості. Тому робиться так звана специфікація тесту – розрахунок того, скільки завдань має бути з кожної складової якості. Таким

чином можна зробити відсотковий підбір складових конструкту, а потім перевести його в кількість завдань залежно від того, скільки завдань планується мати в тесті.

Після генерування необхідної кількості завдань, вони піддаються експертній оцінці. Експерти в даному випадку – це фахівці – психологи, психіатри і п.т., які гарно розбираються в досліджуваних психічних явищах. (Існують два способи експертної оцінки – колективна і середньозважена. Остання може організовуватися через усереднення рейтингів, шкалювань або парних порівнянь). Таким чином змістовна валідність формується однією з перших видів валідності і перевіряти її також можна відразу.

### Конструктна валідність (ідентифікація конструкту)

Конструктна валідність (теоретична) – відбиває ступінь репрезентації досліджуваного психологічного конструкта в результатах тесту; чим більше результати тесту відповідають теоретичній гіпотезі про природу вимірюваної змінної, тим вище конструктна валідність тесту.

При розробці тесту формулюється низка гіпотез про те, як буде пов'язаний досліджуваний тест із іншими тестами, що вимірюють подібні, протилежно спрямовані або незалежні психологічні явища. Ці гіпотези висуваються на підставі теоретичних уявлень про вимірювані властивості як про психологічні конструкти. Підтвердження гіпотез свідчить про теоретичну обґрунтованість методики, тобто про ступінь її конструктної валідності. Цей тип валідності є найбільш складним і комплексним.

Коли ми ідентифікуємо конструкт ми працюємо нібито хіміки, які для того, щоб дослідити якусь невідому речовину, поміщують її у контакт з відомими – кислотою, лугом й іншими. Потім вони спостерігають, як вона реагує з ними і роблять висновок, що ж це за речовина.

Так само з конструктом тесту – його піддають кореляції з іншими вже валідизованими тестами, які вимірюють подібні ж явища і мають корелювати з цим новим тестом (конвергентна валідність); його піддають кореляції з іншими вже валідизованими тестами, які вимірюють незв'язані, зовсім інші явища і не повинні корелювати з цим новим конструктом (дивергентна валідність). Також висуваються

і перевіряються гіпотези про наявність або відсутність статевих розбіжностей у групах чоловіків і жінок за даним тестом; висувуються і перевіряються гіпотези про наявність або відсутність за цим тестом вікових розбіжностей в групах молодих, середніх і старших за віком людей; пропонуються способи психофізіологічної та експериментальної перевірки валідності тесту.

Якщо конструкт тесту в процесі ідентифікації демонструє передбачувані теорією результати, він вважається валідним. Конструктна валідність має особливе значення для валідизації тестів особистості, але тести інтелекту і здібностей теж валідизують за конструктом. Для цього використовують методи декомпозиції завдання (варіювання пред'явлення завдань частково і повністю, з різними інструкціями і підказками), а також аналіз протоколів – голосний синхронний переказ того, що людина думає під час розв'язання завдань.

#### Критеріальна валідність

Критеріальна валідність – комплекс характеристик методики, що відбиває відповідність діагнозу і прогнозу певному колу критеріїв вимірюваного явища. Критеріальна валідність включає як комплекс зв'язків з поточним станом досліджуваного явища (поточна критеріальна валідність), так і ймовірність і обґрунтованість прогнозу про його стан у даної людини в майбутньому (прогностична критеріальна валідність). Найбільше значення цей вид валідності має для валідизації тестів професійних інтересів, мотивів, граничних нервово-психічних розладів, специфічних типів поведінки (асоціальних), здібностей тощо.

Перевіряється критеріальна валідність шляхом статистичного корелювання балів за тестом та показників за зовнішнім параметром, вибраним в якості критерію валідизації.

Валідизація за будь-яким критерієм ефективно здійснюється методом *контрастних груп* (розумово відсталих і з нормальним інтелектом, що страждають і не страждаючими психічними захворюваннями, асоціальних-просоціальних і т.п.).

*Критерій валідизації* є незалежним від тесту показником, що має безпосередню цінність для певних областей практичної діяльності. Наприклад, у

педагогічній психології - це «успішність», у психології праці - «продуктивність», у медичній психології - «стан здоров'я» і т.д.

У якості об'єктивних критеріїв валідизації часто використовують:

- об'єктивні соціально-демографічні й біографічні дані (стаж, освіта, професія, прийоми або звільнення з роботи, кількість браків та розлучень тощо);
- показники успішності навчання, які найчастіше є критеріальною мірою тестів здібностей, досягнень в окремих дисциплінах, тестів інтелекту;
- виробничі показники ефективності виконання певних видів професійної діяльності, що мають найбільше значення під час валідизації методик, використовуваних у профвідборі й профорієнтації;
- результати реальної діяльності (малювання, моделювання, музика, складання розповіді й т.д.) при випробуванні тестів загальних і спеціальних здібностей, креативності, тестів особистості;
- лікарський діагноз або інші висновки фахівця;
- контрольні випробування знань і вмінь.

Суб'єктивними критеріями є оцінки, судження, висновки про досліджуваного, зроблені експертом – людиною, яка знає досліджуваного з боку оцінюваної якості і може надати їй адекватну оцінку (педагогом, керівником, психологом, сусідом, родичем). Об'єктивізація суб'єктивного критерію валідизації досягається збільшенням числа експертів. Крім цього критерій валідизації повинен бути багатокомпонентним, наприклад, експертне опитування плюс біографічні дані, плюс результати тестування контрастних груп.

#### Перехресна валідизація (крос-валідизація)

Важливо, щоб валідність тесту перевірялася на вибірці досліджуваних, відмінної від тієї, за допомогою якої здійснювався аналіз і відбір завдань. Будь-який коефіцієнт валідності, знайдений на вибірці, що застосовувалася для відбору завдань, виявиться завищеним, тому що випадкові особливості цієї вибірки будуть враховуватися як справжні властивості популяції.

Наприклад, для валідизації тесту для відбору майбутніх студентів медичного факультету використовувалася критеріальна валідизація за допомогою контрастних

груп. З кожного курсу відраховується та частина студентів, яка не може оволодіти програмою навчання. За логікою критеріальної валідизації вважається, що можна взяти контрастні групи найкращих і найгірших студентів (успішність навчання буде критерієм) і виявити всі якості, що статистично достовірно розрізняють групи кращих та гірших студентів. За ті якості, що характеризують верхню групу присвоювати позитивні бали до ключа, за ті – що характеризують нижню – негативні. Таким чином за допомогою створеного тесту можна буде одбирати здатну до успішного навчання молодь ще на етапі вступу до вузу. Припустимо, що вибіркою були 100 студентів-медиків і серед них 30 кращих і 30 гірших за успішністю навчання були порівняні за різними характеристиками. Будуть виявлені певні розбіжності, наприклад, в групі кращих було більше дівчат на ім'я Таня і хлопців на ім'я Сергій. А в групі відстаючих було більше Марин та Андріїв. Ми можемо надавати за те, що дівчину звать Таня більше балів, ніж якщо її звать Марина. Але якщо ми візьмемо інший курс – скоріше за все ця закономірність не підтвердиться, або виявиться ще якась нова. Так само і з психологічними ознаками – вони теж можуть бути випадковим феноменом певної вибірки і треба, щоб вони підтвердилися на іншій групі, щоб довести, що вони справді пов'язані з критерієм.

Чим меншими будуть вибірки досліджуваних під час аналізу завдань, менша кількість завдань залишена після аналізу, а також якщо завдання відбиралися суто емпірично, без теоретичного обґрунтування, тим більше вірогідність того, що вони будуть мати після крос-валідизації незадовільну валідність.

### Факторній аналіз (факторизація тесту)

Факторизація тесту є іншим способом проаналізувати і відібрати ефективні завдання тесту, а також остаточною перевіркою гомогенності тесту. Математичний алгоритм факторного аналізу дозволяє виявити, наскільки завдання тесту корелюють між собою, з чого робиться припущення, що вони всі так «поводяться» через кореляцію з гіпотетичним фактором, який і є конструктором тесту – певною латентною психологічною змінною, про яку можна судити через спостережувані

ознаки поведінки людини. Факторний аналіз тесту покаже, які завдання не увійшли до головного фактора, або увійшли до нього з надто низькими кореляціями ( $\leq 0,7$ ). Їх потрібно виключити або переформулювати таким чином, щоб вони більше відповідали психологічній змінній, на діагностику якої спрямовано тест, і потім повторити процедуру факторизації. Ця процедура аналогічна аналізу завдань.

В результаті буде водночас перевірена і надійність тесту за внутрішньою погодженістю, яка ще має назву факторно-дисперсійної надійності. Кумулятивна дисперсія змінних, яка пояснюється значущими факторами, має бути не менше 60% - тоді тест, або частина його факторів – надійний(ні).

Ще факторний аналіз показує наявність одного або більше значущих факторів, що виявляє внутрішню структуру тесту. Якщо виявлено декілька значущих факторів, можна зробити багатфакторний тест зі субтестами, або одразу зробити декілька одно факторних тестів.

Про процедури факторного аналізу і оцінку факторно-дисперсійної надійності можна прочитати у спеціальній літературі.

Факторний аналіз використовується також для з'ясування певних внутрішніх характеристик самого конструкту тесту. Наприклад, за його допомогою (а також за допомогою кластерного аналізу) здійснювалося визначення, чи є новий конструкт «перфекціонізм», тобто прагнення до досконалості в будь-якої діяльності, категоріальним або континуальним. Певні дослідники вважають його протяжним, з негативними виходами на екстремальних значеннях у вигляді обсесивно-компульсивних розладів, порушень харчової поведінки тощо. Інші вважають, що перфекціонізм має дві різні форми – адаптивну і неадаптивну. Проведення цих математичних процедур дозволило одній групі дослідників довести континуальність даної риси.

### **Рекомендована та використана література:**

Анастази А., Урбина С. Психологическое тестирование. - СПб.: Питер, 2003.

Благуш П. Факторный анализ с обобщениями. – М.: Финансы и статистика, 1989. – 247 с.

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.

Лоули Д., Максвелл А. Факторный анализ как статистический метод. – М.: Мир, 1967. – 144 с.

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

Общая психодиагностика / Под ред. А.А. Бодалева, В.В. Столина. СПб.: Изд-во «Речь», 2004.

Окунь Я. Факторный анализ. – М.: Статистика, 1974. – 200 с.

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.

Факторный, дискриминантный и кластерный анализ /Дж.-О. Ким, Ч. У. Мьюллер, У. Р. Клекка и др.; Под ред. И. С. Енюкова. – М.: Финансы и статистика, 1989. – 215 с.