

**Олефір Валерій Олександрович**

Харківський національний університет імені В. Н. Каразіна

## **ОЦІНКА КОНСТРУКТНОЇ ВАЛІДНОСТІ ПСИХОЛОГІЧНОГО ТЕСТУ: ПІДХІД «МНОЖИННІ ОЗНАКИ – МНОЖИННІ МЕТОДИ»**

*Проаналізовано сучасні підходи до оцінки конструктної валідності тесту на основі аналізу мультиознакових-мультиметодних матриць процедурою конфірматорного факторного аналізу.*

**Ключові слова:** конструктна валідність, мультиознакові-мультиметодні матриці, індекси придатності.

**Проблема, її зв'язок із важливими науковими чи практичними завданнями.** Психологічні вимірювання надзвичайно важливі для наук про поведінку та їх практичних додатків. Як наріжний камінь психометричної якості валідність (як і надійність) має фундаментальне значення для розуміння і оцінки якості психологічних вимірювальних інструментів. Сучасний погляд на проблему валідності надає особливого значення теоретичному змісту конструкту, який, як вважається лежить в основі інтерпретації тестових балів. Отож конструктна валідність визначається як «ступінь того, наскільки тестові бали можуть інтерпретуватися як емпіричний показник певного психологічного конструкту» [0, с. 195].

З точки зору Л. Кронбаха та П. Міла, необхідна умова того, що вимірювання має конструктну валідність, полягає в тому, що конструкт повинен бути представлений в номологічній мережі, тобто в мережі «значень», що його оточують [0]. Проте автори не розробили ніякого статистичного строгого методу оцінки конструктної валідності.

Суттєвим кроком вперед у статистичному і методологічному розвитку поняття конструктної валідності стала одна із найбільш цитованих робіт – стаття Д. Кемпбелла і Д. Фіске, які запропонували так звані мультиознакові-мультиметодні матриці (МОМММ) [0]. Такі матриці утворюються в результаті того, що корелюються показники за декількома психологічними характеристиками, кожна з яких виміряна декількома методами. Алгоритм Кемпбелла-Фіске до аналізу матриць МОММ: уважно оглянути кореляції. Шукати:

1. Кореляції між однаковими характеристиками і різними методами (моноознакові-гетерометодні кореляції) і виділити жирним шрифтом («діагонали валідності»). Досить високі кореляції надають свідчення на користь конвергентної валідності інструменту.

2. Кореляції між різними властивостями і різними методами. Вони не повинні бути занадто високими, особливо по відношенню до однакових рис. Це показник дискримінантної валідності: метод вимірювання повинен розрізняти різні риси.

3. Дисперсію методу: якщо різниці між методами не існує, то співвідношення між різними ознаками й однаковим методом будуть такими самими, як і кореляції різних характеристик з різними методами.

Незважаючи на те, що методологія Кемпбелла та Фіске представляла значний прогрес у концептуалізації та оцінці обґрунтованості конструкту, підхід МОММ не був широко використаний протягом багатьох років, безпосередньо після його створення. Крім того, було відзначено кілька обмежень методології МОММ, включаючи суб'єктивний характер його інтерпретації, залежність від кореляцій між невідповідними вимірами [0], а також нездатність експлораторного факторного аналізу отримувати значущі рішення даних МОММ (наприклад, обмеження щодо специфікації корельованих помилок). Проте інтерес до матриць МОММ збільшувався з усвідомленням того, що процедури конфірмаційного факторного аналізу (КФА) можуть бути легко застосовані до їх аналізу, щоб зробити висновки про основні параметри, такі як ознаки та фактори методу.

**Аналіз публікацій (виділення невирішених проблем).** Як свідчать публікації, до даних МОММ можуть бути застосовані декілька різних типів моделей КФА [0; 0], переважають дві форми специфікації КФА. Для використання більш сучасної термінології ці два типи рішень називаються корельованими методами (стандартна модель) та корельованими моделями унікальності [0].

У стандартній моделі КФА МОММ повинні бути навантаження кожного вимірювання на його характеристики і фактори методу. Корельованість як фактори характеристик, так і фактори методу. Зазвичай характеристики і фактори методу незалежні. Повинно бути принаймні три риси і методи для цього підходу. На рис. 1 представлена модель КФА з корельованими рисами і методами для валідації опитувальника соціальної компетентності (див табл. 1).

### Збірник наукових статей

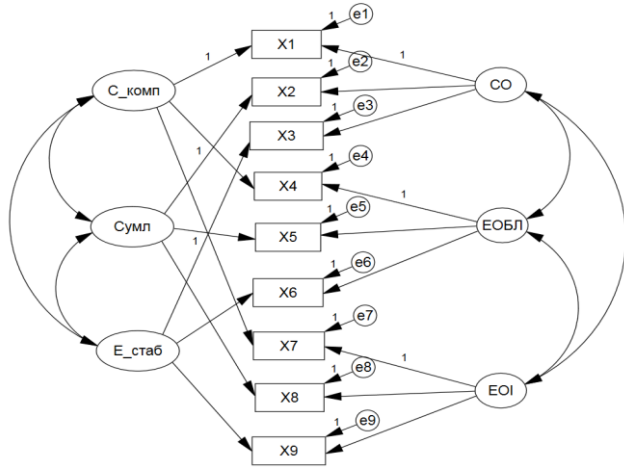


Рис. 1. Стандартна модель КФА для даних МОММ.

При проведенні оцінки параметрів моделі досить часто зустрічаються випадки Хейвуда, неможливі значення (кореляції більші, ніж одиниця і негативні дисперсії), і проблеми сходження. Таким чином, з огляду на ці проблеми, «стандартна» модель може бути проблематичною.

Модель КФА корельованих унікальностей [0] була представлена як альтернативний підхід до аналізу даних МОММ. На рис. 2 зображено шляхову діаграму специфікації КФА моделі корельованої унікальності для матриці МОММ соціальної компетентності (табл. 1).

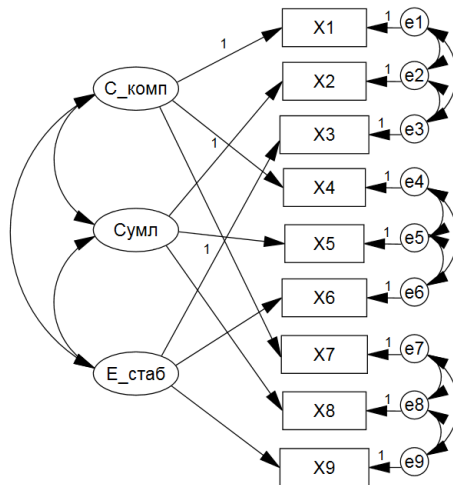


Рис. 2. Корельованих унікальностей специфікація КФА для МОММ

Для того, щоб ідентифікувати модель корельованої унікальності, повинно бути щонайменше дві риси та три методи. Специфікація характерної частини моделі корельованої унікальності така сама, як і методу корельованих методів: (1) Кожен показник завантажується на один фактор ознаки; і (2) кореляції між факторами ознак вільно оцінюються. Таким чином, ключовою різницею між цими параметризаціями є спосіб оцінювання ефектів методу. У моделі корельованої унікальності ефекти методу оцінюються шляхом визначення корельованих унікальностей (помилки) серед індикаторів, що базуються на тому ж методі оцінки, а не за факторами методу.

На відміну від моделей корельованих методів, у моделей корельованих унікальностей рідко виникають проблеми оцінювання. Показано, що специфікації моделей корельованих унікальностей призвели до правильних рішень у 98% випадків. Проте інтерпретація корельованих унікальностей як ефектів методу не завжди проста. Іншим потенційним недоліком моделі корельованої унікальності є її припущення, що кореляції між методами та співвідношення між рисами та методами дорівнюють нулю. Оцінювання параметрів рішення корельованих унікальностей може бути упередженим, коли припущення про нульові кореляції між методами та між методами та рисами не виконується. Якщо ці нульові кореляційні обмеження не є доцільними, сума дисперсії ознак та коваріації між факторами ознак буде переоцінена, що призведе до завищених оцінок конвергентної валідності та нижчих оцінок дискримінантної валідності відповідно.

**Мета** дослідження – ознайомити з оцінкою конвергентної та дивергентної валідності психологічного тесту на основі матриць МОММ з допомогою моделі конфірматорного факторного аналізу.

**Виклад основного матеріалу.** Оскільки модель корельованих факторів є базовою для оцінювання конструктивної валідності тесту на основі матриці МОММ, то КФА даних проведемо згідно цієї моделі. В якості вихідної матриці МОММ скористаємося модифікованими гіпотетичними даними (див. табл. 1) з книги Р. Фера та В. Бакарака [0].

Результати КФА моделі корельованих методів узагальнені в табл. 2, щоб продемонструвати інтерпретацію ефектів рис та методу. Дані мали прийнятні індекси відповідності даних до моделі ( $\chi^2 = 11,24$ ,  $p = 0,59$ ,  $df = 13$ ;  $RMSEA = 1,00$ ). Оцінка риси сумлінності мала найвищу дисперсію помилок при використанні методу експертних оцінок близької людини та експертної оцінки інтерв'юера (помилки = 0,734, відповідно). Тому сумлінність була найбільш складною ознакою для оцінки з використанням будь-якого з цих двох методів. Методи: експертна оцінка близької людини та експертна оцінка інтерв'юера – найкращі для оцінки соціальної компетентності (факторні

### Збірник наукових статей

навантаження = 0,547 та 0,485, дисперсія помилок = 0,072 та 0,586, відповідно).

Таблиця 1

Кореляційні взаємозв'язки в матриці множинних методів і ознак

| Ознаки  | Методи  |       |         |         |       |         |         |       |         |
|---------|---------|-------|---------|---------|-------|---------|---------|-------|---------|
|         | СО      |       |         | ЕОБЛ    |       |         | ЕОІ     |       |         |
|         | С_комп. | Сумл. | Е_стаб. | С_комп. | Сумл. | Е_стаб. | С_комп. | Сумл. | Е_стаб. |
| С_комп. | 1,0     |       |         |         |       |         |         |       |         |
| Сумл.   | 0,20    | 1,0   |         |         |       |         |         |       |         |
| Е_стаб. | 0,35    | 0,19  | 1,0     |         |       |         |         |       |         |
| С_комп. | 0,40    | 0,10  | 0,22    | 1,0     |       |         |         |       |         |
| Сумл.   | 0,09    | 0,36  | 0,14    | 0,14    | 1,0   |         |         |       |         |
| Е_стаб. | 0,20    | 0,11  | 0,41    | 0,30    | 0,18  | 1,0     |         |       |         |
| С_комп. | 0,34    | 0,19  | 0,20    | 0,23    | 0,11  | 0,19    | 1,0     |       |         |
| Сумл.   | 0,09    | 0,30  | 0,14    | 0,09    | 0,20  | 0,06    | 0,24    | 1,0   |         |
| Е_стаб. | 0,14    | 0,08  | 0,33    | 0,13    | 0,06  | 0,19    | 0,44    | 0,29  | 1,0     |

Примітка: Тут і далі. С\_комп. – соціальна компетентність, Сумл. – сумлінність, Е\_стаб. – емоційна стабільність, СО – самооцінка, ЕОБЛ – експертна оцінка близької людини, ЕОІ – експертна оцінка інтерв'юера.

Таблиця 2

МОММ оцінки трьох властивостей трьома методами

|         | Характеристики |       |         | Методи |       |       | Помилка |
|---------|----------------|-------|---------|--------|-------|-------|---------|
|         | С_комп.        | Сумл. | Е_стаб. | СО     | ЕОБЛ  | ЕОІ   |         |
| С_комп. | 0,680          |       |         | 0,253  |       |       | 0,475   |
| Сумл.   |                | 0,691 |         | 0,195  |       |       | 0,488   |
| Е_стаб. |                |       | 0,769   | 0,231  |       |       | 0,361   |
| С_комп. | 0,791          |       |         |        | 0,547 |       | 0,072   |
| Сумл.   |                | 0,516 |         |        | 0,005 |       | 0,734   |
| Е_стаб. |                |       | 0,554   |        | 0,035 |       | 0,691   |
| С_комп. | 0,430          |       |         |        |       | 0,485 | 0,586   |
| Сумл.   |                | 0,414 |         |        |       | 0,310 | 0,734   |
| Е_стаб. |                |       | 0,396   |        |       | 0,329 | 0,734   |

Примітка:  $\chi^2 = 11,245$ ,  $p = 0,590$ ,  $df = 13$ ;  $RMSEA = 1,000$ ;  $n = 240$ .

С\_комп. – соціальна компетентність, Сумл. – сумлінність, Е\_стаб. – емоційна стабільність, СО – самооцінка, ЕОБЛ – експертна оцінка близької людини, ЕОІ – експертна оцінка інтерв'юера.

Конвергентна валідність – середнє навантаження (при аналізі кореляцій):

- за ознакою: соціальна компетентність (0,634), сумлінність (0,540) і емоційна стабільність (0,573);

- за методом: самооцінка (0,713), експертна оцінка близької людини (0,620) і експертна оцінка інтерв'юера (0,413).

Дискримінаційна валідність – кореляції ознак;  $r_{12} = 0,368$ ;  $r_{13} = 0,567$ ; і  $r_{23} = 0,317$ . Хороша дискримінантна валідність.

Дисперсія методу: середня коваріантність помилок: 0,441 – самооцінка; 0,499 – експертна оцінка близької людини і експертна оцінка інтерв'юера – 0,685. Найбільший ефект методу для експертної оцінки інтерв'юера.

**Висновки, перспективи.** Сучасний погляд на конструктну валідність полягає в тому, що теоретичний зміст конструкту лежить в основі інтерпретації тестових балів. Бал за тим чи іншим психологічним тестом повинен інтерпретуватися як емпіричний показник певного психологічного конструкту. Для оцінки конструктної валідності були запропоновані мультиознакові-мультиметодні матриці. Сучасний підхід до їх аналізу полягає в застосуванні процедури конфірматорного факторного аналізу, які дозволяють зробити висновки про основні ефекти ознаки та методів і оцінити конвергентну та дивергентну валідність. Показана можливість оцінки валідності з допомогою моделі корельованих факторів. Перспективним є порівняння оцінок конструктної валідності отриманих на основі різних типів моделей КФА.

### Література

1. Фер Р. М. Психометрика: Введение / Р. М. Фер, В. Р. Бакарак. – Челябинск, Издательский центр ЮУрГУ, 2010. – 445 с.
2. Campbell D. T. Convergent and discriminant validation by the multitrait-multimethod matrix / D. T. Campbell, D. W. Fiske // Psychological Bulletin. – 1959. – Vol. 56. – P. 81-105.
3. Cronbach L. J. Construct validity in psychological tests / L. J. Cronbach, P. E. Meehl // Psychological Bulletin. – 1955. – Vol. 52. – P. 281-302.
4. Marsh H. W. Latent variable models of multitrait-multimethod data / H. W. Marsh, D. Grayson // Hoyle R. H., editor. Structural equation modeling: Concepts, issues, and applications. T– housand Oaks, CA: Sage; 1995. – P. 177-198.
5. Schmitt N. Methodology Review: Analysis of Multitrait-Multimethod Matrices / N. Schmitt, D. M. Stults // Applied Psychological Measurement. – 1986. – Vol. 10 (1). – P. 1-22.

6. Widaman K. F. Hierarchically nested covariance structure models for multitrait–multimethod data / K. F. Widaman // Applied Psychological Measurement. – 1985. – Vol. 9. – P. 1-26.

**Валерий Олефир. Оценка валидности психологического теста: подход «множественные признаки – множественные методы».**

*Проанализированы современные подходы к оценке конструктивной валидности теста на основе анализа матриц множественных признаков – множественных методов процедурой подтверждающего факторного анализа.*

**Ключевые слова:** конструктивная валидность, мультипризнаковые-мультиметодные матрицы, индексы годности.

**Valerii Olefir. Evaluation of the construct validity of the test: the «multitrait- multimethods» approach.**

*Modern approaches to the evaluation of the construct validity of the test based on the analysis of multi-indicative-multimethod matrixes by the procedure of confirmatory factor analysis are analyzed.*

**Keywords:** construct validity, multitrait–multimethod matrix, fit indices.

Стаття надійшла до редакційної колегії 30.09.2017

**Інформація про автора:**

**Олефір Валерій Олександрович** – доктор психологічних наук, доцент, професор кафедри загальної психології, Харківський національний університет імені В.Н. Каразіна.