

УДК 303.214.3

СЕРГІЙ ДЕМБИЦЬКИЙ,

кандидат соціологічних наук, провідний соціолог відділу історії, теорії та методології соціології Інституту соціології НАН України

Теоретична валідизація вимірювальних шкал

Анотація

Статтю присвячено специфіці теоретичної валідизації вимірювальних шкал. Такого ґатунку валідизація передбачає проходження певних етапів: підготовка теоретичної бази, розроблення та коригування емпіричних індикаторів, перевірка конструктивної валідності та ретестової надійності, опис специфіки інтерпретації отриманих результатів. Головною відмінністю поданого у статті підходу від класичної теорії вимірювання, використовуваної в соціальних науках, є орієнтація на валідну теорію, а не на деяку гіпотетичну теорію, що підлягає перевірці. У контексті використовуваної (валідної) теорії здійснюються всі етапи теоретичної валідизації — від вибору типу застосовуваної шкали до особливостей інтерпретації отриманих результатів.

Емпіричний базис статті — результати теоретичної валідизації комплексної 10-бальної шкали, призначеної для вимірювання рівня зусиль, що їх докладають студенти в процесі навчання. У рамках розроблення цієї методики використано дані одного якісного (вивчення феномену навчання студентів за умов обмеженого часу) і двох змішаних досліджень (вивчення поведінкових практик студентів, котрі поєднують роботу і навчання, а також аналіз впливу організаційної структури малих кафедр, що випускають, на особливості навчання студентів).

Ключові слова: *вимірювання, теоретична валідизація*

Вступ

У своїх недавніх працях [Дембицький, 2010а; Дембицький, 2010с; Дембицький, 2009; Дембицький, 2008] я заторкнув широкий спектр питань, що

стосуються проблеми теоретичної валідації. Запропонований мною підхід можна резюмувати в таких положеннях:

1. Теоретична валідність являє собою міру відповідності теоретичного конструкту досліджуваному феномену.
2. Теоретична валідність є комплексною характеристикою, що включає очевидний, змістовий, логічний і композиційний компоненти.
3. Теоретична валідація більшою мірою пов'язана з **досягненням** необхідного рівня відповідності теоретичного конструкту досліджуваному феномену, ніж із перевіркою.
4. Завершений процес теоретичної валідації пов'язаний із забезпеченням інших видів валідності: конструктної, зовнішньої, внутрішньої та прогностичної.
5. Теоретична валідація має здійснюватися із застосуванням комплексних дослідницьких стратегій.
6. Теоретична валідація передусім пов'язана із парадигмою досліджень змішаного типу (mixed methods research).

У рамках пропонованої статті буде продемонстровано роль теоретичної валідації у здійсненні соціологічного вимірювання — від критеріїв побудови шкали і до специфіки інтерпретації отриманих даних. Як приклад використовується шкала, призначена для вимірювання зусиль, що їх докладають студенти в процесі навчання.

Ця шкала базується на теоретичному конструкті, сформульованому в рамках дослідження, присвяченого феномену навчання за умов обмеженого часу. Конструкт описує три типи студентів залежно від їхнього ставлення до навчання: “байдужий”, “декларативно зацікавлений” і “відповідальний”. Попри те, що теорія справедлива стосовно студентів стаціонару, котрі поєднують роботу і навчання, поширення окремих її елементів на всіх студентів стаціонару також справедливе. Відповідні пояснення подано далі.

Перший етап: підготовка теоретичної бази

На відміну від узвичаєного підходу валідації вимірювальних шкал у рамках теоретичної валідації дослідник має розпочинати підготовку вимірювального інструменту, маючи у своєму розпорядженні валідний теоретичний конструкт, а ще краще — валідну теорію. У подальшому всі отримані результати вимірювання треба зіставляти, а в разі потреби — й модифікувати відповідно до цих теоретичних положень.

У процесі побудови шкали як теоретичну базу використано результати двох досліджень. Перше присвячене поведінковому репертуару студентів або поведінковим стратегіям, друге — згаданим вище типам студентів.

Спинімося докладніше на змістових характеристиках цих типів. Кожен із них можна описати двома поведінковими стратегіями — основною та додатковою (рідше трапляються ситуації, коли стратегії мають рівноцінне значення). Відповідні стратегії включають різний поведінковий репертуар (див. табл. 1). Опис поведінкових стратегій було отримано завдяки дослідницькій стратегії узгодження концептів [Дембіцький, 2009], що дає змогу забезпечити змістову й логічну валідність теоретичного конструкту.

Таблиця 1

Змістові характеристики поведінкових стратегій студентів

| Назва групи поведінкових практик | Зміст групи |
|---|--|
| <i>1. Стратегія нероби</i> | |
| 1.1. “Неробство”, сподівання на те, що поталанить | Давати обіцянки, що в подальшому обов’язково навчатимешся; використовувати моменти, коли викладач поспішає, й одержувати списання боргів; бути нахабним, йти на ризик (нічого не відпрацьовувати, йти з лекцій (семінарів) без дозволу тощо); не працювати впродовж семестру і надолужувати заборгованість лише на сесії; складати іспити багаторазово (доки у викладача терпець не увірветься) тощо |
| 1.2. “Мета виправдує засоби” | Використовувати зв’язки; сказати, що роботу забув удома і попросити дозволу відповідати без потрібних матеріалів; якщо здаєш невчасно, списувати готові роботи і здавати їх; умовити викладача і при цьому нічого не робити й т. ін. |
| 1.3. Посилання на зовнішні причини | Пояснювати прогалини і невідвідування зовнішніми причинами і “давити” на те, що був змушений; посилатися на побутові проблеми, що заважають бути присутнім; використовувати тему здоров’я для виправдання тощо |
| <i>2. Стратегія спритника</i> | |
| 2.1. Робота з викладачем | Вивчити викладача і знайти до нього індивідуальний підхід; якщо раніше навчався добре, використовувати свій авторитет; ранжувати викладачів за вимогливістю й залежно від цього виділяти час на підготовку |
| 2.2. Використання слухних ситуацій і моментів | У разі присутності на парах демонструвати активність; коли здаєш борги, перевести обговорення на матеріал, який знаєш; готуватися в транспорті, тощо |
| 2.3. Мінімізація зусиль | Показувати виконані завдання і відпрошуватися з пар; якщо є певний мінімум необхідних завдань, виконувати тільки його; якщо робиш завдання рідко, то принаймні якісно, бо це підкуповує; відвідувати найважливіші види занять (наприклад, практичні, а не лекції) тощо |
| <i>3. Стратегія трудяги</i> | |
| 3.1. Самоорганізація | Планувати, розкласти за пунктами те, що необхідно здійснити; всіляко використовувати самоорганізацію тощо |
| 3.2. Опора на дружні стосунки в навчальній групі | Допомагати студентам, які допомагають тобі, розподіляти підготовку завдань між різними людьми, брати в одногрупників необхідні матеріали та ін. |
| 3.3. Ввічливість | Добре поводитися на парах; демонструвати позитивне ставлення до дисципліни; коли йдеш із пар, завжди погоджувати це з викладачем тощо |
| 3.4. Опертя на власні сили | Завжди намагатися схопити суть матеріалу; якщо необхідно, залишатися після пар, щоби пересклати борги тощо; за потреби витратити весь вільний час на підготовку й т. ін. |

Своєю чергою, типи студентів були визначені в рамках використання дослідницької стратегії “обґрунтованої теорії” [Дембіцький, 2010a] і можуть бути схарактеризовані так. Якщо студент належить до “байдужого” типу, він може використовувати: а) стратегію нероби як основну і стратегію спритника як додаткову; б) стратегію спритника як основну і стратегію нероби як додаткову (або однаковою мірою). Коли студент належить до “декларативно зацікавленого” типу, тоді як основна використовується стратегія спритника, а як додаткова — стратегія трудяги. Зрештою, якщо студент належить до “відповідального” типу, він може використовувати: а) стратегію ловкача і трудяги однаковою мірою; б) стратегію трудяги як основну, а стратегію ловкача як додаткову.

Незважаючи на те, що у фокусі дослідження перебували студенти стаціонару, які поєднують роботу і навчання, отримані дані можна розподілити на дві частини — ту, що стосується тільки студентів, котрі працюють, і ту, що є спільною як для тих, хто працює, так і для тих, хто не працює. Зокрема, співвідношення типів студентів і відповідних поведінкових стратегій є справедливим для всіх студентів, що забезпечує високу зовнішню валідність. Фактично обидві групи з точки зору поведінкового репертуару різняться лише одним — наявністю чи відсутністю поведінкової групи “Компрімі на роботі”.

Другий етап: розроблення та коригування емпіричних індикаторів відповідно до використовуваної теорії

Спершу на підставі цієї інформації для кожної з указаних стратегій був сформульований набір індикаторів (див. табл. 2). Усі індикатори оцінювалися за таким запитанням: “*Студенти використовують різноманітні способи поведінки в різних ситуаціях, що стосуються навчання. Далі описано деякі із цих способів. Скажіть, як часто Ви особисто використовуєте кожен із них, із застосуванням такої шкали: (1) майже ніколи не використовую — (2) використовую доволі рідко — (3) іноді використовую — (4) використовую дуже часто — (5) майже завжди використовую*”.

Також під час пілотного опитування ($N = 16$) було використано запитання щодо відвідування занять, частоти підготовки до практичних занять та успішності складання іспитів останньої сесії. Пілотаж показав, що такий спосіб визначення типу студента є незадовільним.

По-перше, отримані результати погано співвідносилися з контрольними запитаннями щодо частоти відвідувань занять, інтенсивності підготовки й успішності сесії. Головною причиною цього виявилось те, що індикатори, від початку призначені для вимірювання конкретної стратегії, могли трактуватися учасниками дослідження зовсім інакше і, таким чином, не відповідали закладеному у них змісту.

По-друге, результати, отримані на підставі індикаторів, що описують одну спільну стратегію, мали низьку внутрішню узгодженість. Так, найбільше значення α Кронбаха, рівне 0,44, було зафіксовано для індикаторів, що описують стратегію нероби. Для стратегій спритника і трудяги це значення дорівнювало 0,11 і 0,30 відповідно.

Таблиця 2

**Набір індикаторів для вимірювання увиразненості
поведінкових стратегій**

| Назва стратегії | Зміст індикаторів |
|---------------------|--|
| Стратегія нероби | Велика кількість боргів, накопичених упродовж семестру, не є для мене серйозною проблемою, адже з ними можна впоратися й перед самою сесією |
| | Я здаю домашні завдання із відставанням від переважної частини групи. Це дає мені можливість звірити правильність виконання роботи і виправити помилки |
| | Свої пропуски я пояснюю особистими обставинами, що завадили бути присутнім на парі |
| | Для складання багатьох заліків/іспитів мені зазвичай достатньо ознайомитися з матеріалами конспекту за день ба навіть у день складання |
| | Щоб отримати потрібну оцінку, я іноді просто чемно прошу про це викладача |
| | Коли я пояснюю свою відсутність на заняттях викладачеві, то використовую тему здоров'я чи сімейних обставин |
| Стратегія спритника | Коли мені випадає нагода, я демонструю навчальну активність на парі |
| | Я намагаюся знайти правильний підхід до викладача, зважати на його індивідуальні особливості |
| | У разі боргів, складаючи іспит/залік, я намагаюся перевести обговорення на матеріал, який знаю |
| | Я витрачаю більше часу для підготовки до предметів більш вимогливих викладачів |
| | Я переважно відвідую ті види занять (лекції або практичні), які важливіші для складання іспиту з предмета на сесії. Решту занять відвідую не завжди |
| | Якщо викладач задає певний мінімум необхідних завдань на семестр, для мене достатньо виконати тільки їх |
| Стратегія трудяги | Підготовку домашніх завдань ми (я та інші студенти моєї групи) розподіляємо між собою |
| | Я намагаюся планувати виконання моїх навчальних обов'язків |
| | У навчальних цілях я вдаюся до співпраці з моїми одногрупниками |
| | Моя поведінка на парах є правильною і старанною |
| | Я намагаюся схопити суть навчального матеріалу і добре розібратися з ним |
| | За необхідності я використовую весь вільний час для підготовки до навчання |

По-третє, специфіка такого вимірювання не відповідає особливостям наявної теорії. Річ у тім, що згідно з теорією, розробленою на основі вивчен-

ня феномену навчання за умов обмеженого часу, поведінкові стратегії є свого роду континуумом, що розпочинається стратегією нероби, продовжується стратегією спритника і завершується стратегією трудяги. Відтак, поведінковий репертуар студента щонайліпше можна схарактеризувати віднесенням його до двох суміжних стратегій (*нероба і спритник, спритник і трудяга*), як це описано вище. У процесі дослідження за стратегією “обґрунтованої теорії” не трапилося жодного випадку, коли студент використовував усі три стратегії однаковою мірою або ж коли поряд з основною стратегією він мав дві додаткові як рівноцінні. У ситуації вимірювання з використанням набору індикаторів такі випадки траплялися. Гіпотетично можливі навіть такі ситуації, коли результати вимірювання набувають максимального значення для кожної зі стратегій, незважаючи на те, що такі випадки не фіксуються в емпіричній дійсності в процесі спостереження.

На мій погляд, усе це є свідченням того, що такий спосіб вимірювання багато в чому є артефактним і, відповідно, незастосовним для вивчення поведінкових стратегій студентів. Виходячи з цього, було розроблено інший спосіб вимірювання на основі одного комплексного запитання, зміст якого також визначався описаними вище поведінковими стратегіями, тобто студентів безпосередньо запитували про те, яку поведінкову стратегію вони використовують, хоча, власне, йшлося саме про тип студента.

Припустімо, що назагал студентів можна розподілити на три групи, залежно від зусиль, що їх докладають у процесі навчання:

- А) найменш старанні студенти, які в основному не докладають зусиль для навчання які сподіваються на вдалий збіг обставин і на те, що всі борги можна буде закрити наприкінці семестру, а іспити і заліки скласти завдяки допомозі одногрупників ба навіть практично нічого не роблячи;
- В) які балансують між підготовкою найважливіших завдань та ігноруванням завдань, що не мають принципового значення, які використовують ранжування викладачів за суворістю та вимогливістю з метою відповідної концентрації зусиль і зазвичай відвідують ті пари, що мають більшу вагу в рамках дисципліни;
- С) найбільш старанні, які ставляться до навчання максимально відповідально, які намагаються готувати всі завдання, передбачені навчальними вимогами, які намагаються здобути максимум знань і постійно відвідують пари.

А тепер, використовуючи наведену далі шкалу, ВКАЖІТЬ МІСЦЕ, НА ЯКЕ Б ВИ ПОСТАВИЛИ СЕБЕ, порівнюючи з цими трьома групами. При цьому не можна обирати “чисту” групу (А, В або С), необхідно обрати ОДНЕ зі значень між двома “чистими” групами (між А і В або між В і С).

| | | | | | | | | | | | | |
|----------|-------|-------|-------|-------|-------|----------|-------|-------|-------|-------|-------|----------|
| А | A-90% | A-70% | A-50% | A-30% | A-10% | В | B-90% | B-70% | B-50% | B-30% | B-10% | С |
| | B-10% | B-30% | B-50% | B-70% | B-90% | | C-10% | C-30% | C-50% | C-70% | C-90% | |

Третій етап: перевірка конструктивної валідності й ретестової надійності

Мінімальна перевірка конструктивної валідності пов'язана із забезпеченням двох її складових — конвергентної та дискримінантної валідності [Campbell, 1959]. Перша ґрунтується на тому, що результати шкали демонструють зв'язок із даними, отриманими за допомогою методик, призначених для вимірювання концептуально близьких соціальних феноменів. Друга, навпаки, ґрунтується на тому, що результати не пов'язані з даними методик, призначених для вивчення концептуально відмінних феноменів.

Головне побоювання, пов'язане із використанням поданої вище 10-бальної шкали, стосувалося можливості отримання соціально бажаних відповідей, тобто запитання могло виявитися сензитивним і зсувати відповіді в бік завищення оцінок. Для перевірки цього припущення до анкети поряд з іншими запитаннями було включено шкалу обману з опитувальника ММРІ.

Після вдалого пілотного опитування ($N = 24$) шкала була перевірена на двох вибірках різних ВНЗ ($N_1 = 61$, $N_2 = 62$). Згодом, у разі адекватного відображення шкалою теоретичного конструкту результати її застосування мали корелювати із відповідями на запитання про кількість виступів на практичних заняттях і про успішність останньої сесії, негативно корелювати із відповідями на запитання про кількість невідвідувань і не мати кореляції зі шкалою обману. Для відповідної перевірки було використано ранговий коефіцієнт кореляції Спірмена (див. табл. 3).

Таблиця 3

Сила і напрям зв'язку шкали із контрольними змінними

| Контрольна змінна | Сила зв'язку | | |
|---|----------------------------|----------------------------|------------------------|
| | Перший ВНЗ ($N = 61$) | Другий ВНЗ ($N = 62$) | Разом ($N = 123$) |
| Кількість виступів на практичних заняттях | 0,74** | 0,69** | 0,71** |
| Успішність складання останньої сесії | 0,67** | 0,67** | 0,67** |
| Кількість пропусків | -0,56** | -0,64** | -0,59** |
| Шкала обману | -0,16 | 0,20 ¹ | -0,09 ² |

** результати статистично значимі на рівні 1%.

Виходячи з отриманих результатів, було вирішено, що ця шкала відображає теоретичний конструкт, який описує поведінкові стратегії студентів, а також що вона слабо піддається зсувам, пов'язаним із соціально бажаними відповідями.

Разом із тим наведені дані стосуються перевірки шкали, першою чергою на рівні самозвіту, який багато в чому є суб'єктивним. Тому було здійснено додаткову перевірку її із залученням експертів.

¹ Кількість осіб у другому ВНЗ, опитаних за Шкалою обману, дорівнює 17. Надалі цей взаємозв'язок перевіряли на великих вибірках студентів цього ВНЗ, але статистично значимого зв'язку виявлено не було ($N = 41$, $r_s = -0,03$).

² $N = 76$.

До складу останніх увійшли по два викладачі та одному студенту на кожній із кафедр. У кожному ВНЗ оцінювали всіх студентів однієї навчальної групи ($N_1 = 20$, $N_2 = 24$). Викладачі, відібрані як експерти, були добре знайомі з навчальними групами, а студенти-експерти навчалися у складі цих груп і працювали на цих кафедрах. Оцінювання проводили за 10-бальною шкалою комплексного запитання. Крім цього для кожного зі студентів було розраховано середнє значення оцінок на сесіях, отриманих ними за всі роки навчання¹. Для визначення сили зв'язку був використаний ранговий коефіцієнт кореляції Спірмена. Оскільки оцінки експертів високою мірою узгоджені (для першого ВНЗ значення (α Кронбаха дорівнює 0,89, для другого — 0,96), додатково було розраховано середнє значення оцінок експертів для кожної кафедри та обчислені відповідні коефіцієнти кореляції (див. табл. 4).

Таблиця 4

Сила і напрям зв'язку результатів експертного оцінювання із середніми значеннями оцінок, отриманих на всіх сесіях

| Експерти | Сила зв'язку | |
|-------------------------------|--------------|------------|
| | Перший ВНЗ | Другий ВНЗ |
| Перший викладач | 0,83** | 0,93** |
| Другий викладач | 0,72** | 0,91** |
| Студент | 0,80** | 0,93** |
| Середнє значення трьох оцінок | 0,85** | 0,95** |

** результати статистично значимі на рівні 1%.

Наведені дані ґрунтуються на точнішій інформації, ніж самозвіти. По-перше, це пов'язане з узгодженістю результатів, отриманих у перебігу оцінювання різними експертами. По-друге, це забезпечується відсутністю втрати інформації стосовно успішності навчання, оскільки враховано всі роки. Отже, результати експертної перевірки цілковито підтверджують сформульовані раніше висновки.

Після перевірки конструктивної валідності у трьох навчальних групах було проведено опитування з метою перевірки на ретестову надійність. Узгодженість двох опитувань, оцінена із застосуванням коефіцієнта кореляції Спірмена, на рівні 0,75 ($N = 37$, ймовірність помилки — 1%).

Четвертий етап: специфіка інтерпретації отриманих результатів

На особливу увагу заслуговує специфіка інтерпретації результатів, отриманих у разі застосування цієї шкали. Річ у тім, що за змістом шкала ґрунтувалася на інформації про поведінкові стратегії студентів, а кінцевою метою її застосування було отримання інформації про належність кожного студента до одного із трьох типів — “байдужого”, “декларативно зацікавленого” і “відповідального”. Виходячи з цього, інтерпретація шкали здійсню-

¹ Були використані дані щодо успішності складення сесій перших 4-х років навчання.

валася суто на підставі теоретичних положень, що пов'язують поведінкові стратегії з типами студентів (див. табл. 5).

Таблиця 5

Співвідношення значень шкали із типами студентів

| Тип студента | Поведінкові стратегії | Значення шкали |
|--------------------------------|---|----------------|
| 1. “Байдужий” | 1.1. Стратегія нероби – основна, стратегія спритника – додаткова | 1, 2 |
| | 1.2. Стратегія спритника – основна, стратегія нероби – додаткова або ж рівноцінна | 3, 4, 5 |
| 2. “Декларативно зацікавлений” | 2.1. Стратегія спритника – основна, стратегія трудяги додаткова | 6, 7 |
| 3. “Відповідальний” | 3.1. Стратегії спритника і трудяги використовуються рівною мірою | 8 |
| | 3.2. Стратегія трудяги – основна, стратегія спритника – додаткова | 9, 10 |

Отже, студенти, котрі обрали п'ять перших градацій шкали, були віднесені до “байдужого” типу, ті, що обрали шосту і сьому градації, – до “декларативно зацікавленого”, ті, які обрали останні три – до “відповідального”.

Для перевірки правильності такої інтерпретації її результати були співвіднесені з результатами відповідей щодо успішності складання останньої сесії (див. табл. 6, 7), попередньо перетворених у три градації: 1) під час складання сесії були трійки; 2) сесію складено на чотири-п'ять; 3) сесію складено на відмінно.

Для кожної групи студентів було обчислено ранговий коефіцієнт кореляції γ . Для обох ВНЗ він дорівнює 0,83 і статистично значимий на рівні 1%.

Таблиця 6

Зв'язок типів студентів з успішністю складання сесії (перший ВНЗ)

| Результати складання останньої сесії | Тип студента | | | Разом |
|--------------------------------------|--------------|-----------------------------|------------------|-------|
| | “Байдужий” | “Декларативно зацікавлений” | “Відповідальний” | |
| Були трійки | 15 | 6 | 1 | 22 |
| Чотири-п'ять | 4 | 10 | 2 | 16 |
| На відмінно | 1 | 6 | 14 | 21 |
| Разом | 20 | 22 | 17 | 59 |

Таблиця 7

Зв'язок типів студентів з успішністю складання сесії (другий ВНЗ)

| Результати складання останньої сесії | Тип студента | | | Разом |
|--------------------------------------|--------------|-----------------------------|------------------|-------|
| | “Байдужий” | “Декларативно зацікавлений” | “Відповідальний” | |
| Були трійки | 21 | 4 | 3 | 28 |
| Чотири-п'ять | 5 | 11 | 8 | 24 |
| На відмінно | 0 | 1 | 9 | 10 |
| Разом | 26 | 16 | 20 | 62 |

Попри те, що сила і напрям зв'язку однакові для обох ВНЗ, його характер істотно різниться. Проте в рамках цієї статті це питання не розглядається, оскільки пов'язане не зі специфікою самої шкали, а з особливостями навчального закладу. Це питання розглянуто в моїй публікації [Дембіцький, 2010с]. Назагал ці результати підтверджують правомірність використання запропонованої інтерпретації, хоч і вказують на необхідність її специфікації стосовно вищих навчальних закладів різних типів, що, своєю чергою, є важливою умовою досягнення композиційної валідності шкали.

Разом із тим у контексті групування пунктів шкали цілком логічним може виглядати сумнів стосовно пріоритетності саме такого варіанта групування. Найочевиднішими альтернативами є такі подані далі варіанти розподілу шкали, які б поділяли її на приблизно рівні групи.

Перший варіант (3/4/3):

| 1 група / "байдужий" тип | | | 2 група / "декларативно зацікавлений" тип | | | | 3 група / "відповідальний" тип | | |
|--------------------------|-------|-------|---|-------|-------|-------|--------------------------------|-------|-------|
| A-90% | A-70% | A-50% | A-30% | A-10% | B-90% | B-70% | B-50% | B-30% | B-10% |
| B-10% | B-30% | B-50% | B-70% | B-90% | C-10% | C-30% | C-50% | C-70% | C-90% |

Другий варіант (3/3/4):

| 1 група / "байдужий" тип | | | 2 група / "декларативно зацікавлений" тип | | | 3 група / "відповідальний" тип | | | |
|--------------------------|-------|-------|---|-------|-------|--------------------------------|-------|-------|-------|
| A-90% | A-70% | A-50% | A-30% | A-10% | B-90% | B-70% | B-50% | B-30% | B-10% |
| B-10% | B-30% | B-50% | B-70% | B-90% | C-10% | C-30% | C-50% | C-70% | C-90% |

Третій варіант (4/3/3):

| 1 група / "байдужий" тип | | | | 2 група / "декларативно зацікавлений" тип | | | 3 група / "відповідальний" тип | | |
|--------------------------|-------|-------|-------|---|-------|-------|--------------------------------|-------|-------|
| A-90% | A-70% | A-50% | A-30% | A-10% | B-90% | B-70% | B-50% | B-30% | B-10% |
| B-10% | B-30% | B-50% | B-70% | B-90% | C-10% | C-30% | C-50% | C-70% | C-90% |

Якщо ці варіанти групування є вдалішими в плані відображення типу студента, то вони мають демонструвати вищі порівняно з ініціальним групуванням (5/2/3) коефіцієнти кореляції (γ і Спірмена) з успішністю навчання студентів (див. табл. 8).

Таблиця 8

Величина коефіцієнта γ для різних варіантів групування пунктів шкали

| Варіант групування шкали | Сила зв'язку для першого ВНЗ | Сила зв'язку для другого ВНЗ | Сила зв'язку для обох ВНЗ* |
|--------------------------|------------------------------|------------------------------|----------------------------|
| 5/2/3 | 0,83** | 0,83** | 0,81** |
| 3/4/3 | 0,84** | 0,80** | 0,79** |
| 3/3/4 | 0,82** | 0,89** | 0,83** |
| 4/3/3 | 0,81** | 0,78** | 0,77** |

* масиви об'єднано;

** результати статистично значимі на рівні 1%.

Як можна бачити з наведених даних, найефективнішим варіантом групування є варіант 3/3/4, а вже потім 5/2/3. Проте якщо порівняти розподіли таблиць зіставності для двох варіантів групування, то матимемо протилежну картину (див. табл. 9, 10). В обох таблицях наводиться інформація одразу про два ВНЗ.

Таблиця 9

Зв'язок типів студентів з успішністю складання сесії (групування 3/3/4)

| Результати складання останньої сесії | Тип студента | | | Разом |
|--------------------------------------|--------------|-----------------------------|------------------|-------|
| | “Байдужий” | “Декларативно зацікавлений” | “Відповідальний” | |
| Були трійки | 16 | 29 | 5 | 50 |
| Чотири-п'ять | 1 | 20 | 19 | 40 |
| На відмінно | 1 | 3 | 27 | 31 |
| Разом | 18 | 52 | 51 | 121 |

Таблиця 10

Зв'язок типів студентів з успішністю складання сесії (групування 5/2/3)

| Результати складання останньої сесії | Тип студента | | | Разом |
|--------------------------------------|--------------|-----------------------------|------------------|-------|
| | “Байдужий” | “Декларативно зацікавлений” | “Відповідальний” | |
| Були трійки | 36 | 10 | 4 | 28 |
| Чотири-п'ять | 9 | 21 | 10 | 24 |
| На відмінно | 1 | 7 | 23 | 10 |
| Разом | 46 | 38 | 37 | 121 |

Якщо уявити функціональний зв'язок цієї залежності, логічним буде припустити, що всі байдужі студенти мали би трійки, декларативно зацікавлені навчалися би на чотири-п'ять, а відповідальні — на відмінно. Виходячи з такого критерію, можна оцінити, наскільки кожен варіант групування відхиляється від цієї ідеальної моделі. У разі варіанта 3/3/4 відхилення спостерігається у 48% випадків (58 із 121), а за варіанта 5/2/3 — у 34% випадків (41 із 121)¹. У таблиці, побудованій для варіанта 3/3/4, видно, що значна частина студентів, які мають трійки, потрапили в декларативно зацікавлену групу, що не відповідає поданій вище теорії. Власне, саме останній аспект, а не сила установлених взаємозв'язків, має вирішальне значення для оцінювання різних варіантів групування.

Назагал вищі коефіцієнти кореляції для варіанта 3/3/4 пояснюються у цьому випадку насамперед особливостями обчислення самого показника γ_i , відповідно, несуть радше технічне, ніж смислове навантаження².

¹ Для варіантів групування 3/4/3 і 4/3/3 величина відхилення становить 43% і 42% відповідно.

² Для детального ознайомлення з особливостями розрахунку коефіцієнта γ_i див., напр.: [Хили, 2005: с. 416–420].

Утім, якщо для оцінювання сили зв'язку використовувати коефіцієнт кореляції Спірмена, то групування 5/2/3 буде доцільнішим за інші варіанти (див. табл. 11).

Таблиця 11

**Величина кореляції Спірмена
для різних варіантів групування пунктів шкали**

| Варіант групування шкали | Сила зв'язку для першого ВНЗ | Сила зв'язку для другого ВНЗ | Сила зв'язку для обох ВНЗ* |
|--------------------------|------------------------------|------------------------------|----------------------------|
| 5/2/3 | 0,65** | 0,68** | 0,66** |
| 3/4/3 | 0,57** | 0,62** | 0,58** |
| 3/3/4 | 0,66** | 0,64** | 0,64** |
| 4/3/3 | 0,57** | 0,64** | 0,59** |

* масиви об'єднано;

** результати статистично значимі на рівні 1%.

Висновки

При формулюванні висновків необхідно сказати як про методологічні аспекти теоретичної валідизації, так і про методичні аспекти використання комплексної 10-бальної шкали, розглянутої у статті.

У методологічному плані можна дати дві головні рекомендації стосовно теоретичної валідизації вимірювальних шкал:

1. Перш ніж переходити до соціологічного вимірювання, треба розробити теорію, що описує конструкти тих соціальних феноменів, що будуть об'єктом вимірювання. При цьому розроблена теорія має бути валідною (тобто відповідати емпіричній дійсності, що лежить у фокусі дослідження) уже до вимірювання. У цьому сенсі теорія є первинною, а вимірювання — вторинним, похідним від неї. Вимірювальна шкала розробляється, модифікується й оцінюється в контексті наявної теорії. Завдяки цьому досягається змістова і логічна валідність вимірювальної шкали, які слугують базою для досягнення конструктної валідності.

2. Після того як шкала пройшла перевірку на конструктну валідність, дослідник має здійснити вимірювання щодо тих соціальних груп, для вивчення яких придатна ця шкала. Це необхідно переважно для того, щоб мати уявлення про засади інтерпретації результатів, отримуваних у разі її використання. Тут, знову ж таки, неможливо обійтися без валідної теорії, що вписує конструкт, покладений в основу вимірювання, у ширше концептуальне поле. У цьому випадку вже вимірювальну шкалу (а точніше, її результати) використовують для збагачення тієї теорії, на підставі якої вона була побудована.

У методичному плані можна рекомендувати таке¹:

¹ Наведені далі рекомендації ґрунтуються на досвіді, отриманому в процесі збирання інформації, не відображеному в цій статті. Ця інформація доволі звична у плані проведення прикладних досліджень, але конче важлива для тих, хто відчує необхідність використання шкали у своєму дослідженні.

1. Не включати шкалу в анкети, призначені для самозаповнення. Слова Сеймура Садмена про те, що респонденти є когнітивними покупцями, на всі 100% вочевидноються в цьому випадку. Часто студенти навіть не дочитують інструкцію до кінця і позначають один із описаних у запитанні типів, навіть після прохання уважно прочитати інструкції запитання до самого кінця. Тому бажано використовувати відповідну анкету в разі face-to-face інтерв'ю. Якщо ж проводити групове інтерв'ю (наприклад, з навчальною групою), необхідно докладно пояснювати специфіку запитання, а також перевіряти результати відповідей, щоб мати можливість внести поправки у разі неправильного заповнення.

2. У комплексних освітніх дослідженнях, котрі передбачають формування бази даних щодо кожного студента, слід опитувати викладачів, а не студентів. У цьому разі отримувана інформація про студентів має більш точний характер.

3. Слід дуже акуратно ставитися до продемонстрованих у цій статті взаємозв'язків. Річ у тім, що специфіка освітньої установи може вносити в них суттєві корективи. Останнє має конче важливе значення з точки зору інтерпретації отриманих результатів. Найліпший спосіб уникнути хибних висновків — триангуляція методів (зокрема — включене спостереження). З цих міркувань описана вище шкала найбільш придатна в дослідженнях змішаного типу¹, що передбачають отримання даних із різних джерел.

Джерела

Дембицкий С. Исследования смешанного типа в социологии / Сергей Дембицкий // Социальные измерения общества. Выпуск 2 (13). — 2010b. — С.128–139.

Дембицкий С. “Обоснованная теория”: стратегия сбора и анализа качественных данных при теоретической валидации / Сергей Дембицкий // Социология: теория, методы, маркетинг. — 2010a. — № 2. — С. 64–83.

Дембицкий С. Применение стратегии согласования концептов в теоретической валидации (на примере исследования поведенческих практик студентов) / Сергей Дембицкий // Социология: теория, методы, маркетинг. — 2009. — № 4. — С. 99–114.

Дембицкий С. Теоретическая валидация на различных уровнях социологического исследования / Сергей Дембицкий // Социология: теория, методы, маркетинг. — 2010c. — № 4. — С. 152–178.

Дембицкий С. Теоретическая валидность измерительной процедуры и смещение данных в социологическом исследовании / Сергей Дембицкий // Социология: теория, методы, маркетинг. — 2008. — № 3. — С. 99–118.

Хили Дж. Статистика. Социологические и маркетинговые исследования / Джозеф Хили. — Киев: ООО “ДиаСофтЮП”; СПб.: Питер, 2005. — 638 с.

Campbell D. Convergent and Diskriminant Validation by the Multitrait-multimethod Matrix / Donald Campbell, Donald Fiske // Psychological Bulletin. — 1959. — № 2. — P. 81–105.

¹ Докладніше про дослідження змішаного типу див., напр.: [Дембицкий, 2010b].