

СЕМАНТИКА ПОНЯТЬ ПРОЦЕДУРИ ЗМІСТОВОЇ ВАЛІДИЗАЦІЇ ДИДАКТИЧНОГО ТЕСТУ

словник: У 2 т. Т.2 / За ред. С.В. Мочерного. – Львів: Світ, 2005. – 616 с.

6. Ничкало Н.Г. Професійна освіта і навчання: проблеми взаємозв'язку з ринком праці // Формування широкої кваліфікації робітників. Вклад ПТО у розвиток трудового потенціалу ХХІ століття: Зб. матер. підгот. у рамках реалізації укр.-нім. проекту

“Підтримка реформи професійно-технічної освіти в Україні” / Ю. Вайс, Н. Ничкало, А. Сімак та ін. – Ніжин: ТОВ “Видавництво “Аспект-Поліграф”, 2007. – С. 134 – 148.

7. Психологічна енциклопедія / Автор-упорядник О.М. Степанов – К. “Академвидав”, 2006. – 424 с.

8. Соціологічна енциклопедія / Укладач В.Г. Городяненко. – К.: Академвидав, 2008. – 456 с.

Стаття надійшла до редакції 11.08.2009

УДК 348.147

Лілія Морська, доктор педагогічних наук, завідувач кафедри англійської філології
Тернопільського національного педагогічного університету
імені Володимира Гнатюка

СЕМАНТИКА ПОНЯТЬ ПРОЦЕДУРИ ЗМІСТОВОЇ ВАЛІДИЗАЦІЇ ДИДАКТИЧНОГО ТЕСТУ

У статті здійснено аналіз семантики понять процедури змістової валідизації дидактичного тесту та обґрунтовано її складові компоненти, а також охарактеризовано змістове наповнення критерійної валідності як того поняття, що сприяє розумінню феномену змістової валідизації.

Ключові слова: валідність, дидактичні тести, критерії.

Постановка проблеми. Дидактичні тести стали активно використовуватися у практиці української освіти. Вони виконують роль одночасно інструменту прикладної діагностики для внесення корективів у педагогічний процес та засобу поповнення фактологічної бази педагогічної теорії й окремих методик викладання предметів регулярною інформацією про реальні досягнення в опануванні освітніх стандартів, про розвиток тих, хто навчається, про недоліки та труднощі у навчанні тощо. Отже, такі тести є важливим складовим компонентом освітнього процесу загалом. І тому на часі є проблема забезпечення якості дидактичних тестів, вимоги до яких постійно зростають.

Дидактичні тести створюються у багатьох навчальних закладах для внутрішніх потреб. Їх, як правило, не стандартизують, не валідизують, не перевіряють на надійність, бо, зазвичай, не бачать у цьому необхідності або не вміють робити. Проте, зауважуємо, що така необхідність є, бо вважаємо, що доцільно працювати з діагностичним інструментом, цінність якого доведена, аніж використовувати тест, в результатах якого немає впевненості.

У зв'язку з відносно молодим (принаймні в Україні) віком досліджень з педагогічної діагностики в наукових публікаціях з'являється велика кількість нових понять і термінів (в тому

числі перекладених із арсеналу західної тестології). В дидактичному тестуванні, що, зазвичай, супроводжується ідеями психодіагностики, особливо на етапі його експериментальної адаптації або перевірки методики, необхідно якомога коректніше користуватися психолого-педагогічною термінологією. Не дивлячись на те, що стосовно проблематики валідизації тестів терміносистема у психології розроблена значно детальніше, ніж у загальній педагогіці, це не дає підстави для її перенесення у дидактичну тестологію у повному обсязі без критичного аналізу [6, 38].

Тому метою статті є визначити загальні та специфічні елементи процесуального аспекту валідизації, проаналізувати загальні риси концепції валідизації в дидактичній тестології з точки зору становлення її терміносистеми. Результатом такого аналізу стане аргументоване зіставлення двох паралелей термінології валідизації – тестологічної та експериментальної, що буде основою для визначення їх співвідношення, спільного та відмінного, доцільної сфери та умов застосування.

Огляд останніх досліджень та публікацій.

Хоч термін “валідність” відомий для європейської науки вже давно, в Україні (а ще раніше в колишньому Радянському Союзі) за останні 25 років він вживався дуже рідко у психологічних, соціологічних і тим більше у педагогічних

СЕМАНТИКА ПОНЯТЬ ПРОЦЕДУРИ ЗМІСТОВОЇ ВАЛІДИЗАЦІЇ ДИДАКТИЧНОГО ТЕСТУ

дослідженнях. Лише в публікаціях останніх років педагоги з більшою сміливістю стали вводити це поняття, проте не зовсім розуміючи його основний зміст і сферу використання.

Характеристика валідності подається у провідних виданнях довідкового змісту. У словнику “Психологія” (під загальною редакцією О.В. Петровського і М.Г. Ярошевського) це поняття визначається як один із основних критеріїв якості тесту (від англ. valid – придатний). Тут також зазначається, що проблема валідності виникає у процесі розробки та практичного застосування тесту, коли ставиться завдання встановити відповідність між ступенем вираженості якості особистості і методом її вимірювання. Чим валіднішим є тест, тим краще відображається в ньому ця якість (чи властивість), заради вимірювання якої він створювався [9, 47].

Німецький учений Г. Вітцлак характеризує валідність методу як відповідність того, що вимірює психодіагностичний метод до того, що він призначений вимірювати. На думку згаданого автора, визначенням валідності встановлюється спектр дійсності, для якої метод подає репрезентативні дані [8, 93].

Не дивлячись на те, що термін “валідність” вважають загальнонауковим, ми не знайшли його висвітлення у фундаментальних загальнокультурних словниках. Немає цього терміна в “Педагогічному енциклопедичному словнику” (2002 р.).

У “Великому тлумачному словнику сучасної української мови” валідність трактується як обґрунтованість й адекватність дослідницьких інструментів (операціоналізованих понять, вимірювальних операцій та експериментів) [3, 73].

На початку 80-х років минулого століття соціолог і фахівець з педагогічного вимірювання В.С. Аванесов під валідністю тесту розумів його здатність вимірювати те, для чого він був задуманий. На думку вченого, це поняття стосується не лише самого тесту, але й мети, ситуації та процедури його застосування [1, 136].

Позитивним є той факт, що рівень ознайомленості педагогічної громадськості з поняттями та ідеями валідації за останні 5 – 7 років суттєво змінився на краще, проте, не лише в педагогіці, але й у суміжних соціальних науках ситуація з оцінювання якості тестів ускладнилася заплутаністю термінології, а відтак залишається лише пріоритетом теорії, а практика продовжує дотримуватися традиційних спрощених процедур створення та використання педагогічних тестів.

У зв’язку з цим виникає питання про

семантичне насичення терміну “валідність”, тобто про його основні складові.

Нами досліджено, що вчені пов’язують змістову валідність тесту не лише з аналізом його відповідності до навчальних планів і програм (у західній педагогіці цей компонент називають курикулярною валідністю), але й визначення конструктивної та концептуальної валідності тесту.

Так, характеризуючи необхідність і значимість змістової валідації тесту, відомий польський дидакт В. Оконь зазначав, що тест для певного предмету не буде вдалим, якщо за його допомогою визначається лише рівень знань і не враховуються інші критерії, якщо надмірно виділяється лише один розділ програми за рахунок інших розділів, або якщо визначається лише інтелектуальний розвиток, пам’ять, а знання предмету залишається поза увагою тесту [7, 166].

Німецький учений К. Інгенкамп вважає, що жодний з елементів вимірювального інструментарію не є валідним загалом, він може бути таким лише у певному смислі [4, 43]. Ми, у свою чергу, погоджувачись із позицією відомого теоретика сучасної педагогічної діагностики, вважаємо, що всі ознаки валідності повинні бути конкретними, стосуватися певних критеріїв, аспектів методики діагностування.

Систематизація видів валідності, як методологічна проблема, є одним із головних питань у дослідженнях західних психологів і педагогів. Наприклад, американський учений А. Анастасіє виділяє три підходи до визначення валідності: валідність за змістом, валідність за критерієм, конструктивна валідність. Як додатковий вид автором згадується очевидна валідність [2].

Відзначаючи існування різних видів валідності, Г. Вітцлак вказує на змістову, конструктивну, прогностичну й педагогічну [8, 133 – 142].

Методолог психологічного експерименту Д. Кемпбел виділяє внутрішню та зовнішню валідність, вживаючи поряд зі згаданими двома видами ще й конструктивну валідність [5].

Виклад основного змісту дослідження. Отже, спектр термінів, що характеризують валідність, досить широкий та різноманітний, проте це не пояснює проблему систематизації понятійного апарату процесу валідації. Здійснивши дослідження типології видів валідності, ми прийшли до висновку про необхідність їх поділу на дві групи понять, що стосуються змістової та критерійної валідації. Слід враховувати те, що валідність власне дидактичного тесту пов’язана з адекватністю відображення того змісту засвоєного раніше навчального матеріалу, який учень чи студент опанував на момент тестування.

СЕМАНТИКА ПОНЯТЬ ПРОЦЕДУРИ ЗМІСТОВОЇ ВАЛІДИЗАЦІЇ ДИДАКТИЧНОГО ТЕСТУ

На основі здійсненого зіставно-функціонального аналізу згаданих у статті наукових досліджень, вважаємо, що у структуру поняття змістової валідизації входять такі компоненти: конструктна, концептуальна, куррикулярна та психологічна валідності.

Конструктна валідність відображає основні блоки діагностованого змісту навчального матеріалу, який на момент тестування повинен бути засвоєний учнями чи студентами, а також співвідношення цих блоків, ієрархію їх значущості та послідовність діагностування.

Конструктну валідність ще називають передбачуваною валідністю (*assumed validity*), якщо до певної міри підтверджується висунута тестологом гіпотеза про те, що створений тест і його результати корелюватимуть з іншими тестами, які вимірюють аналогічні ознаки (властивості, якості) тестованих (тобто, учнів чи студентів) [8]. Проте передбачувана і конструктна валідності не є аналогами одного і того ж поняття. Перший термін визначає не постійну, а певну тимчасово-ситуативну характеристику тесту на період здійснення методологічного експерименту щодо складання двох тестів.

Відображення змістового аспекту навчальної програми у структурі конструкту тесту визначає його куррикулярну валідність. Згідно цього показника тест може володіти високою, адекватною та неадекватною куррикулярною валідністю. Останній тип не позначає негативну оцінку тесту, оскільки така ситуація є можливою, коли метою тесту є діагностика певних психологічних, ситуативних впливів на якість засвоєного навчального матеріалу або коли існує декілька паралельних дидактичних тестів, створених різними розробниками.

Концептуальна валідність, у нашому розумінні, повинна бути орієнтована на відображення певної дидактичної та методичної концепції бачення навчального предмету автором навчальної програми, навчального підручника чи посібника та розробника тесту. Усі ці три позиції можуть радикально відрізнитися, наприклад, у прихильника традиційного лінійного викладу матеріалу і тих, хто дотримується проблемного, проектного чи інтерактивного методів навчання.

Наведемо інший приклад із галузі навчання іноземних мов: автори сучасної навчальної програми орієнтують вивчення іноземної мови у середній загальноосвітній школі на комунікативний метод навчання, а підручники розробляються за різними авторськими концепціями (наш аналіз дає підстави виділити щонайменше 6 авторських

підходів до розробки підручника з англійської мови для 7-го класу, серед яких: метод читання, граматико-перекладний метод, свідомо-перекладний метод, метод кооперативного навчання, ситуативний підхід до вивчення іноземної мови, інтерактивна технологія та метод проектів). Дидактичні тести, на жаль, як правило, розробляються без урахування концепції авторів підручників, а більше орієнтовані на програмну концепцію предмету.

При розробці чи експертній оцінці завдань дидактичного тесту доцільно ставити питання про те, в якій мірі тест може діагностувати прояв тих психічних і психологічних якостей тестованого, які є об'єктивно необхідними для виконання певних інтелектуальних операцій в ході вирішення тестових завдань. Саме з такої позиції розуміємо суть психологічної валідності тесту.

З метою розрізнення групи понять змістової валідності вважаємо за необхідне пояснити свою позицію стосовно критерійної валідності тесту.

Г. Вітцлак вважає, що критерійна валідність демонструє те, наскільки за результатами тесту можна робити висновок про той аспект особистості тестованого, який нас цікавить, причому як в теперішньому, так і в майбутньому. Для її визначення результати тесту порівнюють з певними зовнішніми критеріями (наприклад, зі шкільними оцінками) або порівнюють між собою результати різних тестів. В такому випадку мова йде про конкретну валідність критеріїв, яка встановлюється шляхом визначення кореляції між новими тестами й тими, що вже використовувалися [8].

Критерійна валідність у сучасному розумінні передбачає наявність не одного показника, а сукупності експериментально одержаних показників різних аспектів функціонування тесту як вимірювального інструменту. Саме тому інколи зустрічається синонім згаданої валідності у формулюванні “функціональна”.

Одна із найскладніших проблем визначення критерійної валідності інструменту вимірювання – це вибір прийнятного критерію. Яким, наприклад, буде прийнятний критерій навчальних досягнень для перевірки навчальних здібностей? Дати відповідь на це питання непросто. Можна виділити чотири аспекти вибору критерійного виміру. По-перше, критерій повинен бути відповідним. Проте тут перешкодою стає суб'єктивність, оскільки людям властиві різні думки стосовно відповідності критерію. По-друге, критерійний вимір повинен бути неупередженим, що виключає можливість впливу зовнішніх чинників, що можуть надавати вимірюванню

СЕМАНТИКА ПОНЯТЬ ПРОЦЕДУРИ ЗМІСТОВОЇ ВАЛІДИЗАЦІЇ ДИДАКТИЧНОГО ТЕСТУ

характеру тенденційності. Інший аспект означає, що критерій повинен бути надійним, тобто якомога точнішим і чіткішим. По-четверте, необхідно, щоб критерій був доступним, тобто таким, якого можна реально отримати. Як правило, під час створення тестів важко знайти такий критерійний вимір, який би поєднував усі перелічені вимоги.

Загалом у дидактичних тестах побутують умовно прийняті критерії (показник успішності, досягнутий рівень освіти, результат навчання у вигляді елементів спеціальної підготовки), а також конкретні проміжні критерії (міцність сформованості умінь виконувати певні завдання тощо).

У розумінні критерійної валідності використовують конвергентну та дискримінативну валідності (терміни, введені Д. Кемпбелом), де конвергентна валідність є кореляцією тесту з іншими змінними, з якими він теоретично повинен корелювати, а дискримінативна валідність, навпаки, визначає відсутність значної кореляції тесту з тими змінними, з якими він не повинен корелювати [2, 144].

Для дидактичних тестів рубіжного та підсумкового контролю, а також для селективних тестів особливе значення має прогностична та ретроспектна валідності [6, 62]. Прогностична валідність тесту пов'язана з можливістю за його допомогою прогнозувати поведінку тестованого через певний проміжок часу, при цьому особливого значення набуває фактор довготривалого процесу. Ретроспектна валідність визначає валідність самого вимірювання, що проводилося раніше, в тому числі його умов й чіткості дотримання процедур тестування.

Оскільки дидактичне тестування завжди має ознаку корекційності, для нього на практиці важливого значення набуває педагогічна валідність, яка характеризує можливість для педагога за умови однотипних результатів застосування певної методики давати однозначні типові діагнози та рекомендації, що ведуть до успішної корекції навчальних успіхів.

Висновок. Отже, на основі концептуального підходу до аналізу змісту й структури понять, що відображають хід валідації тестів, ми прийшли до висновків: 1) про необхідність чітко розрізняти поняття “змістова”, “концептуальна”, “конструктивна” і “куррикулярна” валідності, виокремлюючи їх як такі, що не співпадають за обсягом охоплення процесу валідації, мають свою специфіку відображення різних аспектів складного процесу діагностування цінності тесту; 2) про те, що у структурі змістової валідації її складовими є: концептуальна, конструктивна,

куррикулярна, психологічна й очевидна валідності, причому останній компонент відображає неправдиві діагностичні ознаки тесту, виступає в ролі так званої фонові валідності, але водночас є значеннєвим для сприйняття користувачами (тестованими) та зацікавленими у процедурі тестування особами.

1. Аванесов В.С. Тесты в социологическом исследовании / В.С. Аванесов. – М.: Наука, 1982. – 200 с.

2. Анастаси А., Урбан С. Психологическое тестирование / А. Анастаси, С. Урбан. – СПб.: Питер, 2004. – 688 с.

3. Великий тлумачний словник сучасної української мови / Гол. ред. В.Т. Бусел. – Київ: Ірпінськ: ВТФ “Перун”, 2004. – 1440 с.

4. Ингенкамп К. Педагогическая диагностика. – М.: Педагогика, 1991. – 240 с.

5. Кемпбел Д. Модели экспериментов в социальной психологии и прикладных исследованиях. – М.: Прогресс, 1980. – 348 с.

6. Нестеренко Н.Д. Система содержательной валидации дидактических тестов в современной педагогике / Наталья Дмитриевна Нестеренко. – Дисс... канд. пед.наук: 13.00.01. – Ростов-на-Дону, 2006. – 206 с.

7. Оконь В. Введение в общую дидактику. – М.: Высшая школа, 1990. – 232 с.

8. Психодиагностика: теория и практика. Пер. с нем. / Под общ. ред. Н.Ф. Талызиной – М.: Прогресс, 1986. – 132 с.

9. Психология: Словарь / Под ред. А.В. Петровского, М.Г. Ярошевского. – 2-е изд. – М.: Политиздат, 1990. – 494 с.

Стаття надійшла до редакції 07.05.2009

a b c d a b c d a b

“Героїзмом можна назвати тільки таке діло, де мука і терпіння одиниці здобуває або окуплює добро цілого народу”.

Іван Франко

*“Правда против сили!
Боем проти зла!
Міжнарод похилий
Вольності слова!”*

Іван Франко

a b c d a b c d a b