

УДК 303.732.4/37.014.6:005.642.4

Камишин В.В., к.т.н.

СИСТЕМНО-ІНФОРМАЦІЙНИЙ АНАЛІЗ ЕФЕКТИВНОСТІ ШКАЛ КВАЛІМЕТРІЇ АКАДЕМІЧНОЇ ОБДАРОВАНОСТІ

Інститут обдарованої дитини
Національної академії педагогічних наук України

Враховуючи важливість забезпечення чіткої кваліметрії академічної обдарованості та застосовуючи методи системного аналізу, проведений порівняльний аналіз найбільш відомих шкал оцінювання знань (4-бальної, 7-бальної (ECTS), 9-бальної стенойнів, 10-бальної стенов, 100-бальної та 200-бальної) за критеріями номінального та кількісного розрізнення тих, хто навчається.

Актуальність проблеми

Ефективність процесу кваліметрії навчально-виховного процесу (НВП), насамперед, рівнів навчальних досягнень (РНД) залежить як від ефективності відповідних систем оцінювання, так і від якості навчальних завдань, що при цьому застосовуються. На сьогоднішній день в закладах освіти України застосовуються 4-бальна (скорочена 5-бальна), 12-бальна, 100-бальна та 200-бальна (під час зовнішнього незалежного оцінювання - ЗНО). Ще більше розмаїття спостерігається у об'єднаній Європі [1; 2]. До того ж, запровадження європейської «полегшеної шкали оцінювання» ECTS, як витікає з недоліків її формування кардинально не вирішило спірні питання взаємного визнання РНД студентів країн-членів Євросоюзу [3].

Міністерство освіти і науки, молоді та спорту (МОНМС) України приділяє суттєву увагу проблемам кваліметрії знань студентів і вимагає від вищих навчальних закладів (ВНЗ) «уведення механізму об'єктивного педагогічного контролю з вивчення результатів діяльності системи вищої освіти і рівнів навчальних досягнень тих, хто навчається». Сюди віднесено також «технологію стандартизованого тестування» [4; 5]. При цьому рекомендується орієнтуватись на вимоги Міжнародної стандартної класифікації занять (ISCO-88), Міжнародної стандартної класифікації освіти (ISCED-97), Міжнародного стандарту якості серії ISO 9000, а також на критерії та стандарти, узгоджені учасниками Болонського про-

цесу. Однак варто наголосити, що відповідні процеси стримуються у вітчизняній освітянській системі відсутністю достатньої кількості наукових результатів, що не лабораторно-фрагментарно, а саме системно дослідили б ці процеси.

Аналіз досліджень і публікацій

Існує безліч визначень поняття «вимірювання», що відрізняються одне від одного залежно від точки зору конкретного дослідника. Провівши всебічний аналіз сутності кваліметрії дидактиці, Н.М. Розенберг сформулював таке [6, 15]: «вимірювання в педагогіці – пізнавальний процес, який складається з того, що на підставі числової системи (або системи класів), отриманої раніше, ізоморфній емпіричній системі з відношеннями, експериментально визначають кількісні значення величин, що характеризують деякі ознаки педагогічних об'єктів або явищ, або вказують на клас, до якого вони відносяться». Скажімо, вимірювання коефіцієнта інтелектуальності (IQ) – це присвоєння чисел характеру зворотної реакції випробуваного на типові тестові завдання.

Процес будь-якого вимірювання має здійснюватися за допомогою певних шкал. Спираючись на дослідження П. Суреса та Дж. Зінеса, можна визначити впорядковану трійку, що складається з емпіричної системи з відношеннями, кількісної системи з відношеннями та функції, яка гомоморфно зображує вихідну систему в іншу [7]. Такого роду перевтілення умовно подано на рис. 1 [3; 8].

На жаль, проблеми кваліметрії РНД

висвітлено в обмеженій кількості сучасних фундаментальних наукових праць [3; 9-11], відповідні процедури з єдиних системних позицій не розглянуто і не узагальнено. Недостатньо використовується багатий досвід психологічної, соціологічної, біологічної кваліметрії [12-15].

Постановка завдання

Таким чином, варто зазначити, що вчені та фахівці з кваліметрії НВП не спираються у своїх дослідженнях одночасно на властивості відомих шкал вимірювання, що застосовуються у будь-якій галузі людської діяльності, у поєднанні з аналізом дидактичних особливостей бальних шкал, що застосовуються для оцінювання РНД студентів (РНДС). Усунення цього недоліку і є *метою* цієї статті.

Системно-кваліметричний підхід до порівняння бальних шкал оцінювання знань

Отже, вимірювання перетворює певні властивості сприйняття людини на так звані «числа». Необхідну інформацію одержують шляхом обробки експериментальних даних. І зрозуміло, що чим точ-

нішою є відповідність між станами та їх позначками, тим більше інформації можна витягти з обчислення даних. Менш очевидно, що ступінь цієї відповідності залежить не тільки від організації вимірювань, а й від природи явища, що досліджується, і що ступінь цієї відповідності, у свою чергу, визначає допустимі (недопустимі) засоби обробки даних.

Уявлення про «шкали вимірювань» створюють певну групу понять. Розглянемо особливості їх застосування у дидактиці, спираючись на [3; 6; 8; 16; 17 та ін.]. Природно, що кожний тип шкали має свою інформативність і свій клас допустимих перетворень, за межі яких не можна виходити без ризику отримати хибні (навіть безглузді) результати. Особливого значення набуває це зауваження в дослідженнях ефективності процесів кваліметрії НВП, що залежать від великої кількості різноманітних чинників, складно взаємопов'язаних та притаманних викладачеві, випробуваному (студентові чи учню), а також навчальному завданню, що виконується під час перевірки.

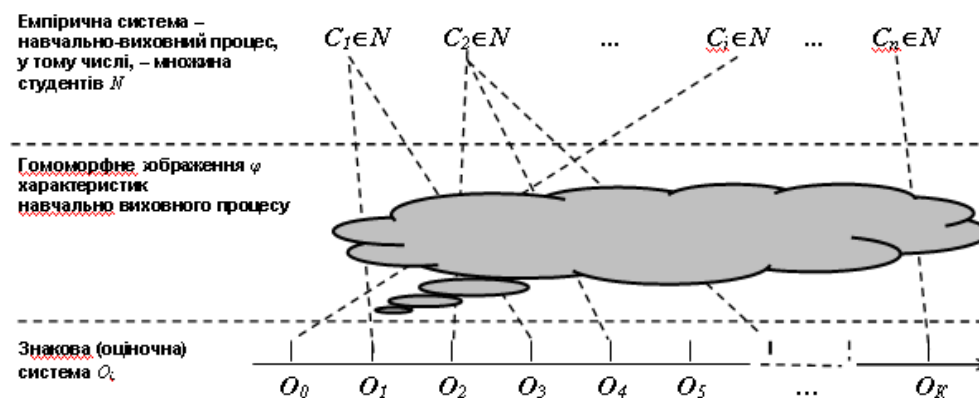


Рис. 1. Загальна ілюстрація кваліметрії навчально-виховного процесу

Нас цікавлять лише такі об'єкти (студенти, учні), про будь-які два стани РНД яких можна сказати, чи розрізняються вони, чи ні, і застосовувати лише такі алгоритми вимірювань, котрі ставлять у відповідність різним станам різні позначення, у той час як станам, що не розрізняються, – однакові позначення. Це означає, що як стани досліджуваних об'єктів, так і їх позначення мають задовольняти такі аксіоми тотожності:

$$\begin{aligned}
 1^0. \hat{A} \hat{=} B, \hat{B} \hat{=} A & \quad \begin{matrix} \text{Б} \\ \text{П} \end{matrix} \\
 2^0. \beta \hat{=} \hat{A} \hat{=} B, \hat{B} \hat{=} A & \quad \begin{matrix} \text{П} \\ \text{Э} \end{matrix}, (1) \\
 3^0. \beta \hat{=} \hat{A} \hat{=} B \text{ i } B \hat{=} C, \hat{B} \hat{=} A \hat{=} C & \quad \begin{matrix} \text{П} \\ \text{П} \\ \text{Ю} \end{matrix}
 \end{aligned}$$

де « $\hat{=}$ » – позначка відношення еквівалентності. І якщо A і B – числа, вона означає рівність.

У вимірюванні будь-яких показників найбільш поширені *метричні* (абсолютні

шкали, інтервальні та відношень), номінальні шкали та шкали впорядкування. Тип шкали враховується завжди, коли йдеться про те, які дії має сенс здійснювати з чисельними значеннями показника, що розглядається. У теорії вимірювань це називають *проблемою адекватності (усвідомлення)* числових тверджень. Її суть полягає в тому, що числове твердження вважають адекватним лише тоді, коли значення його істинності є інваріантним відносно допустимих перетворень шкали. Розглянемо вищеперелічені типи шкал у порядку зростання їх досконалості.

1. *Номінальна (класифікаційна, найменувань)* – це найменш досконала, простіша за змістом, якісна шкала. У ній об'єкти групуються в класи за ідентичною (або майже ідентичною) ознакою. Класам присвоюються певні позначення. Замість позначень класів для ідентифікації часто використовують числа, що слугують поясненням поняття «номінальне вимірювання». Номінальна шкала зберігає відношення еквівалентності та відмінності між об'єктами. Числа використовують тільки для позначення і відокремлення об'єкта: усім об'єктам одного класу присвоюється одне й те саме число чи якась інша спеціальна позначка, а об'єктам різних класів – різні числа (позначки). Це можуть бути бали будь-якої шкали оцінювання знань, літерні позначки, як це зроблено, скажімо, у шкалі ECTS. Однак переваги між об'єктами не встановлюються. Ця шкала – єдина з точністю до взаємно-однозначності відповідності, тому у ній відсутні поняття масштабу і точки початку відліку.

Особлива потреба у класифікації виникає у випадках, коли стани (РНД), що класифікуються, розглядаються на безперервній множині – континуумі знань (рис. 2). Тоді завдання кваліметрії вирішується ефективно, якщо всю множину знань розбити на кінцеву множину підмножин, штучно створюючи тим самим класи еквівалентності, кожний з яких відповідає тій чи іншій оцінці прийнятої шкали. Тоді належність стану до якогось класу знову можна реєструвати у шкалі найменувань. Проте

умовність введених класів (не їх шкальних позначок, а самих класів) рано чи пізно проявиться на практиці.

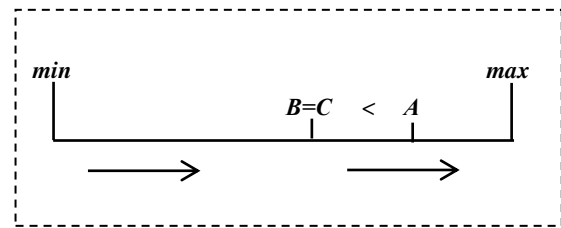


Рис. 2. Уявлення знань як точок континууму

Обробляючи будь-які дані, зафіксовані у номінальній шкалі, безпосередньо із самими даними можна здійснювати тільки операцію перевірки їх збігу чи незбігу. Така операція зображується за допомогою так названого *символу Кронекера*:

$$d_{ij} = \begin{cases} 1 & : x_i = x_j \\ 0 & : x_i \neq x_j \end{cases}, \quad (2)$$

де x_i, x_j – записи різних вимірювань.

За результатами операції зі застосування символу Кронекера можна здійснювати складніші перетворення:

– обчислювати відносні частоти класів, наприклад частота k -го класу:

$$p_k = n_k / n;$$

– рахувати кількість збігів, наприклад, число спостережень k -го класу:

$$n_k = \sum_{j=1}^n d_{k,j}, \text{ де } n - \text{ загальна кількість}$$

спостережень;

– порівнювати ці частоти між собою, визначаючи, наприклад, моду – номер класу, що зустрічається найчастіше:

$$k_{max} = \arg \max_k p_k;$$

– виконувати різноманітні статистичні процедури, суворо спостерігаючи, щоб у них з вихідними даними не здійснювалось нічого, окрім операції перевірки їх на збіг. Із цією метою можна використати критерій χ^2 , інші тести і критерії на відносних частотах, коефіцієнт згоди та ін. [3; 6; 11-13; 16; 17].

Природно, що чим більша розмірність шкали оцінювання, тим більше можливостей має викладач для розрізнення класифікації тих, хто навчається за РНД. Варто обов'язково враховувати психофізіо-

логічні можливості людини (викладача, студента) щодо запам'ятовування класів цього розрізнення, тому що: *по-перше*, відоме так зване *магічне число Міллера* (7 ± 2), що визначає ємність оперативної пам'яті людини та пояснює зручність практичного використання тієї чи іншої шкали [18]; *по-друге*, з психологічної теорії розпізнавання образів відомо, що найдосвідченіші фахівці спроможні ефективно розрізнити класи об'єктів, коли їх загальна кількість не перевищує 11 [19].

Отже, можна зазначити, що усі найбільш відомі бальні шкали оцінювання РНД (4-бальна, 7-бальна (ECTS), 9-бальна, 10-бальна та 12-бальна) знаходяться у відповідних інтервалах. Порівнюючи їх виявлено, що найдосконалішою з них з точки зору розрізнення тих, хто навчається, є 100-бальна шкала (табл. 1). До того ж, встановлено [17], що певним викладачам не вистачає навіть розмірності навіть і цієї шкали, тому вони здійснюють спроби її «покращення», застосовуючи ще й проміжні (!) градації вимірювання знань шляхом введення «дробових» балів. Тому

впровадження для ЗНО 200-бальної шкали є закономірним, однак вона не застосовується в НВП закладів освіти.

2. Шкала впорядкування (ранжирування, рейтингова) – більш досконала та суттєво інформативніша, аніж номінальна, тому що відповідні показники, які нею вимірюються, дають змогу судити про відношення типу «краще – гірше», «більше – менше», що існують між об'єктами, які, власне й властиві людському мисленню [20]. У цій шкалі відсутні поняття масштабу і точки початку вимірювань. Порядкове вимірювання здійснюється, коли можна знайти в об'єктах відмінність ступенів ознаки або властивості. Тоді застосовується властивість «впорядкування» чисел.

Найвідомішою порядковою шкалою є «*ранг у класі школи*» (або дуже розповсюджений на Заході «*ранг випуску в університеті*»). Номери встановлюються від 1 для максимального середнього значення оцінок до n – для мінімального середнього значення оцінок у класі (навчальній групі, потоці) з n студентів (учнів).

Таблиця 1. Порівняльна ефективність шкал оцінювання знань

Зміст завдання вимірювання рівнів навчальних досягнень	Ранжирування шкал оцінювання знань за ефективністю
2	3
Класифікаційне розрізнення тих, хто навчається	200-бальна >> 100-бальна >> 12-бальна > 10-бальна > 9-бальна > 7-бальна > 5-бальна
Рейтингова оцінка тих, хто навчається	200-бальна >> 100-бальна >> 12-бальна > 10-бальна > 7-бальна > 5-бальна
Кількісне оцінювання знань	200-бальна >> 100-бальна >> 12-бальна \approx 10-бальна \approx 9-бальна \approx 7-бальна \approx 5-бальна
Лінгвістична (нечітка) оцінка тих, хто навчається	12-бальна > 10-бальна > 9-бальна > 7-бальна > 5-бальна >> 100-бальна >> 200-бальна
ПРИМІТКИ: > – ознака переваги; «>>» – ознака суттєвої переваги; « \approx » – ознака еквівалентності.	

Порядкові шкали застосовуються для вимірювання в дидактиці РНД, дисциплінованості, ступеня складності навчальних дисциплін (НД) або майбутньої професії. Щодо таких шкал Н.М. Розенберг зазначає [6, 27]: «в педагогічній практиці майже монопольним способом оцінювання засвоєння слугує позначка на порядковій шкалі. Таке широке розповсюдження порядкових шкал у навчально-виховному процесі та

дослідженнях навчання і виховання пояснюється, зокрема, тим, що в багатьох випадках, вивчаючи педагогічні об'єкти і явища, ми маємо кількісної міри, сильнішої, ніж оцінка порівняльної інтенсивності тієї самої ознаки у різних об'єктів». Він також вказує на те, що «тільки такий чутливий, складний і рухливий інструмент, як людина-спостерігач, може забезпечити задовільну інформацію щодо діяльності або харак-

терних особливостей іншої особи чи об'єкта. Людина виступає в таких ситуаціях як своєрідний вимірювальний прилад з усіма своїми індивідуальними особливостями і характеристиками: вибіркоким сприйняттям, пам'яттю, інтелектом і нерідко, на жаль, відсутністю чутливості до того, що важливо в педагогічному або психологічному відношенні...»

Не існує математичного закону або правила, що забороняли б робити арифметичні дії чи здійснювати інші операції над числами, які приписані об'єктам у ході порядкового вимірювання. Тому ще раз підкреслимо, що числа у шкалі впорядкування визначають тільки порядок розташування об'єктів, проте не дозволяють стверджувати «на скільки (у скільки разів) один об'єкт кращий (гірший) від іншого».

Застосування шкали впорядкування сприяє формуванню систем переваг учасників НВП. Під системою переваг буде розумітися будь-яка форма ранжирування досліджуваних характеристик [3; 16; 20]. Найбільшого розповсюдження шкали впорядкування набули в обробці експертної інформації про порівняльні оцінки якісних властивостей об'єктів. Такі оцінки даються у балах, а самі шкали часто називають бальними (ранговими, рейтинговими). Бальним оцінкам завжди притаманний суб'єктивізм, але ж його питома вага може бути різною. З цієї точки зору розрізняють два види бальних оцінок: за наявності загальноприйнятого еталона і за відсутності такого. Клас шкал упорядкування (рангів) з'являється, якщо окрім аксіом 1⁰–3⁰ класи задовольняють такі аксіоми упорядкування:

$$\begin{aligned} 4^0. \beta \hat{e} \hat{u} \hat{i} \quad A > B, \quad \delta \hat{i} \quad B < A & \quad \frac{\beta}{\delta} \\ 5^0. \beta \hat{e} \hat{u} \hat{i} \quad A > B \text{ i } B > C, \quad \delta \hat{i} \quad A > C & \quad \frac{\beta}{\delta} \end{aligned} \quad (3)$$

Позначивши такі класи символами та встановивши між ними ті самі відношення порядку, ми отримуємо *шкалу простого впорядкування*. Скажімо, це призові місця в конкурсі, зокрема шкільних чи студентських олімпіадах. Причому іноді виявляється, що не кожену пару класів можна суворо впорядкувати за перевагами: деякі пари вважаються рівними. Тоді аксі-

оми 4⁰–5⁰ видозмінюються:

$$\begin{aligned} 4'. \beta \hat{e} \hat{u} \hat{i} \quad A \hat{J} B, \quad \alpha \hat{i} \quad A > B & \quad \frac{\beta}{\alpha} \\ 5'. \beta \hat{e} \hat{u} \hat{i} \quad A \hat{i} \hat{A} \hat{z} \hat{A} \hat{i} \hat{N}, \quad \delta \hat{i} \quad A \hat{i} \hat{N} & \quad \frac{\beta}{\delta} \end{aligned} \quad (4)$$

Шкала, що відповідає аксіомам 4'–5', називається *шкалою слабого впорядкування*. Інша ситуація виникає, коли існують пари класів, *непорівнянні* між собою, тобто ані $A \leq B$, ані $B \leq A$ (це відрізняється від умови слабого впорядкування, коли одночасно $A > B$ і $B > A$, тобто $A = B$). Тоді говорять про *шкалу часткового впорядкування*. Такі шкали часто виникають у дослідженнях суб'єктивних переваг.

Характерною особливістю порядкових (у суворому розумінні) шкал є те, що відношення порядку нічого не говорить про дистанції між класами, об'єктами, що порівнюються. Тому порядкові експериментальні дані, навіть якщо вони відображені цифрами, не можна розглядати як числа, з ними не можна виконувати дії, що приводять до отримання різних результатів при перетворенні шкали, не порушуючи порядку. Скажімо, не можна обчислити вибіркове середнє порядкових вимірювань, тобто $\frac{1}{n} \sum_{i=1}^n x_i$, тому що пере-

хід до монотонно перетвореної шкали $x' = f(x)$ при усередненні дасть:

$$\frac{1}{n} \sum_{i=1}^n x'_i \neq \frac{1}{n} \sum_{i=1}^n x_i$$

Проте допустимою є операція, що дозволяє встановити, яке з двох спостережень x_i чи x_j краще, хоча формально цю процедуру ми можемо подати через різницю $x_i - x_j$. Уведемо індикатор позитивних чисел – функцію:

$$C(t) = \begin{cases} 1 & t \geq 0 \\ 0 & t < 0 \end{cases} \quad (5)$$

Якщо $x_i \geq x_j$ та введено цифрову шкалу порядку, то $C(x_i - x_j) = 1$, або $C(x_i - x_j) = 0$, що й дозволяє встановити перевагу x_i перед x_j . Отже, число $R_i = \sum_{j=1}^n C(x_i - x_j)$, де n

– кількість об'єктів, що порівнюються, називається *рангом* i -го об'єкта, що пояснює походження іншої назви шкал упорядкування –

рангові. Якщо наявний слабкий порядок, то частина спостережень збігається (у статистиці така група спостережень називається зв'язкою) і всі члени зв'язки отримують однаковий (старший для них) ранг. Коли це незручно, звертаються або до призначення рангу, середнього для цієї зв'язки, або призначають ранги від меншого до більшого у звичайний спосіб. Отже, під час вимірювання у порядкових (у суворому розумінні) шкалах обробка даних повинна ґрунтуватися лише на операціях, що допустимі для цих шкал – обчислення δ_{ij} і R_i . З цими числами далі вже можна «працювати» довільно: окрім знаходження частот і мод (як для номінальної шкали), з'являється можливість:

- визначити вибірккову медіану, тобто спостереження з рангом R_i , найближчим до числа $n/2$;
- розбити всю вибірку на частини у будь-якій пропорції, знаходячи вибірккові квантили будь-якого рівня p , $0 < p < 1$ (тобто спостереження з рангом R_i , найближчим до величини np);
- визначити коефіцієнти рангової кореляції R_S Спірмена між серіями порядкових спостережень чи коефіцієнт множинної рангової кореляції – коефіцієнт конкордації Кендала W [3; 6; 11; 13-16];
- будувати за допомогою отриманих величин інші статистичні процедури.

У тих випадках, коли стани об'єктів дозволяють робити лише порядкові порівняння, хоч і в експерименті вони вимірюються через величини, що пов'язані з ними опосередковано, однак такі, що фіксуються у числових шкалах, ці вимірювання все ж залишаються вимірюваннями у порядковій шкалі. І. Пфанцагль наводить показові приклади вищезазначеного [21], одним з яких є оцінювання розумових здібностей, коли орієнтуються на час, що витрачає випробуваний на розв'язання тестової задачі.

Ми спеціально до назви шкали впорядкування додали вислів «у суворому розумінні». Справа в тім, що порядкові (у суворому розумінні) шкали визначаються лише для заданого набору об'єктів, що порівнюються. У цих шкал немає загальноп-

рийнятого, тим більше абсолютного стандарту. Тому за певних умов правомірний вираз «перший у світі, другий у Європі» означає, що, скажімо, український студент – переможець всесвітньої олімпіади (наприклад, з інформатики), посів друге місце на загальноєвропейських змаганнях. Як витікає з розміреності розглядуваних нами різноманітних шкал оцінювання, найефективнішою з точки зору рейтингової оцінки РНД є 100-бальна шкала (табл. 1).

3. *Метричні шкали. Інтервальне вимірювання.* Якщо впорядкування об'єктів можна виконати так, що відомими стануть відстані між властивостями будь-яких двох з них, то вимірювання буде суттєво сильнішим, ніж у шкалах упорядкування. Ця шкала може мати довільні точки відліку та масштаб. Також необхідно встановити одиницю вимірювання (наприклад, кількість правильних відповідей студента на тест-завдання). Об'єкту присвоюється число, яке дорівнює певній кількості еталонних одиниць вимірювання, що еквівалентне кількості властивості, притаманній об'єкту. Наслідком цього є незалежність відношень двох інтервалів від того, у якій зі шкал ці інтервали виміряно (тобто, яка одиниця довжини інтервалу та яке значення прийнято за початок відліку). Зв'язок між вимірами у таких шкалах є лінійним: $y = ax + b$, $a > 0$, $-\infty < b < +\infty$. Отже, шкала інтервалів єдина з точністю до лінійних перетворень. Побудовані в такий спосіб шкали тому і називають *інтервальними*. Важлива особливість, що відрізняє інтервальне вимірювання від вимірювання відношень, яке буде розглянуто далі, полягає у тому, що властивість оцінюваного об'єкта зовсім не втрачається, коли результат вимірювання дорівнює нулю. Тому, якщо студент отримав оцінку «2» за 4-бальною шкалою, то це не означає, що він не має жодних знань з певної НД. Точка нуль на інтервальній шкалі є довільною. Проте невідомо, в яких саме «еталонних одиницях» необхідно вимірювати довжину відповідних інтервалів та їх різницю, і чому має дорівнювати нуль (мінімальне значення) та максимальне значення недисциплінованості. Числа, що приписуються у процесі

такого вимірювання, мають властивість однозначності та впорядкування і становлять собою кількість одиниць вимірювання, що властиві досліджуваному об'єктам. Окрім того, у даному разі є істотною і різниця між числами.

Назва «шкала інтервалів» наголошує, що в цій шкалі лише інтервали мають значення дійсних чисел і тільки з інтервалами необхідно здійснювати арифметичні операції: якщо виконати арифметичні операції із самими відкликами, забувши про їх відносність, то маємо ризик отримати безглузді результати.

Подібно до того, як визначення показника символу Кронекера є єдиною операцією, що допускається здійснювати над спостереженнями у номінальній шкалі, а обчислення рангу спостереження – у шкалі впорядкування, в інтервальній шкалі єдиною новою допустимою операцією зі спостереженнями є визначення інтервалів між ними. З інтервалами можна виконувати будь-які арифметичні операції, а також вдаватися до будь-яких зручних способів статистичної обробки даних. Наприклад, центральні моменти, у тому числі дисперсія, мають об'єктивний фізичний зміст, а вихідні моменти, у тому числі середнє значення, є відносними поряд із початком відклику. Тому поняття відносної похибки (зокрема, коефіцієнта варіації $n = s / \bar{x}$, де s – стандартне відхилення, \bar{x} – математичне очікування) не має сенсу для інтервальної шкали. Це не означає, що результати вимірювань взагалі не можна підсумовувати у шкалі інтервалів (наприклад, обчислювати вибіркоче середнє $\bar{x} = (1/n) \sum x_i$). Однак з такою величиною необхідно поводитись так само, як і з іншими вихідними спостереженнями: вона залишається інтервальною величиною і набуває числового змісту лише в процесі визначення інтервалів. Тому вибіркоче дисперсія має об'єктивний зміст, хоча і визначається через \bar{x} за виразом $S^2 = M(X - \bar{x})^2$. Річ у тому, що $X - \bar{x}$ є інтервалом.

Звертаючись до рис. 2, нескладно уявити, що РНД випробуваного визначається точкою на континуумі знань і розглядається для оцінювання як різниця між мінімально визначеним (точка відліку) знанням і його реальним виміром. Завдання успішно вирішується лише в разі застосування для виміру 100-бальної шкали, що можна сформулювати за допомогою відповідних тестів, хоча й тут виникає запитання: які саме знання необхідно вважати «нульовими», а які – «абсолютними»? Таку проблему можна вирішити шляхом застосування методів лінгвістичних змінних (ЛЗ) і нечітких множин (НМ) [3]. Наведене й дістало відображення у відповідному ранжируванні ефективності шкал оцінювання знань за ефективністю (табл. 1).

Отже, стає зрозумілим, що показники РНД можуть мати шкали різних типів. Це дає можливість дати точніше визначення кількісного та якісного показника. *Кількісними показниками* будемо називати такі, що мають шкалу не менш досконалу, ніж інтервальна. *Якісними показниками* називатимемо такі, що мають шкалу менш досконалу, ніж інтервальна.

4. Вимірювання відношень. Нехай величини РНД, що вимірюються, задовольняють не тільки аксіоми 4⁰ і 5⁰, але ж і аксіоми адитивності:

$$\begin{aligned} 6^0. \beta \hat{e} \hat{i} \hat{A} = \mathcal{D}^3 B > 0, \delta \hat{i} A + B > P & \begin{array}{l} \text{Ь} \\ \text{П} \\ \text{П} \\ \text{Э} \end{array} \\ 7^0. A + B = B + A & \begin{array}{l} \text{П} \\ \text{П} \\ \text{Э} \end{array} \\ 8^0. \beta \hat{e} \hat{i} A = P \text{ і } B = Q, \text{ то } A + B = P + Q & \begin{array}{l} \text{П} \\ \text{П} \\ \text{П} \end{array} \\ 9^0. (A + B) + C = A + (B + C) & \begin{array}{l} \text{П} \\ \text{П} \\ \text{Ю} \end{array} \end{aligned} \quad (5)$$

Це істотне посилення шкали: вимірювання допускають порівняння не тільки самих значень РНД або їх різниць, але ж і різноманітних арифметичних комбінацій цих значень, якщо, безумовно, йдеться про те, що вони мають фізичний зміст. Уведена у такий спосіб шкала називається *шкалою відношень*. Цей клас шкал має такі особливості: відношення двох спостережених значень величини, що вимірюється, не залежить від того, у якій із таких шкал проведено вимірювання. Цю вимогу задовольняє вираз типу $y = a \cdot x$, $a \neq 0$, $a > 0$, тобто числа

відображають, у скільки разів властивість одного об'єкта перевищує ту саму властивість іншого. Така процедура називається перетворенням розтягання. Величини, що вимірюються у шкалі відношень, мають природний абсолютний нуль, хоча й залишається свобода у виборі одиниць. Тому, якщо нульова точка не довільна, а абсолютна, то має сенс твердження, що у студента *A* вдвічі, втричі або в чотири рази більше деякої властивості (скажімо, знань), ніж у *B*. Проте одразу ж зазначимо, що йдеться не про знання взагалі, навіть з чітко визначеної НД, а про оцінку знань за допомогою певного тесту, а результати виконання саме цього тесту дозволяють робити висновки щодо співвідношення навчальної компетентності студентів *A* і *B*.

Зазначимо, що у науках про поведінку людини переважна більшість вимірювань належить до номінального, порядкового та інтервального рівнів. Лише найменш важливі змінні допускають вимірювання відношень. Іноді змінні шкали відношень, такі як час розв'язання задачі або завчання тексту, можуть становити інтерес, але ж це відбувається не часто.

5. *Абсолютна шкала.* Розглянемо таку досконалу шкалу, що має і абсолютний нуль, і абсолютну одиницю. Ця шкала не єдина з точністю до якогось перетворення, а просто унікальна. Саме такі властивості має числова вісь, що природно називають *абсолютною шкалою*. Важливою особливістю шкали у порівнянні зі всіма іншими є абстрагованість (безрозмірність) і абсолютність її одиниці. Це дозволяє здійснювати над вимірами абсолютної шкали такі операції, які є недоречними для вимірів інших шкал, зокрема застосовува-

ти їх як показник ступеня і аргументу логарифма. Числова вісь використовується як шкала вимірювання у явній формі переліку предметів, і як допоміжній засіб, присутня у всіх інших шкалах. Внутрішні властивості числової осі, зважаючи на її начебто простоту, є надзвичайно різноманітними, і теорія чисел до цього часу не вичерпала їх до кінця. В абсолютній шкалі визначається, наприклад, кількість об'єктів, що можна виміряти одиничним способом за допомогою ряду натуральних чисел.

З усіх шкал оцінювання знань кількісним вимірюванням в абсолютній шкалі відповідають лише 100-бальна та 200-бальна шкали, використовувати які зручно ще й тому, що вони сприяють асоціативному уявленню викладача як вимірювача та оцінювача знань, а також тих, хто навчається, щодо відсотка ($0 \div 100\%$) засвоєння навчального матеріалу.

Наведене відобразилося у ранжируванні систем оцінювання (табл. 1). У результаті наших досліджень виявлено, що 100-бальна шкала може бути основою створення різноманітних оцінних систем, незалежно від статистично-імовірнісного чи нечіткого підходів до їх формування [3]. У табл. 2 подано зведені відомості про розглянуті шкали вимірювань. І природно, що чим сильніша шкала, тим більше відомостей про НВП дають вимірювання.

6. *Лінгвістичні (нечіткі) шкали.* Шкали вимірювань у дидактиці, що були розглянуті, начебто дозволяють у своїй сукупності провести «всебічну» кількісно-якісну кваліметрію РНД. Однак разом з тим вважаємо доцільним ще раз процитувати

Таблиця 2. Узагальнені характеристики шкал вимірювання

Шкала	Характеристика	Визначаючи відношення	Еквівалентне перетворення шкал	Допустимі операції над даними	Вторинна обробка даних
1	2	3	4	5	6
Номінальна	Об'єкти класифіковані, а класи позначені номерами. Якщо номер одного класу більше або менше іншого, це ще нічого не каже про властивості об'єктів, за виключенням того, що вони відрізняються	еквівалентність (а)	перестановка найменувань	обчислення символу Кронекера δ_{ij}	обчислення відносних частот і операції над ними
Впорядкування	Відповідні значення чисел, що присвоюються предметам, відображають кількість властивості, що належить предметам. Рівні різниці чисел не означають рівних різниць у кількості властивостей	а + переваги (б)	не змінює порядку (монотонне)	обчислення δ_{ij} і рангів R_i	обчислення відносних частот і вибіркового квантилів, операції над ними
Інтервальна	Існує одиниця вимірювання, за допомогою якої предмети можна не тільки впорядкувати, але ж і приписати їм числа так, щоб рівні різниці чисел, що присвоюються предметам, відображали рівні відмінності у кількості властивості, що вимірюється	а + б + зберігання відношень інтервалів (в)	лінійне перетворення	обчислення δ_{ij}, R_i та інтервалів (різностей між спостереженнями)	арифметичні дії над інтервалами
Відношень	Числа, що присвоюються предметам, мають всі властивості інтервальної шкали, але ж, крім того, на цій шкалі існує абсолютний нуль. Значення нуль свідчить про відсутність властивості, що оцінюється. Відношення чисел, що присвоюються у вимірюванні, відображають кількісні відношення властивості, що вимірюються	а + б + в + зберігання відношень двох значень (д)	зрушення: $y = x + nb$, $b = const$, $n = 0, 1, 2, \dots$	усі арифметичні операції	будь-яка зручна обробка, що допускається
Абсолютна	Уся сукупність характеристик попередніх шкал, а також різноманіття їх сукупності	а + б + в + д + безрозмірна одиниця + безрозмірний нуль	шкала унікальна	усі арифметичні операції, використання у якості показника ступеня, основи і аргументу логарифму	будь-яка необхідна обробка

Н.М. Розенберга, який справедливо зазначає [6, 6–7]: «Ніколи дидактичні вимірювання не будуть якоюсь панацеєю від усіх негараздів, закінченою системою точного знання, у межах якої можна досить легко розв'язувати складні дидактичні проблеми. Єдино правильний шлях раціонального синтезу якісного і кількісного аналізу». На цей *раціональний синтез* і спрямовано лінгвістичні (нечіткі) шкали. Їх розробці сприяла нагальна потреба вирішення питання: чи можуть звичайні кількісні методи вимірювань бути справді ефективними під час аналізі *освітніх гуманістичних систем* (у розумінні Л. Заде [22, 6–7]). Зазначений підхід хоч й інтуїтивно, але ж часто та реально застосовується у практиці психолого-педагогічної кваліметрії. Зокрема, всі характеристики оцінок шкали ECTS мають виражений, нечіткий характер, тому їх можна поєднати у таку терм-множину (ТМ) відповідної ЛЗ [3]:

$$T_{ECTS}^M(\mathcal{D}\mathcal{N}\mathcal{A}) = \tilde{T}_X + \tilde{T}_{FX} + \tilde{T}_E + \tilde{T}_D + \tilde{T}_C + \tilde{T}_B + \tilde{T}_A =$$

$$= \overset{\mathcal{O}}{i} \overset{FX}{\hat{a}\hat{c}\hat{a}\hat{i}\hat{i} \hat{a}\hat{z}\hat{e}\hat{i}\hat{i} \hat{i}} + \overset{E}{i} \overset{D}{\hat{a}\hat{c}\hat{a}\hat{i}\hat{i} \hat{a}\hat{z}\hat{e}\hat{i}\hat{i} \hat{i}} + \overset{C}{\hat{a}\hat{i} \hat{n}\hat{o} \hat{a}\hat{o} \hat{i} \hat{i}\hat{i}} + \overset{B}{\hat{c}\hat{a}\hat{i}\hat{i} \hat{a}\hat{z}\hat{e}\hat{i}\hat{i} \hat{i}} + \overset{A}{\hat{a}\hat{i} \hat{a}\hat{d}\hat{a}}$$

$$\overset{B}{\hat{a}\hat{o}\hat{a}} \overset{A}{\hat{a}\hat{i}\hat{i} \hat{a}\hat{d}\hat{a} + \hat{a}\hat{z}\hat{e}\hat{i}\hat{i} \hat{z}\hat{i}\hat{i}}, \quad (6)$$

де $\tilde{T}_A, \tilde{T}_B, \tilde{T}_C, \tilde{T}_D, \tilde{T}_E, \tilde{T}_{FX}, \tilde{T}_X$ – терми (назви) РНД балів шкали ECTS;

$T_{ECTS}^M(\mathcal{D}\mathcal{I}\mathcal{A})$ – ТМ ЛЗ «РНД» шкали ECTS;

+ – позначка операції об'єднання термів.

Нечіткою є класифікація РНД навчених, що була проведена МОНМС України, що теж дозволяє сформулювати ТМ відповідної ЛЗ, але вже для 4-бальної шкали оцінювання:

$$T_4^M (\text{ĐÍ Ä}) = \check{O}_{,2''} + \check{O}_{,3''} + \check{O}_{,4''} + \check{O}_{,5''} =$$

$$= \begin{matrix} \text{í è çüèèé, " í àçàüí à³èüí í "} \\ \text{ì àððàü ò èàí í - ì ðí äóèð èàí èé +} \\ \text{ñàðàüí ³é, " çàüí à³èüí í "} \\ \text{+ ðäü ðí äóèð èàí èé +} \\ \text{äü ñò àð í ³é, " äü äðä "} \\ \text{êí í ñò ðóèð èàí í - äàð³àð èàí èé +} \end{matrix}$$

$$\begin{matrix} \text{àèñ èèé, " à³àü ³í í "} \\ \text{ò äü ð-èé} \end{matrix}, \quad (7)$$

Де $T_4^{\check{I}}$ (ĐÍ Ä) – ТМ ЛЗ «РНД» 4-бальної шкали оцінювання знань;

$\check{T}_{,2''}, \check{T}_{,3''}, \check{T}_{,4''}, \check{T}_{,5''}$ – терми (назви) РНД-балів шкали оцінювання.

Оскільки будь-яка шкала оцінювання складається за якісним принципом «сильно – середнє – слабко – 0 (нейтрально) – слабко – середнє – сильно» [16], то можна отримати якісні, характерологічні, лінгвістичні оцінки 12-бальної шкали [3]. До речі, у контексті аналізу цієї шкали не втратила практичного значення та актуальності шкала Бофорта для вимірювання сили вітру, що може слугувати зразком для розробки лінгвістичних шкал [3].

Експліцитна, якісна, лінгвістична характеристика-диференціація РНД учнів і студентів є принципово важливою в процесі кваліметрії знань ще й тому, що пов'язана з рівнем домагань (РД), якому притаманні всі властивості мотивуючого чинника [3; 20]. Така оцінка має здійснюватися, спираючись не тільки на логіку, здоровий глузд та досвід, але ж і на чітко визначені правила [3; 16; 22 та ін.]. Отже, порушується питання оптимізації процесу якісної класифікації-характеристики кількісно виміряного РНДС. Виходячи з вищезазначеного, оціночні шкали були ранжировані за критерієм лінгвістичної оцінки тих, хто навчається (табл. 1).

Висновки

Узагальнюючи наукові результати з застосування системно-інформаційного підходу та найбільш розповсюджених шкал вимірювань до аналізу бальних шкал кваліметрії РНДС, виділимо такі найбільш істотні положення.

1. Проаналізовано сутність поняття «вимірювання в дидактиці», його характерні риси та ознаки. Розглянуто можливості та сфери застосування шкал вимірювань, показано їхні специфічні особливості. Проведено порівняльний аналіз ефективності комплексу різноманітних систем оцінювання знань. Визначено, що найбільш ефективною для проведення розрізнення студентів за РНД, ранжируванням їх за цим самим критерієм, а також визначенням обсягу їхніх знань є 100-бальна шкала, якій притаманні всі властивості унікальної абсолютної шкали.

2. Запропоновано додати до традиційних напрямів кваліметрію знань за допомогою ЛЗ, що сприятиме появі змішаної якісно-кількісної оцінки та відкриватиме перспективи для якісної оцінки-характеристики РНДС, спираючись на певну сукупність точок континууму знань при визначеному ступені належності кожного знання до цієї характеристики. Використання такої оцінки додатково мотивуватиме та спрямовуватиме студента на опанування знаннями.

3. Подальші дослідження з розвитку системно-інформаційної методології досліджень у дидактиці необхідно проводити за напрямками побудови нечітких моделей кваліметрії та порівняння РНД для усіх бальних шкал, що застосовуються у ВНЗ та школах; а також розробки й адаптація методів системного аналізу для визначення основної домінанти діяльності учасників НВП у показниках навчання.

Список літератури

1. Системы высшего образования стран Запада. – М.: УДН, 1991. – 353 с.
2. Организация, уровни и квалификация образования в зарубежных странах: Справ.-метод. пособ. / Под ред. В.М. Филиппова. – М.: Центр сравнительной образовательной политики, 2004. – 416 с.
3. Камишин В.В. Методы системного анализа у кваліметрії навчально-виховного процесу: Монографія / В.В. Камишин, О.М. Рева. – К.: Інформаційні системи, 2012. – 270 с.
4. Журавський В.С. Болонський

процес: головні принципи входження в Європейський простір вищої освіти / В.С. Журавський, М.З. Згуровський. – К.: ІВЦ Вид-во «Політехніка», 2003. – 200 с.

5. Модернізація вищої освіти України і Болонський процес / Уклад. М.Ф. Степко, Я.Я. Болюбаш, К.М. Лівківський, Ю.В. Сухарніков; Відп. ред. М.Ф. Степко. – К.: Освіта України, 2004. – 60 с.

6. Розенберг Н.М. Проблеми измерений в дидактике / Н.М. Розенберг; Под ред. Д.А. Сметанина. – К.: Вища школа, 1979. – 175 с.

7. Супес П. Основы теории измерений / П. Супес, Р. Зинес // Психологические измерения. – М.: Мир, 1967. – С.9–110.

8. Анфилатов В.С. Системный анализ в управлении: Учеб. пособ. / В.С. Анфилатов, А.А. Емельянов, А.А. Кукушкин. – М.: Финансы и статистика, 2002. – 368 с.

9. Субетто А.И. Введение в квалиметрию высшей школы / А.И. Субетто // Общие основания квалиметрии высшей школы. – М.: Исследоват. центр проблем качества подготовки специалистов, 1991. – Кн.1. – 84 с.

10. Одерій Л.П. Кваліметрія вищої освіти: методологія та інструментарій: Монографія / П.П. Одерій. – К.: МКА. УЗМН, 1996. – 264 с.

11. Євтух М.Б. Математичне моделювання в психологічних та соціологічних дослідженнях: Підручн. / М.Б. Євтух, М.С. Кулік, Е.В. Лузік, Т.В. Ільїна. – К.: Інформаційні системи, 2012. – 428 с.

12. Анастаси А. Психологическое тестирование: Пер. з англ. В 2-х кн. / А. Анастаси; Под ред. К.М. Гуревича, В.И. Лубовского. – М.: Педагогіка, 1982. – Кн.1. – 320 с.; Кн.2. – 336 с.

13. Бурлачук Л.Ф. Словарь-

справочник по психодиагностике / Л.Ф. Бурлачук, С.М. Морозов. – К.: Наукова думка, 1989. – 200 с.

14. Лакин Г.Ф. Биометрия / Г.Ф. Лакин. – М.: Высшая школа, 1990. – 352 с.

15. Толстова Ю.Н. Математико-статистические модели в социологии / Ю.Н. Толстова. – М.: ГУ ВШЭ, 2008. – 244 с.

16. Надежность и эффективность в технике: Справочник в 10 т. / Под общ. ред. В.Ф. Уткина, Ю.В. Крючкова // Эффективность технических систем. – М.: Машиностроение, 1988. – Т.3. – 328 с.

17. Перегудов Ф.И. Введение в системный анализ: Учеб. пособ. / Ф.И. Перегудов, Ф.П. Тарасенко. – М.: Высшая школа, 1989. – 367 с.

18. Герасимов Б.М. Організаційна ергономіка: методи та алгоритми досліджень і проектування: Монографія / Б.М. Герасимов, В.В. Камишин. – К.: Інформаційні системи, 2009. – 212 с.

19. Психология труда: Пер. со словац. / П. Крбатя, Й. Мюллнер и др.; Общ. ред. и предисл. К.К. Платонова. – М.: Профиздат, 1979. – 216 с.

20. Козелецкий Ю. Психологическая теория решений / Ю. Козелецкий; Под ред. Б.В. Бирюкова; Пер. с польск. Г.Е. Минца, В.Н. Поруса. – М.: Прогресс, 1979. – 504 с.

21. Пфанцагль И. Теория измерений / И. Пфанцагль. – М.: Мир, 1976. – 248 с.

22. Заде Л. Понятие лингвистической переменной и его применение к принятию приближенных решений / Под ред. Н.Н. Моисеева, С.А. Орловского; Пер. с англ. Н.И. Ринго. – М.: Мир, 1976. – 165 с.