

Інститут педагогіки і психології професійної
освіти Академії педагогічних наук України

На правах рукопису

БУЛАХ Ірина Євгенівна

ТЕОРІЯ І МЕТОДИКА КОМП'ЮТЕРНОГО
ТЕСТУВАННЯ УСПІШНОСТІ НАВЧАННЯ
(на матеріалах медичних навчальних закладів)

13.00.01 — теорія та історія педагогіки

Автореферат
дисертації на здобуття наукового ступеня
доктора педагогічних наук

КИЇВ — 1995

АВ 32.983

Дисертацією є рукопис.

Робота виконана у Київському університеті імені Тараса Шевченка.

Науковий консультант — доктор педагогічних наук, професор А. М. Ксюк А.М.

Офіційні опоненти: доктор педагогічних наук, професор Підласий І.П.
доктор технічних наук, професор Биков В.Ю.
доктор педагогічних наук, професор Черпінський М.В.

Провідна установа: Київський державний економічний університет

Захист відбудеться " 4 " лютого 1995 р. о 14⁰⁰ г. на засіданні спеціалізованою вченої ради Д 01.61.01. в Інституті педагогіки і психології професійної освіти АПН України за адресою: 254060, м. Київ, вул. М.Берлінського, 9.

З дисертацією можна ознайомитись у бібліотеці Інституту педагогіки і психології професійної освіти АПН України (254060, м. Київ, вул. М.Берлінського, 9).

Автореферат розісланий " 1 " березня 1995 р.

Вчений секретар спеціалізованої вченої ради

Г.М.Цибульська

ЛНБ ім. В. Стефаника
АН України

ЛНБ України ім. В. Стефаника
00777126 (U)



Загальна характеристика дисертаційної роботи

АКТУАЛЬНІСТЬ І СТУПІНЬ ДОСЛІДЖЕНОСТІ ТЕМАТИКИ. Реформування усіх сфер суспільної діяльності, яке почалося з 1991 р. в Україні, наближення їх до загальноосвітніх показників на перший план висуває проблему співставлення рівнів підготовки вітчизняних та іноземних фахівців, з метою узгодження вимог щодо їхньої кваліфікації. Вирішення цієї проблеми можливе лише при впровадженні сучасних методів вимірювання та оцінювання рівня знань, які активно розвивалися й широко використовуються в Європі та США. Їх коріння — у надбанні фахівців у галузі психолого-педагогічних досліджень, які на початку 50-х років поставили завдання побудувати теоретично обґрунтовані психологічні теорії навчання та на їх основі сучасні технології навчання. Саме зусиллями цих вчених та їх послідовників в галузі теорії навчання сформований такий напрям, як інформаційні технології навчання. Їх джерелом є перш за все здобутки вчених, які розробили загальну методологію та апарат системного дослідження: педагогічних систем — А.Алексюк, Ш.Амонашвілі, С.Архангельський, Ю.Бабанський, В.Безпалько, Б.Гершунський, С.Гончаренко, Т.Львіна, Н.Кузьміна, М.Нікандров, І.Підласий, В.Сухомлинський, М.Черпінський; психології особистості — Б.Ананьев, Г.Балл, Л.Виготський, П.Гальперін, Г.Костюк, О.Леонтьев, С.Рубінштейн, В.Семиченко, Н.Талізїна, Д.Узнадзе; інформатизації освіти — В.Биков, В.Глушков, М.Жалдак, В.Козаков, Ю.Машбиць, О.Мінцер, В.Міхалевич, В.Нетяженко, Ю.Попов, А.Стогній, О.Чалий, І.Чекман, М.Шкіль, В.Шкурба, М.Ядренко та інші. Серед зарубіжних

вчених це перш за все праці Є.Гарсі, Н.Краудера, Б.Скіннера, Р.Сміта, Р.Тайлера, Є.Толмана, Є.Торндайка, К.Хала.

Початок 60-х років визначається формуванням напрямів психолого-педагогічних досліджень, серед яких окреме місце посідають вимірювання якісних ознак. Розвиток цього напряму реалізувався у створенні самостійної галузі наукових досліджень — тестології. А поширення методу тестування у розвинутих країнах Заходу набуло масового характеру і впроваджено майже в усі сфери діяльності людини. Що стосується системи освіти — то це перш за все поширення стандартизованих загальнонаціональних тестів, тестування за якими набули характеру обов'язкових. Отже, кінець 70-х початок 80-х років в Західних країнах — це створення нормативної бази щодо прийняття рішень про досягнення навчальної мети та основ для переходу на нові форми вимірювання — орієнтовані на критерії. Цей напрям у теорії вимірювання рівня знань пов'язаний з працями Р.Ебеля, Дж.Макка, Х.Уельса, Р.Фрике, Дж.Ханта та інших. Саме ці вимірювання є теоретичним підґрунтям впровадження інформаційних технологій навчання.

Інформаційні технології навчання, а йдеться перш за все про інтелектуальні навчальні системи, ґрунтуються на нових підходах щодо подання інформації й перетворення її на знання та нових категоріальних поняттях: база знань, модель предметної області, генератор стратегії навчання, модель учня. Але будь-який засіб навчання, включаючи і новітні технології, є лише засобом реалізації методичних ідей з метою розкриття дидактичних властивостей цього методу. В той же час, складовою будь-якого методу навчання є метод вимірювання рівня знань. *Інформаційні технології потребують об'єктивних та комп'ютеропридатних методів вимірювання знань, до яких належить тестування.* Отже, вивчення закономірностей процесу комп'ютерного тестування є

передумовою впровадження комп'ютерних навчальних систем.

З іншого боку, в деяких галузях діяльності людини вимоги щодо її кваліфікації мають особливе значення. Це стосується перш за все галузей, де помилка може мати фатальні наслідки. Серед них не останнє місце посідає медична практика. З метою контролю за кваліфікацією медиків у більшості розвинутих країн Заходу введено ліцензування права лікарської практики та багаторівневу підготовку майбутніх лікарів. Ліцензія видається за результатами державних випробувань, що проходять у формі тестування за загальнонаціональними тестами.

Характерною особливістю нової концепції вітчизняної медичної освіти є перехід на комп'ютерне тестування як форму вимірювання рівня теоретичних знань. Комп'ютерне тестування — це складова загальних іспитів, яке доповнюється ще перевіркою практичних лікарських навичок. З 1993/94 н.р. така форма іспитів введена МОЗ України по закінченню інтернатури та на випускних іспитах в медичних інститутах, при присвоєнні кваліфікації лікаря-практика.

Практичний досвід автора з реалізації переходу на нові форми вимірювання рівня знань свідчить про те, що державні тестування проводяться в умовах недостатньої теоретико-методологічної проробки цієї проблеми. Тестуванню, результати якого мають адміністративні наслідки, не передують попередні випробування, відсутні методики його реалізації, теоретичні положення щодо аналізу та інтерпретації результатів. Завдяки цьому, поширення набули нестандартизовані тести, впровадження яких, як правило, призводить до дискредитації самого методу вимірювання. Це вимагає ґрунтовних досліджень у галузі тестування як методу вимірювання рівня знань, а також розробки теоретичних основ та методів його впровадження.

Узагальнюючи вище наведене, можна дійти висновків, що

проблеми вимірювання та оцінювання рівня знань в сучасних умовах є надзвичайно важливими у зв'язку з постійним збільшенням обсягу знань та змінами у формі їх подання, впровадженням комп'ютерної техніки та інформатизацією суспільства, перетвореннями, що відбуваються у суспільному житті нашої країни. В той же час механічне перенесення досягнень зарубіжних вчених в галузі тестології неможливе з багатьох причин, серед яких: теорія тестів знаходиться у стані розвитку, нормативна база прийняття рішень щодо досягнення навчальної мети має створюватися з урахуванням традицій та досягнень вітчизняної педагогіки.

Отже, актуальність дисертаційної роботи зумовлена:

1. Необхідністю постійного підвищення ефективності навчального процесу шляхом впровадження інформаційних технологій навчання, які потребують нових, більш продуктивних форм подання знань та методів їх вимірювання, що пов'язане з постійним зростанням обсягу наукових знань, у тому числі й медичних.

2. Узгодженням та співставленням рівня набутих знань випускників медичних вузів України з рівнем, який існує у розвинутих країнах Заходу, що є наслідком розпочатого процесу інтегрування України у загальносвітову спільноту. Але методи вимірювання та оцінювання рівня знань, що прийняті в наших вузах, унеможливають таке співставлення.

3. Завданнями реформування вищої медичної освіти та впровадженням комп'ютерного тестування як обов'язкової складової випускних іспитів у медичних вузах. Введенням МОЗ України обов'язкового комп'ютерного тестування при визначенні кваліфікації лікаря-практика.

4. Необхідністю дослідження впливу використання комп'ютерних технологій на управління навчальним процесом.

ОБ'ЄКТ ДОСЛІДЖЕННЯ — процес контролю навчання у вищих навчальних закладах.

ПРЕДМЕТ ДОСЛІДЖЕННЯ — процес вимірювання успішності навчання, в якому: інструментом вимірювання є комп'ютерна тест-система, процедурою вимірювання є комп'ютерне тестування, процедурою оцінювання є шкалювання.

МЕТА ДОСЛІДЖЕННЯ — виявлення сутності та закономірностей процесу вимірювання успішності навчання, розробка теорії комп'ютерного тестування успішності й на цій основі кількісних методів оцінки результатів тестування, методик реалізації нормативних та критеріальних вимірювань, принципів оцінювання, рекомендацій щодо впровадження комп'ютерного тестування у навчальний процес медичних вузів.

КОНЦЕПЦІЯ ДОСЛІДЖЕННЯ включає положення:

— на сучасному етапі розвитку теорії вимірювання комп'ютерне тестування є найбільш об'єктивним та інформативним методом вимірювання успішності навчання, а повільне його впровадження пов'язане з відсутністю теоретично обґрунтованих та практично перевірених дидактичних основ його ефективного застосування;

— впровадження тестування у навчальний процес без врахування дидактичних закономірностей процесу вимірювання завдає шкоди не лише самому методу, а й навчальному процесу.

ГІПОТЕЗА ДОСЛІДЖЕННЯ — підвищення ефективності процесу навчання в умовах інформатизації освіти можливе лише при впровадженні об'єктивних та інформативних методів контролю, що потребує розвитку теорії вимірювання успішності навчання і створення інструментарію кількісного аналізу результатів тестування для об'єктивного введення нормованих стандартних шкал оцінювання та критеріїв щодо прийняття рішення про досягнення навчальної мети.

Мета та наукова проблема, об'єкт, предмет, концепція та гіпотеза дослідження зумовили постановку таких ЗАВДАНЬ:

- виявити основні закономірності процесу вимірювання успішності навчання, встановивши сутність та оцінивши властивості основних категорій цього процесу,
- сформулювати теоретичні основи комп'ютерного тестування успішності та критерії кількісного аналізу результатів,
- розробити методики експертизи тесту та тестових завдань, а також валідазації та стандартизації тесту,
- визначити принципи оцінювання рівня знань та критерії щодо прийняття рішення про досягнення навчальної мети,
- дослідити вплив комп'ютерного контролю на навчальний процес та визначити умови його ефективного застосування,
- розробити методичні рекомендації щодо впровадження комп'ютерного тестування навчальної успішності у навчальний процес медичних вузів.

МЕТОДОЛОГІЧНУ ОСНОВУ ДОСЛІДЖЕННЯ становлять загальнонаукові засоби пізнання та системно-структурного аналізу об'єктів, визначаючи стратегію дослідження. Тактика дослідження ґрунтувалася на методології другого рівня і пов'язувалася з принципами, методами та теоріями наукових досліджень у предметній області, а саме: психолого-педагогічними теоріями навчання, теорією вимірювання, принципами педагогічної та психологічної діагностики, законами тестології. З метою одержання вірогідних результатів важливе місце при дослідженні приділялося використанню математичних методів: *вибіркового методу* для визначення вимог, що забезпечують статистичну достовірність висновків; *методу перевірки статистичних гіпотез* для встановлення статистичної достовірності впливу умов проведення дидактичного експерименту; *дисперсійного аналізу* для визначення статистичної

достовірності впливу зовнішніх факторів на результати; *методу перевірки гіпотез законів розподілу*: оцінки коефіцієнтів асиметрії та ексцесу, критерію χ^2 -розподіл; *кореляційного та регресійного аналізу* для визначення наявності та форми зв'язку між стохастичними величинами, які характеризують процес вимірювання. Математичний аналіз здійснювався із застосуванням персонального комп'ютера у пакеті статистичних програм "STATGRAF".

Проведений на системному рівні аналіз надав можливість, побудувавши факторну, ієрархічну та структурну моделі, вивчити предмет дослідження з точки зору одночасної цілісності та диференційованості на окремі складові частини. Це дослідження проводилося на трьох рівнях: I — комп'ютерне тестування як самостійний предмет дослідження, II — комп'ютерний тестовий контроль як елемент навчального процесу, III — вплив на управління навчальним процесом системи вимірювання, що ґрунтується на взаємозв'язку і цілісній методиці об'єктивного вимірювання знань та системному аналізі його результатів.

Дидактичний експеримент, який проводився у роботі, пов'язаний із застосуванням таких методів як: *спостереження, бесіди, анкетування, моделювання*.

Методи: *спостереження, бесіди, анкетування* дозволили авторові: по-перше, при проектуванні комп'ютерної експертної діагностичної системи обрати оптимальну форму діалогу "комп'ютер — особа тестування"; по-друге, провести класифікацію та визначити дію факторів, які впливають на процедуру тестування; по-третє, визначити вплив методу вимірювання знань, як на особу, яка екзаменується, так і на процес навчання.

Експеримент проводився у комп'ютерному класі Українського державного медичного університету. В ньому брали участь пред-

ставники усіх груп, які проходять студіювання у системі вітчизняної медичної освіти: школярі 10, 11 класів медичних ліцеїв м.Києва (243 особи), студенти 1-5 курсів медичного та стоматологічного факультетів (4028 осіб), інтерни (післядипломна підготовка) (137 осіб), аспіранти (67 осіб), викладачі медичних інститутів та медичних училищ України — слухачі факультету підвищення кваліфікації (124 особи). Тестування проводилося з різних дисциплін: основ інформатики, основ математичного моделювання, медичної інформатики, внутрішніх хвороб, ендокринології, нервових хвороб, інфекційних хвороб, ЛОР захворювань, терапевтичної стоматології, а також з усіх спеціальностей для випускних іспитів інтернів за програмою МОЗ України.

Такий діапазон контингенту та дисциплін засвідчує, що розроблені теоретичні положення та одержані результати не мають вузького прикладного значення, а носять загальний характер і можуть бути поширені для застосування у різних вузах не тільки медичного профілю для вимірювання та оцінювання рівня знань.

НАУКОВА НОВИЗНА ДОСЛІДЖЕННЯ. Вперше на системному рівні із застосуванням сучасних інформаційних технологій проведено комплексний аналіз процесу вимірювання рівня знань, теоретично обгрунтовано та експериментально підтверджено закономірності цього процесу. Визначено основні категорії процесу комп'ютерного тестування, їхню ієрархічну підпорядкованість та залежність результатів тестування від характеристик усіх категорій процесу вимірювання. Вперше сформульовано критерії об'єктивності результатів тестування та критерії кількісного аналізу основних категорій процесу тестування, розроблено математичну модель тестів, що закладає теоретичні основи введення критеріїв щодо прийняття рішення про досягнення навчальної мети. Обгрунтовано теоретичні основи конвертації результатів тестування за

інтервальною чотирибальною шкалою оцінювання: початок відліку, масштаб, формулу конвертації за умов мінімізації помилки.

ТЕОРЕТИЧНЕ ЗНАЧЕННЯ ДОСЛІДЖЕННЯ полягає: — у розробці основ теорії вимірювання успішності навчання методом комп'ютерного тестування та створенні методу кількісного аналізу результатів, що дає змогу об'єктивно визначити норми та розробити критерії, для введення критеріальних вимірювань у вітчизняній педагогіці та впровадження інтелектуальних навчальних систем; — у розробці принципів шкалювання у чотирибальну шкалу, загально прийняту шкалу оцінювання, визначенні її обмежень та підготовці теоретичних основ для переходу на багатобальні стандартні шкали оцінювання рівня знань.

ПРАКТИЧНА ЦІННІСТЬ ДОСЛІДЖЕННЯ полягає: — у розробці методики валідазації та стандартизації тестів, методики аналізу результатів тестування щодо визначення валідності процесу вимірювання, методики оцінювання рівня знань та порівняльного аналізу результатів; — у створенні інструментарію кількісного аналізу результатів тестування — *статистичного сертифікату тестів*, алгоритму впровадження комп'ютерного тестування навчальної успішності у навчальний процес медичного вузу, програмного забезпечення — пакету комп'ютерних програм для організації та проведення комп'ютерного тестування.

НА ЗАХИСТ ВІНОСЯТЬСЯ:

1. Експериментально підтверджене положення, що рівень знань, як випадкова величина, підпорядковується статистичним закономірностям і розподіляється за нормальним законом, тому застосування вимірювань, орієнтованих на норми, при підсумковому контролі рівня знань є науково обґрунтованим.

2. Процес тестування як метод вимірювання має багатовимірну структуру, кожна категорія якої має свою сутність та властивості,

що впливають на кінцеві результати вимірювання.

3. Ступінь відхилення від нормального закону розподілу результатів тестування, визначений за коефіцієнтами асиметрії та ексцесу, безпосередньо характеризує об'єктивність результатів вимірювання, що конкретизовано у сформульованих критеріях об'єктивності.

4. Критерії об'єктивності результатів тестування забезпечують на кількісному рівні визначення валідності та методичку валідації основних категорій процесу тестування: тесту як інструментарію вимірювання, процедури тестування, процедури оцінювання.

5. Результати тестування мають конвертуватися у інтервальну шкалу балів; при конвертації за чотирибальною інтервальною шкалою статистично обґрунтовується її початок відліку, масштаб та умови мінімізації помилки; введення інтервальної чотирибальної 0-шкали є першим етапом до переходу на багатобальні шкали оцінювання.

6. Використання комп'ютерного тестового контролю приводить до параметричної адаптації системи навчання, викликаючи зміни характеристик її підсистем: мотиваційного елементу у того, хто навчається; активізації та вдосконалення професійної діяльності у того, хто навчає; збільшення інформаційної значущості та обсягу інформації, яку повинні засвоювати як викладач, так і студент.

АПРОБАЦІЯ та ВПРОВАДЖЕННЯ наукових розробок здійснювалися:

1. У публікаціях матеріалів у монографії, брошурах, методичних рекомендаціях, авторському свідостві, статтях.

2. У доповідях на конференціях, нарадах, з'їзді.

3. У лекціях для слухачів ФПКв медичних інститутів України.

4. В організації сесійних іспитів в УДМУ з внутрішніх хвороб, ендокринології, нервових хвороб, стоматології, медичної інформа-

тики у формі комп'ютерного тестування, що затверджено на Вченій та Методичних Радах університету.

5. У програмах навчальних дисциплін, затверджених МОЗ України, з основ медичної інформатики для медичних училищ, медичної інформатики для студентів медичних інститутів.

6. У положенні про впровадження у навчальний процес медичних вузів рейтингової системи оцінювання, затвердженому МОЗ України.

Публікації здобувача, що висвітлюють основний зміст дисертації, складають 37 авторських друківаних аркушів.

ВІРОГІДНІСТЬ ОДЕРЖАНИХ РЕЗУЛЬТАТІВ. В основу дослідження покладено експериментальні дані: результати екзаменаційної сесії за 1993/94 н.р. УДМУ на 43 кафедрах; результати тестувань, що проводилися за 12 тестами з різних дисциплін для студентів і 42 тестами з різних спеціальностей для інтернів та експериментальних тестів для ліцейстів і слухачів ФПК. Для забезпечення достовірності висновків, виконувалася умова масовості, тобто об'єм виборок був репрезентативним і коливався від 90 до 1009 осіб. Статистичні висновки робилися із застосуванням сучасних комп'ютерних програм обробки даних на рівні значущості $p=0,05$. Сучасний методичний підхід, теоретична обґрунтованість наукових положень, репрезентативна кількість експериментальних даних забезпечують логічність, обґрунтованість і достовірність наукових положень та висновків, сформульованих у дисертації.

ОСОБИСТІЙ ВНЕСОК. Автором самостійно одержані усі наукові результати, що виносяться на захист. Комп'ютерне тестування проводилося разом з викладачами курсу медичної інформатики Українського державного медичного університету — доц. Т.І.Жегрій, ас. Л.Г.Лесько, ас. І.М.Шило та викладачами відповідних профільних кафедр.

Основний зміст дисертації

Дисертація складається із вступу, п'яти розділів, підсумкових висновків, списків цитованої та основної використаної літератури, додатків. Зміст дисертації викладений на 380 сторінках рукопису, містить 14 таблиць і 50 малюнків (загальний обсяг — 430 сторінок).

У ВСТУПІ обгрунтовано вибір наукової проблеми, що підлягає осмисленню та вирішенню, аргументована її актуальність та ступінь розробленості; визначені об'єкт, предмет, мета та завдання дослідження; висвітлені наукова новизна, теоретична та практична цінність дослідження, основні положення, які захищає автор, особистий внесок, апробація та впровадження результатів.

У РОЗДІДІ I "Контроль навчання в умовах інформатизації освіти" аналізуються чинники, що зумовлюють виникнення новітніх технологій навчання, умови та засоби їхньої реалізації. Досліджуються взаємозв'язок та тенденції розвитку методів контролю навчання і впровадження інформаційних технологій. Дається літературний огляд з історії педагогічної тестології, визначається сучасний стан застосування тестування, як методу вимірювання та оцінювання знань.

Виходячи із загальної тенденції розвитку суспільства — інформатизації, технізації та автоматизації усіх сфер діяльності людини, ефективне ведення навчального процесу потребує створення сучасної технології навчання. Традиційні методи викладання та спілкування підлягають суттєвому перегляду внаслідок впровадження в навчальні заклади нових технічних засобів на базі комп'ютерної техніки. Інформатизація суспільства обумовила виникнення новітніх інформаційних технологій, бурхливий розвиток

яких припадає на початок 80-х років, і є наслідком розробки та впровадження персональних комп'ютерів. Впровадження новітніх технологій навчання потребує наявності апостеріорної, статистично достовірної інформації для розробки критеріїв, правил та моделей прийняття рішень щодо досягнення навчальних цілей. Ці технології навчання, ґрунтуючись на нових підходах щодо подання та засвоєння знань, потребують нових, сучасних методів їх вимірювання. Застосування традиційних методів вимірювання знань — усного опитування та чотирибальної системи їх оцінювання, гальмує не лише інтегрування випускників вищих медичних навчальних закладів у міжнародну спільноту практикуючих лікарів, а й процес інформатизації освіти взагалі [1, 15]. Пошук досконалих методів вимірювання рівня знань на сучасному етапі розвитку інформаційних технологій набуває надзвичайної актуальності, оскільки об'єктивізація процесу вимірювання, забезпечуючи зворотний зв'язок, дає можливість координувати цей розвиток.

Сучасна психолого-педагогічна діагностика розвивається по двох напрямках: використання методів, орієнтованих на норми, та методів, орієнтованих на критерії. Але на думку Р.Фрике, одного з провідних фахівців у галузі вимірювань, відмінностям цих методів надається занадто багато уваги, "доречніше було б вважати вимірювання, орієнтовані на критерії, не іншим типом тестів, а, можливо, новим аспектом інтерпретації результатів вимірювання". В той же час впровадженню вимірювань, орієнтованих на критерії, які забезпечують умови створення *інтелектуальних навчальних систем*, має передувати період розробки нормативної бази, що буде покладена в основу критеріїв та моделей прийняття рішень щодо досягнення навчальних цілей. Отже, ці обидва методи, доповнюючи та уточнюючи один одного, нададуть можливість об'єктивно і точно вимірювати та оцінювати рівень знань.

Проблемою педагогічної діагностики займалися і займаються педагоги та психологи, серед яких: В.Аванесов, К.Берка, П.Воловик, Ю.Гільбух, Р.Ебель, К.Інгенкамп, А.Клименко, В.Козаков, П.Лазарсфельд, Ф.Лорд, Дж.Макк, Є.Марченко, Ю.Машбиць, В.Міхеєв, М.Олійник, І.Підласий, Дж.Раш, Н.Розенберг, Р.Сміт, Р.Тайлер, К.Хал, Р.Хамбет, В.Хант, Р.Фрике та інші. Але необхідно констатувати, що на відміну від розвинутих країн Заходу, у нашій Державі далі наукових дискусій ця проблема не розвинулася.

Останнім часом необхідність у змінах змушує як МО, так і МОЗ України розпочати роботи по впровадженню рейтингової системи оцінювання. Розробляються положення про впровадження цієї системи, але не враховується, що система оцінювання узгоджується із системою вимірювання. Отже, вирішувати потрібно загальну проблему, а не окрему.

РОЗДІЛ 2 "Методи та методики дослідження процесу вимірювання знань" присвячений опису методів наукових досліджень процесу вимірювання знань. У ньому розглянуті методи експериментального дослідження, які базуються на використанні сучасних комп'ютерних технологій, а саме експертної діагностичної системи вимірювання навчальної успішності. За допомогою цієї комп'ютерної системи реалізувався дидактичний експеримент. Авторські права на цю систему захищені у роботі [14]. Для експертної діагностичної системи вимірювання навчальної успішності характерні такі особливості: зручність у використанні та відсутність вимоги до користувача щодо спеціальної підготовки з інформатики. Вона об'єднує три підсистеми: *інструментальну, тестування, аналізу та інтерпретації*. Кожна з цих підсистем виконує свої специфічні функції. *Інструментальна підсистема* є автоматизованою системою роботи з банком тестових завдань; її призначення — автоматизація процесу вимірювання, формування

тесту та накопичення тестових завдань. Вона виконана за принципом "оболонки", тобто відкритої системи, що дозволяє проводити формування тестів з різних дисциплін та виконувати різні процедури з базою тестових завдань. *Підсистема тестування* призначена для автоматизації процесу вимірювання, накопичення та збереження результатів тестування, формування їх у відповідні звітні форми. *Третя підсистема* виконує функції валідації тесту і, з точки зору її організації та функціонування, є діалоговою експертною системою.

Діагностична експертна система вимірювання дозволила авторів, використовуючи бази даних та бази знань, формувати тести з різних дисциплін в залежності від цілей вимірювання. Вона також надала широкі можливості моделювання: по-перше, різних моделей комп'ютерного тестування з використанням різних видів тестів; по-друге, різних умов процедури тестування та можливості детального математичного і статистичного аналізу результатів.

У РОЗДІЛІ 3 "Теоретичні основи процесу вимірювання та оцінювання знань", використовуючи інформаційно-логічний підхід до навчання як процесу трансформування інформації на знання, визначаються основні принципи процесу вимірювання рівня знань, проводиться порівняльний аналіз існуючих методів вимірювання та оцінювання знань, аналізуються переваги та недоліки кожного з них.

Вимірювання якісних характеристик, до яких належить і рівень знань, мають відповідати загальним вимогам теорії вимірів, найважливішими з яких є критерії: *об'єктивності, надійності, валідності й точності.*

Проведений у роботі за критеріями, що названі вище, порівняльний аналіз методів вимірювання знань, які використовуються в сучасній педагогіці: *усна форма перевірки знань або усне*

опитування, письмова форма перевірки знань або письмова робота, співбесіда у формі інтерв'ю, тестування, свідчить про те, що останній метод, тобто тестування, найбільше відповідає критеріям якості при визначенні рівня знань. У той же час, усна форма контролю знань як метод діагностики ефективно може використовуватися при вимірюванні знань, пов'язаних з мовним розвитком, або мовними умовами спілкування; форма письмового контролю — при вимірюванні знань, пов'язаних із визначенням когнітивного розвитку та проблемного мислення; інтерв'ю або співбесіда — для одержання інформації щодо особистості респондента: його зовнішнього вигляду, поведінки, мови тощо.

Вимірюванням та оцінюванням якісних характеристик займається тестологія — галузь науки, де методом вимірювання є тестування. Цей метод передбачає, що: *інструментом вимірювання є тест, складений з тестових завдань, процедурою вимірювання є тестування, методом оцінювання є шкалювання.*

У сучасній тестології існує два напрями щодо стратегії вимірювання, що призвело до диференціації не лише тестів, а й тактики тестування на два типи: вимірювання, орієнтовані на норми, та орієнтовані на критерії. Крім відмінностей щодо стратегії та тактики процесу вимірювання, вони мають суттєву різницю щодо закономірностей, яким підпорядковуються результати вимірювань.

Так, вимірювання, орієнтовані на норми, ґрунтуються на класичній теорії тестів, в основі якої лежать такі положення:

- гіпотеза нормального розподілу, яка допускає, що якісні характеристики, притаманні людині, розподіляються у відповідності з кривою Гауса (за нормальним законом),
- ранговий принцип, який базується на уяві, що якісний параметр визначається щодо середньостатистичного значення досліджуваної групи.

У той же час, при вимірюваннях, орієнтованих на критерії, закладається відхід від нормального закону розподілу результатів тестування, оскільки ймовірності дати вірну p або невірну q відповідь не рівні. Це призводить до відхилення від нормального закону розподілу та накопичення результатів тестування на одному з кінців шкали. При вимірюваннях, орієнтованих на критерії, необхідно визначити: коли можна прийняти рішення щодо досягнення навчальної мети; яку кількість завдань для цього потрібно виконати; до якої групи за показниками навчальної успішності потрібно віднести особу, яка тестувалася. В теоріях вимірювань, орієнтованих на критерії, використовуються такі моделі обробки результатів: а) дихотомні; б) квантитативні; в) політомні. Усі ці моделі виходять з того, що результати тестування повинні розподілятися за біноміальним законом розподілу. Але, варто враховувати те, що з математичної точки зору при рівних ймовірностях дати вірну p або невірну q відповідь, біноміальна крива симетрична і, при збільшенні випробувань, наближається до своєї границі — нормальної кривої. Біноміальна крива має суттєву асиметрію лише при дуже великій різниці ймовірностей дати вірну p або невірну q відповідь. Крім того, усі моделі обробки даних передбачають наявність нормативної бази прийняття рішень, яка створюється лише при вимірюваннях, орієнтованих на норми. Все це, на думку автора, підтверджує наведене раніше висловлювання Р.Фрике про інтегральну узгодженість обох методів вимірювання.

Нормативна база щодо прийняття рішень створювалася в різних галузях, де використовуються діагностичні методи, наприклад, медицині та психології. Що стосується вітчизняної педагогіки, то можна констатувати, що об'єктивна нормативна база щодо прийняття рішень про досягнення навчальної мети не розроблена

насамперед з причин застосування малооб'єктивних методів вимірювання та оцінювання рівня знань.

Отже, гіпотеза щодо закону розподілу є предметом широкого обговорення, оскільки дає можливість надати педагогічним процесам кількісного визначення та встановити основні його закономірності. Але на даний момент у літературі відсутні науково обгрунтовані та експериментально підтвержені докази того, що рівень знань як якісна характеристика, пратаманна людині, розподілена за тим чи іншим законом розподілу. А оскільки нормативна база має створюватися на вимірюваннях або вихідних, або підсумкових знань, що створює нормативні інтервали, тому першочерговим завданням було вирішення саме проблеми закону розподілу цих величин. Для цього здійснювався статистичний аналіз результатів іспитів екзаменаційної сесії в УДМУ з 43-х дисциплін, які проводилися досвідченими екзаменаторами у формі усного опитування, письмової роботи, бланкового та комп'ютерного тестування; умови масовості забезпечувалися. Доведено, що оцінка конвертована за чотирибальною шкалою, як результат вимірювання рівня підсумкових знань, розподілена за нормальним законом (у 79%), а відхилення від такого розподілу (у 21%) пов'язані з порушенням процедури вимірювання.

Підсумовуючи, був зроблений висновок, що використання процесу тестування, як методу вимірювання та оцінювання рівня підсумкових знань є науково обгрутованим.

РОЗДІЛ 4 "Основні закономірності комп'ютерного тестування успішності навчання" присвячений питанням розвитку теорії тестів для комп'ютерного тестування, використовуючи основні положення якої, сформульовані критерії кількісного аналізу категорій процесу тестування. Це дало можливість розробити алгоритми валідазації тестів та тестових завдань, інших категорій процесу

вимірювання знань. Визначено принципи оцінювання знань та розроблено методики конвертації результатів тестування у шкальовані бали для різних шкал і для різних типів тестів.

Впровадження тестування у систему освіти гальмується наявністю термінологічних проблем, що зумовлені відсутністю чіткої диференціації основних категорій цього процесу, визначенням їхнього змісту та дослідженням їхніх властивостей. Серед основних категорій процесу вимірювання знань автором виділені такі: метод вимірювання, інструментарій вимірювання, процедура вимірювання, процедура оцінювання [1, 4].

Метод вимірювання — це категорія, що визначає форму інших категорій процесу вимірювання, а саме: *інструментарію, процедури вимірювання, процедури оцінювання*.

Аналіз усіх категорій процесу вимірювання рівня знань ґрунтується на визначальному понятті — валідність.

Термін валідність у зарубіжній літературі використовується для загальної характеристики тесту, який відповідає вимогам та меті тестування. Найбільш загальне визначення валідності тесту дає А.Анастасі: валідність тесту це — "поняття, котре вказує нам на те, що тест вимірює та наскільки добре він це робить".

На основі системного аналізу, проведеного у роботі, загальне поняття валідність щодо процесу вимірювання рівня знань диференціюється за категоріальною та функціональною ознаками на: валідність методу (валідність змісту, валідність відповідності), валідність інструментарію (валідність тесту, валідність тестових завдань), валідність процедури тестування, валідність процедури оцінювання.

Валідність методу, яка поділяється на валідність відповідності та валідність змісту, детально розглянуті у 3-му розділі дисертації. Зазначимо, що валідність змісту визначає ступінь репрезентатив-

ності обсягу знань, визначеному у кількості тестових завдань, що містить тест. Вона пов'язана з двома параметрами: загальним обсягом знань, які повинен засвоїти студент з даної дисципліни та обсягом інформації, яка виноситься на тестування, і при вимірюванні якої роблять статистично ймовірні висновки щодо загального рівня знань. Кількість тестових завдань (довжина тесту), що виноситься на тестування, розраховується статистичними методами для забезпечення умови репрезентативності вибірки та достовірних висновків. На рівні значущості $p = 0,05$ для забезпечення точності вимірювання у 5% довжина тесту повинна становити приблизно 400 тестових завдань, тоді як для точності у 10% приблизно 100.

Враховуючи особливість комп'ютерного тестування, а саме, формування тесту за рандомізованим вибором тестових завдань із загальної кількості, виконання умов репрезентативності довжини тесту та загального обсягу бази тестових завдань є необхідною умовою достовірності вимірювання рівня знань.

Валідність інструментарію — це відповідність характеристик інструментарію вимірювання певним критеріям, що забезпечують об'єктивність результатів. Вона поділяється на валідність тесту та валідність тестових завдань.

Тест навчальної успішності є валідним, якщо при додержанні валідності інших категорій процесу вимірювання, одержані результати відповідають критеріям об'єктивності. Валідність тесту визначається його характеристиками: *складністю* та *розподільною здатністю*. *Складність тесту* визначається у відповідності до середніх результатів тестування середнім значенням репрезентативної вибірки. Якщо одержані результати завишені — робимо висновок: тест за своєю складністю легкий. І навпаки, якщо результати тестування занижені — тест складний. *Розподільна*

здатність — це спроможність при тестуванні з достатньою точністю, що закладена у методі, розрізнити тих, хто тестується, з різним рівнем знань. Розподільна здатність тесту це не лише внутрішня характеристика тесту, а й параметр, на який має суттєвий вплив валідність процедури тестування.

Тестове завдання є валідним, якщо воно відповідає вимогам тесту, тобто забезпечує валідність тесту. Валідність тестового завдання залежить від його характеристик: *складності* та *розподільної здатності*. *Складність* тестового завдання визначається рівнем засвоєння інформації та перетворенням її на знання і поділяється на три рівні: 1 — запам'ятовування, 2 — розуміння, 3 — аналіз. *Розподільна здатність* залежить від зрозумілості запитань, однозначності відповідей, форми запиту.

Визначення цих характеристик для забезпечення валідності інструментарію пропонується проводити на двох рівнях — апіорному та апостеріорному.

Апіорний рівень — це експертиза тестових завдань, що поділяється на два етапи: 1) усунення грубих помилок; 2) усунення неточностей формулювань.

На першому етапі виправляються граматичні та інші помилки. На другому проводиться якісний аналіз, що стосується в цілому корекції розподільної здатності тестових завдань — їхньої зрозумілості та однозначності. Для розробки стандартних тестів, враховуючи наявність різних медичних шкіл, потрібно враховувати думки принаймні трьох незалежних експертів і відбирати тільки ті запитання, відповіді на які співпадають.

Апостеріорний рівень — це аналіз результатів тестування, що забезпечує експертизу тесту на основі теорії тестів. Теоретичною основою процесу тестування є класична теорія тестів, яка активно розвивається і набула нового етапу в теорії незалежних запитань

("Item Response Theory" IRT). Вона детально розкрита Р.Хамблетом у роботі "Фундаментальна теорія незалежних запитань" (1991р.). Серед нерозв'язаних проблем в теорії IRT у першу чергу постають такі: розробка методу кількісного аналізу категорій процесу тестування; обґрунтування оптимальної конструкції тесту та оптимальної моделі тестування; розробка методу порівняльного аналізу результатів тестування за різними тестами. Для розв'язання цих проблем розроблено теорію комп'ютерного тестування, що ґрунтується на ймовірнісній моделі тестування, на основі якої введено критерії визначення об'єктивності результатів тестування та математична модель тестів. Ця модель є розвитком ідей, запропонованих П.Лазарсфельдом, Ф.Лордом та Дж.Рашем, які, визнаючи принципові відмінності психолого-педагогічних вимірювань від фізичних, започаткували математичні моделі взаємозв'язку двох основних параметрів процесу тестування: ступені прояву ознаки — у нашому випадку рівня знань та складності тестового завдання.

Теорія комп'ютерного тестування

Рівень знань X , виміряний методом комп'ютерного тестування, залежить від таких основних факторів: F_1 — параметрів тесту, F_2 — моделі тестування, F_3 — форми діалогу "комп'ютер - особа тестування", F_4 — умов тестування, F_5 — критеріїв оцінювання.

Аналіз дії цих факторів засвідчує, що вплив F_2 , F_3 , F_4 та F_5 може бути зафіксованим при стандартизації процедур тестування та оцінювання. Тоді як вплив фактору F_1 робить результати тестування неінваріантними, що впливає на їхню об'єктивність.

Теорія комп'ютерного тестування ґрунтується на таких припущеннях:

- існує параметр, який характеризує нормований рівень знань особи тестування v ; він визначається за тестом, сконструйовано-

ваним із "стандартних запитань"; за його допомогою можна провести стандартизацію результатів тестування;

- цей параметр може бути виміряний при комп'ютерному тестуванні з певною однозначністю та точністю;
- параметр нормованого рівня знань особи тестування має тісний прямий достовірний кореляційний зв'язок з рівнем знань.

Для побудови моделі був обраний апарат теорії ймовірностей, оскільки педагогічні процеси носять стохастичний характер.

Проведені дослідження дали підставу сформулювати основні положення теорії комп'ютерного тестування:

1. Між рівнем знань X та нормованим рівнем знань v існує прямий тісний кореляційний зв'язок.

2. Аналіз рівняння регресії $X = f(v)$ дає можливість проводити стандартизацію процесу вимірювання рівня знань, забезпечуючи умови, при яких воно буде лінійним.

3. Рівень вихідних та підсумкових знань X , виміряний при комп'ютерному тестуванні, розподілений за нормальним законом.

4. Відхилення від нормального закону розподілу результатів тестування пов'язані з впливом факторів F_1, F_2, F_3, F_4, F_5 .

F_1 — характеристики тесту;

F_2 — модель тестування;

F_3 — форма діалогу "комп'ютер — особа тестування";

F_4 — умови тестування;

F_5 — критерії оцінювання.

Сформульовані основні положення теорії комп'ютерного тестування є основою критеріїв, які дають змогу встановити об'єктивність результатів тестування та виявити чинники, які могли призвести до порушення цієї об'єктивності. Усунення дії цих факторів може розглядатися як валідація методу вимірювання.

В основі запропонованих критеріїв об'єктивності результатів

лежить оцінка міри відхилення від нормального закону, яка визначається за такими коефіцієнтами як асиметрія A_s і ексцес E_x з визначенням їхньої статистичної достовірності за t - критерієм Стьюдента. Залежно від значення та знаку асиметрії або ексцесу результати були поділені на 5 критеріїв:

Критерій I . Якщо дані аналізу засвідчують, що: $A_s = 0$ $E_x = 0$, то робляться висновки: *результати вимірювання є об'єктивними, а отже і процес вимірювання є об'єктивним, що являє собою необхідну умову його валідності.*

Критерій II . Якщо дані аналізу засвідчують наявність від'ємної асиметрії та додатнього ексцесу $A_s = "-"$ $E_x = "+"$, то робляться висновки: *результати вимірювання завищені, тобто процес вимірювання є невалідним.*

Це може бути результатом дії факторів F_1, F_2, F_4 — а саме: невалідності тесту та моделі тестування. Така ситуація, крім названих випадків, частіше усього має місце у разі втручання екзаменаторів у процес тестування або при наявності у осіб тестування вірних відповідей чи інших матеріалів.

Критерій III. Якщо дані аналізу засвідчують про наявність від'ємних асиметрії та ексцесу $A_s = "-"$ $E_x = "-"$, то робляться висновки: *результати завищені за рахунок невалідності тесту.*

Критерій IV . Якщо дані аналізу засвідчують про наявність додатньої асиметрії та від'ємного ексцесу: $A_s = "+"$ $E_x = "-"$, то робляться висновки: *результати занижені за рахунок невалідності тесту.*

Критерій V. Якщо дані аналізу засвідчують що: $A_s = "+"$ $E_x = "+"$, робляться висновки: *результати вимірювання занижені за рахунок невалідності тесту, форми діалогу, умов тестування.*

Застосування критеріїв об'єктивності результатів тестування

дало можливість оцінити також валідність основних категорій процесу тестування.

Для аналізу параметрів інструментарію розроблена методика, яка ґрунтується на аналізі:

- функції розподілу результатів тестування,
- характеристичної кривої тесту і тестового завдання.

Ці методи доповнюють один одний і дають можливість не лише визначити валідність інструментарію, а й з'ясувати причини, з яких вона порушується. Такі висновки набувають свого визначення у критеріях якісного аналізу тестів, сформульованих автором [5]:

Критерій I. Тест є валідним, якщо асиметрія та ексцес закону розподілу дорівнюють нулю ($A_s = "0"$ $E_x = "0"$).

Критерій II. Якщо асиметрія від'ємна та ексцес додатний ($A_s = "-"$ $E_s = "+"$), то тест невалідний, а саме: легкий та з малою розподільною здатністю. Це є наслідком дії декількох факторів: а) тест сконструйований із невалідних тестових завдань і повинен валідизуватися шляхом ускладнення завдань та збільшенням їхньої розподільної здатності; б) невалідна процедура тестування.

Критерій III. Якщо асиметрія та ексцес від'ємні ($A_s = "-"$ $E_s = "-"$), то тест є легким, але з великою розподільною здатністю. Валідизація полягає в ускладненні тестових завдань.

Критерій IV. Якщо асиметрія від'ємна, а ексцес відсутній ($A_s = "-"$ $E_s = "0"$), то тест є легким. Валідизація його полягає в ускладненні тестових завдань.

Критерій V. Якщо асиметрія додатна та ексцес від'ємний ($A_s = "+"$ $E_s = "-"$), то тест є складним з великою розподільною здатністю. Валідизація його полягає в усуненні легких тестових завдань.

Критерій VI. При додатній асиметрії та ексцесі ($A_s = "+"$ $E_s = "+"$) тест є складним з малою розподільною здатністю. Це є наслідком дії двох факторів, що впливають один на одний, тобто: а) порушення валідності змісту; б) невалідність тестових завдань. Валідизація тесту полягає у валідизації тестових завдань.

Критерій VII. При додатній асиметрії та відсутності ексцесу ($A_s = "+"$ $E_s = "0"$) тест є складним. Це є наслідком дії двох факторів, що впливають один на одний, тобто: а) порушення валідності змісту; б) невалідність тестових завдань. Валідизація тесту полягає у валідизації тестових завдань.

Окреме місце посідають два останні критерії VIII та IX, які дають можливість визначити конструктивну валідність тесту, а саме визначити тест на однорідність.

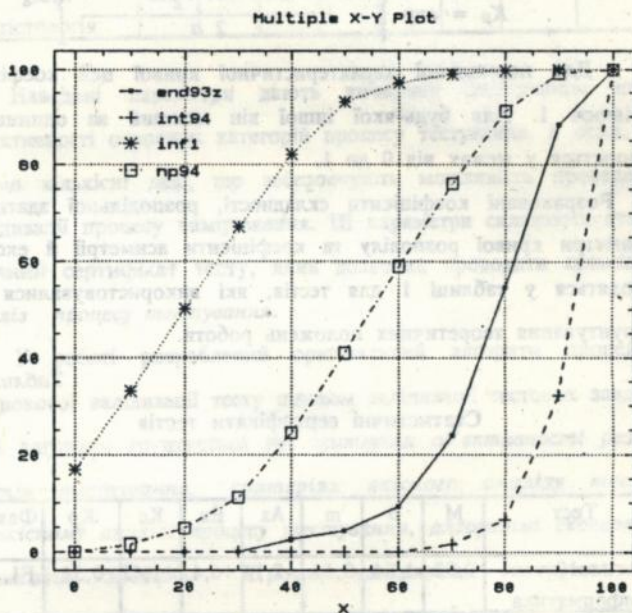
Критерій VIII. Крива розподілу симетрична, асиметрія дорівнює нулю, а ексцес додатній ($A_s = "0"$ $E_s = "+"$) — тест неоднорідний і сконструйований із двох підтестів: складного та легкого. Валідизація тесту полягає у його розщепленні на два окремих тести.

Критерій IX. Крива розподілу симетрична, асиметрія дорівнює нулю, а ексцес від'ємний ($A_s = "0"$ $E_s = "-"$) — тест неоднорідний і сконструйований з кількох (більше двох) підтестів різної складності. Валідизація тесту полягає у його розщепленні на окремі тести.

Математична модель тестів

Визначивши характеристичну криву тесту як функціональну залежність ймовірності одержати відповідь з результатом, меншим фіксованого, від значення результату, у роботі доводиться, що в аналітичному вигляді це є інтегральна функція закону розподілу $\Phi(X) = P \{ \xi < X^* \}$.

На малюнку 1 подані інтегральні функції розподілу результатів за тестами, що мають за критеріями якісного аналізу різні характеристики. Аналіз графіків дає можливість наочно оцінити характеристики тесту. Для їхнього кількісного визначення виведені формули розрахунку: K_c — коефіцієнту складності, K_p — коефіцієнту розподільної здатності, які функціонально залежать від параметрів закону розподілу M та σ .



Мал. 1 Експериментальні характеристичні криві тестів: *end93z* — легкий тест з невалідною процедурою тестування, *int94* — дуже легкий тест з невалідною процедурою тестування, *inf1* — складний тест з валідною процедурою тестування, *np94* — валідний тест з валідною процедурою тестування.

Коефіцієнт складності:

$$K_c = \frac{|\bar{X} - \frac{X_{\max} + X_{\min}}{2}|}{2\sigma} A_s \quad (1)$$

Для валідного тесту, а отже тесту з нормальною характеристичною кривою, $K_c = 0$. Додатне значення мають складні тести, від'ємне — легкі.

Коефіцієнт розподільної здатності:

$$K_p = \exp \left\{ - \frac{|\bar{X} - \frac{X_{\max} + X_{\min}}{2}|}{2\sigma} \right\} \quad (2)$$

Для нормальної характеристичної кривої цей коефіцієнт дорівнює 1. Для будь-якої іншої він менший за одиницю і змінюється у межах від 0 до 1.

Розраховані коефіцієнти складності, розподільної здатності, параметри кривої розподілу та коефіцієнти асиметрії й ексцесу наводяться у таблиці 1 для тестів, які використовувалися для обґрунтування теоретичних положень роботи.

Таблиця 1

Статистичні сертифікати тестів

Тест	M	σ	m	A_s	E_x	K_c	K_p	Фактор
ss10 інформатика	2,82	1,64	0,14	+2,14	+6,4	+1,42	0,51	F1, F2
inf1 мед.інформ.	22,1	18,15	0,76	+0,63	0	+0,48	0,46	F1
ter94 терапія	84,36	9,2	0,76	-1,05	+1,19	-1,96	0,15	F1, F2, F4
end93l ендокринологія	65,3	15,61	1,12	-0,32	-0,35	-0,15	0,61	
end93z ендокринологія	77,36	12,75	0,99	-1,6	+2,2	-1,29	0,34	F4
ner93 нерв.хвороби	63,13	22,15	1,03	-0,99	+0,56	-0,29	0,74	F1

Тест	M	σ	m	As	Ex	Kc	Kp	Фактор
пр94 нерв.хвороби	55,08	19,50	1,73	+0,34	-0,21	-0,04	0,88	
п194 нерв.хвороби	65,06	20,06	1,04	-0,66	+0,07	-0,25	0,68	F1
vb91 внутр.хвороби	80,9	15,4	0,8	-1,4	+2,47	-1,40	0,37	F1, F2, F4
vb93 внутр.хвороби	70,78	25,53	1,2	-1,2	+0,84	-0,49	0,66	F1, F2, F4
int94 інтернатвга	91,36	6,77	0,58	-1,35	+1,99	-4,12	0,05	F1, F2, F4
gic94 гістологія	60,39	10,72	0,34	-0,44	-0,46	-0,21	0,62	F1

Наведені параметри дають вичерпну інформацію щодо об'єктивності основних категорій процесу тестування, а отже, ми маємо кількісні дані, що забезпечують можливість проведення валідації процесу вимірювання. Ці параметри складають статистичний сертифікат тесту, який дозволяє проводити *кількісний аналіз процесу тестування*.

У роботі розроблений оригінальний алгоритм процедури покрокової валідації тесту шляхом валідації тестових завдань. Цей алгоритм ґрунтується на: *критеріях об'єктивності результатів тестування, критеріях якісного аналізу тестів, кількісному аналізу процесу тестування, алгоритмі експертизи тестових завдань*. Процес валідації інструментарію вимірювання проводиться методом покрокової ітерації з постійним спостереженням за змінами характеристик тесту та наближенням їх до валідних. Або, використовуючи термінологію інформатики, в основу методу валідації покладений принцип організації умовного циклу, в якому передумовою закінчення покрокової валідації системи є валідність тесту, що визначається за статистичним сертифікатом. Валідацію тесту потрібно проводити за результа-

тами пілотного тестування на репрезентативній вибірці. Для реалізації покрокової валідизації при пілотному тестуванні, тест повинен мати довжину у 1,5—2 рази більшу остаточної довжини валідного тесту, щоб при експертизі тестових завдань після кожного циклу валідизації його довжина задовольняла умовам необхідності та достатності для статистично достовірних висновків.

Розроблений алгоритм має реалізуватися на комп'ютері за принципом експертної системи, що надає можливість у діалоговому режимі з комп'ютером формувати валідний тест. Запропонований метод покрокової валідизації тесту потенційно може реалізовувати процедуру валідизації за результатами одного тестування, що звичайно скорочує шлях стандартизації тесту.

Визначаючи основні категорії процесу вимірювання рівня знань, ми з'ясували, що валідність методу залежить від валідності кожної категорії. Отже поруч з валідністю інструментарію важливе місце посідає також валідиція процедури вимірювання. Тому що невідповідність характеристик цієї категорії критеріям валідності робить також невалідним метод вимірювання, навіть якщо інструментарій атестовано як валідний.

Розроблені критерії об'єктивності результатів тестування та критерії кількісного аналізу тестів дають можливість ідентифікувати порушення валідності процедури вимірювання.

Найчастіше наслідком порушення валідності процедури вимірювання є випадки, які класифіковані за *критеріями об'єктивності результатів тестування* як II та V. У цих випадках порушення нормального закону розподілу результатів тестування пов'язані з наявністю додатнього ексцесу, що засвідчує зменшення розподільної здатності тесту. Аналіз кожного з цих випадків довіє наступне. У першому, коли результати класифікуються за критерієм II ($A_s = -$ $E_x = +$), вони є завищеними за рахунок

дії декількох факторів, у тому числі й порушення процедури тестування. Наші спостереження та аналіз доводять, що виникнути ці порушення могли внаслідок таких основних чинників: порушення вимог щодо підготовки тестування, а саме, людина, яка тестувалася, мала варіанти тестових завдань з еталонами вірних відповідей і могла користуватися ними або одержувала іншу нерегламентовану допомогу; порушення вимог щодо рандомізації тестових завдань або валідності змісту, або довжини тесту, яка, як правило, обирається меншою за статистично достатню, тощо.

У другому випадку, коли результати тестування підпадають під критерій V ($A_s = +$ $E_x = +$), вони є занижені за рахунок порушення умов проведення тестування.

Отже, на процедуру вимірювання, внаслідок чого порушується валідність цієї категорії, впливають фактори суб'єктивні та об'єктивні, внутрішні та зовнішні. Тому при стандартизації методу вимірювання необхідно чітко, однозначно, зрозуміло визначитися щодо процедури та умов вимірювання.

Принципи і методика оцінювання знань

Оцінювання — це один із завершальних етапів вимірювання рівня знань. Процедура та методика оцінювання суттєво впливають на остаточні результати, на можливість аналізу та статистично достовірних висновків, тому визначення валідності оцінювання є надзвичайно важливим етапом у процесі вимірювання.

Процедура оцінювання полягає в конвертації одержаного при тестуванні первинного результату у певну нормовану шкалу балів — оцінку, тоді як методика оцінювання полягає у визначенні алгоритму виконання цієї процедури.

Теорія шкалювання, яка покладена в основу процедури оцінювання результатів тестування, дає можливість одержати достовірні дані про рівень знань як у відносних, так і в абсолютних

одинацях. У роботі розроблені поетапно узгоджена система та загальні принципи процедури оцінювання рівня знань, виміряних при комп'ютерному тестуванні як за чотирибальною шкалою, так і за інтервальною Т-шкалою американських стандартизованих тестів.

Загальним при оцінюванні результатів тестування є вирішення запитання: *стандартизований тест чи ні?*

Стандартизовані — це тести, тестування за якими відповідає таким умовам: тест і процедура тестування є валідними та регламентованими згідно з умовами стандартизації; результати тестування підпорядковуються нормальному закону розподілу; первинний бал розраховується за загально визначеною для даного тесту формулою; конвертація первинного бала проводиться за стандартною шкалою; визначення кінцевих результатів проводиться за стандартизованою нормою у стандартизованих одиницях. Але справа в тому, що на сучасному етапі впровадження тестування в Україні, тести та процедура оцінювання є нестандартними.

Якщо тест нестандартизований, потрібно провести його стандартизацію, яка передбачає наявність інформації :

- 1) *одержані результати тестування об'єктивні чи ні,*
- 2) *яка методика підрахунку первинного бала,*
- 3) *за якою шкалою проводиться конвертація результатів.*

Методика оцінювання залежить від того — одержані результати пілотного тестування об'єктивні чи ні. Цей аналіз проводиться за *критеріями кількісного аналізу результатів*. При наявності необ'єктивності потрібно усунути чинники, які зумовили цю необ'єктивність, і тільки після повторного тестування проводити оцінювання.

Одержані при тестуванні результати первинного бала можуть бути згруповані за принципами об'єктивності та стандартизованості

таким чином: *стандартизовані тести; нестандартизовані тести, з результатами тестування, які є об'єктивними; нестандартизовані тести з результатами тестування, які є необ'єктивними.*

Одним з етапів процедури стандартизації є підрахунок первинного бала. Його значення розраховується за певною, визначеною конкретно для кожного тесту, формулою, яка враховує кількість вірних та невірних відповідей. Ця формула виконує функцію нормалізації результатів тестування з різних тестів, що полягає у тому, щоб зробити середнє значення первинного бала для вибірки пілотного тестування рівним 50: $\bar{X}\% = 50$.

При шкалюванні використовується чотири шкали: номінальна, порядкова або рангова, шкала інтервалів та шкала відношень. Тільки з величинами конвертованими за шкалами інтервалів та відношень можна проводити адитивні операції, визначати середнє значення. Шкала оцінок навчальної успішності при традиційному оцінюванні — це порядкова шкала, тому використовувати середню оцінку з математичної точки зору некоректно. Але чотирибальна шкала може бути інтервальною, якщо результати тестування конвертувати через Z-шкалу за формулою :

$$O_i = \text{int} \left\{ 4,0 + 1,0 \left(\frac{X_i - \bar{X}}{\sigma} \right) \right\} \quad (3)$$

Ця шкала названа O-шкалою. Вона є інтервальною і тому з результатами, конвертованими за нею, можна проводити адитивні операції. Її статистичною характеристикою є математичнє очікування. Початок її (X_{\min}) обирається на основі статистичних властивостей нормальної кривої розподілу із міркувань щодо кількості незадовільних оцінок, що прогножуються. У загальному вигляді X_{\min} може бути визначено, як $X_{\min} = \bar{X} - d \sigma$, де d — число, значення якого залежить від цілей тестування і

може змінюватися від 1 до 3, що звичайно позначиться на кількості незадовільних оцінок. Ця величина для даного тесту встановлює норму, яка може бути покладена в основу критерію визначення мінімально необхідного рівня знань для прийняття рішення щодо досягнення навчальної мети.

Конвертація у чотирибальну О-шкалу накладає певні обмеження на довжину тесту, при яких мінімізувалася б помилка. В основу запропонованого підходу прикладено принцип, названий *оптимальним тестуванням*. Суть його полягає у тому, що особа, яка тестується, відповідає на кількість запитань, необхідну та достатню для статистично достовірних висновків з мінімально можливою помилкою.

В основі шкалювання за Т-шкалою лежить принцип конвертації первинного бала, що розподілений за нормальним законом, також через Z-шкалу. Загальна формула такої конвертації має вигляд:

$$T_i = 50 + int \left\{ 10 \left(\frac{X_i - M}{\sigma} \right) \right\} \quad (4)$$

Конвертовані бали у Т-шкалу змінюються від 20 до 80, отже вона є шестидесятибальною. Це, звичайно, збільшує точність оцінювання у порівнянні з чотирибальною О-шкалою у 15 разів.

Перехід на багатобальну шкалу оцінок повинен проходити з урахуванням основних принципів шкалювання та можливістю визначення кінцевого результату як у балах, так і у перцентілях. Проводити цей перехід потрібно поступово в декілька етапів. Автором пропонується, як один із варіантів, такий:

- 1 етап — перехід на О-шкалу з введенням кількісного визначення нижнього граничного значення початку відліку,
- 2 етап — перехід на багатобальну Т-шкалу та її похідні.

У роботі розроблена загальна схема процесу вимірювання

рівня знань. Варто зауважити, що методика, яка пропонується, застосовується для тестів, які ще нестандартизовані. Саме її застосування дасть змогу провести стандартизацію усіх категорій процесу вимірювання та оцінювання.

Загальна схема процесу складатиметься з таких основних кроків:

Крок 1. Провести тестування та підрахувати для кожної особи, що тестується, первинний бал.

Крок 2. Проаналізувати результати тестування за *критеріями кількісного аналізу*, визначивши *статистичний сертифікат тесту*.

Можливі висновки:

А. Результати тестування є об'єктивними.

Б. Результати тестування є необ'єктивними.

А. Результати тестування є об'єктивними.

Крок 3. Визначити формулу підрахунку первинного бала $X = f(n_e, n_{ne})$ в аналітичному вигляді, користуючись *принципом вирівнювання*.

Крок 4. Провести конвертацію первинного бала у кінцевий бал та, якщо це можливо, у перцентілі за обраною шкалою, використовуючи формулу або для *О-шкали*, або для *T-шкали*.

Б. Результати тестування є необ'єктивними.

Якщо результати необ'єктивні, користуючись критеріями кількісного аналізу, потрібно з'ясувати причини, що це зумовили та усунути їх при наступних тестуваннях. Конвертацію результатів провести, виходячи з міркувань щодо статистичних характеристик закону розподілу первинного бала. Використовуючи математичні підходи, пропонується замість середнього значення у формулах конвертації застосовувати медіану Me — статистичну величину, по обидва боки від якої знаходиться однакова кількість даних вибірки. Тоді продовження схеми матиме такий вигляд:

Крок 3. Провести конвертації первинного бала у кінцевий бал за обраною шкалою, використовуючи формулу:

$$\text{для } O\text{-шкали: } O_i = \text{int} \left\{ 4,0 + 1,0 \left(\frac{X_i - Me}{\sigma} \right) \right\}$$

$$\text{для } T\text{-шкали: } T_i = 50 + \text{int} \left\{ 10 \left(\frac{X_i - Me}{\sigma} \right) \right\}$$

Методика порівняльного аналізу

Існує декілька рівнів нормованого порівняння кінцевих результатів, одержаних при вимірюванні та оцінюванні навчальної успішності: *соціальна відносна норма* — порівняння з результатами інших учасників тестування; *індивідуальна відносна норма* — порівняння результатів, одержаних раніше або з інших дисциплін для одного й того ж учасника тестування; *предметна відносна норма* — порівняння результатів кожної особи, що тестується, з навчальними цілями.

Порівняння на рівні соціальної та індивідуальної відносних норм здійснюються за результатами нормативно-групових вимірювань, теорія яких для комп'ютерного тестування успішності навчання розроблена у даній роботі. Результати учасників тестування, співставляючи з результатами тестування репрезентативної вибірки попередніх тестувань, порівнюють один з одним. Методика та умови порівняльного аналізу для цих рівнів наводяться у дисертації.

На рівні предметно відносної норми відбувається порівняння щодо виконання особою тестування певних критеріїв, у нашому випадку — досягнення навчальної мети. Цей рівень порівняльного аналізу може проводитися лише за результатами, що орієнтовані на критеріальні вимірювання. У роботі доводиться, що вимірювання, орієнтовані на критерії, ґрунтуються на нормативно-групових вимірюваннях, які забезпечують нормативну базу для введення

критеріїв щодо прийняття того чи іншого рішення про досягнення навчальної мети. Ці висновки ґрунтуються на математичних міркуваннях щодо основних закономірностей, яким повинні підпорядковуватися результати тестування, орієнтовані на групу та критерії. Нормативно-групові результати мають розподілятися за нормальним законом розподілу, а результати, орієнтовані на критерії, — за біноміальним або законом розподілу Вейбулла. В той же час, порушення валідності категорій процесу нормативно-групових вимірювань, як це доведено, призводить до викривлення нормального закону розподілу. Він стає асиметричним, що є ознакою як біноміального, так і закону розподілу Вейбулла. У роботі запропонована методика, що надає можливість відрізяти асиметрію, що пов'язана з невалідністю процесу вимірювання, від дійсної асиметрії, пов'язаної із збільшенням рівня знань у процесі навчання і досягненням навчальної мети.

У РОЗДІЛІ 5 "Управління навчальним процесом в умовах використання комп'ютерного тестового контролю" із застосуванням системного аналізу та кібернетичного підходу до управління навчальним процесом досліджується процес вимірювання знань в умовах використання комп'ютерного тестового контролю. Встановлені суттєві зміни в управлінні процесом навчання.

Застосування методу вимірювання знань, що відповідає критеріям якості, має позначитися на процесі навчання. Комп'ютерне тестування застосовувалося для вимірювань вихідних та підсумкових знань і було складовою частиною сесійних іспитів. Особливості цього методу вимірювання, у якому мінімізується вплив суб'єктивних факторів, повинні позначитися й на схемі управління системою навчання. У даному випадку для неї характерна наявність двох зворотних зв'язків, походження та функції яких різні за своєю природою та характером впливу. β_1

— основний зворотний зв'язок, пов'язаний з впливом кінцевої інформації на вихідну. Він додатний, стійкий і виконує корегуючі та адміністративні функції. β_2 — зворотний зв'язок, який є наслідком специфіки методу вимірювання та оцінювання знань — комп'ютерного тестування. За класифікацією це також додатний зворотний зв'язок зі стійкою дією. Впливаючи безпосередньо на об'єкт управління, він призводить до суттєвих його змін.

Дидактичний експеримент засвідчив, що у такій системі змінюються як суб'єкти, так і об'єкт навчання. Загальну систему було змодельовано трьома підсистемами з різними об'єктами управління: 1) той, хто навчається; 2) той, хто навчає; 3) форма подання знань. Проаналізовано функціонування кожної підсистеми і виявлено такі особливості.

Підсистема з об'єктом управління «той, хто навчається».

Метою управління цією підсистемою є досягнення певного кінцевого рівня знань. Встановлені зміни у мотивації навчання, що відбуваються через дію зворотнього зв'язку β_2 . Той, хто навчається, адаптується під систему контролю і, оскільки вона забезпечує підпорядкованість результатів вимірювання статистичним закономірностям, це позитивно позначається на його навчанні.

Підсистема з об'єктом управління «той, хто навчає».

Метою управління у цій підсистемі є зміни професійної діяльності викладача. Це пов'язано з тим, що уніфікована форма контролю знань вимагає й уніфікованої форми подання інформації. Застосування комп'ютерного тестування з автоматизованою обробкою даних, яка відбувається під час іспиту, тобто без затримання у часі, та можливість проведення достовірного статистичного аналізу його результатів забезпечує візуалізацію діяльності кожного викладача. Отже, праця кожного може бути об'єктивно оцінена через результати об'єктивного вимірювання знань тих, кого вони

навчають.

Підсистема з об'єктом управління формою подання інформації. Метою управління у цій підсистемі є зміни змісту, форми подання та збереження інформації. Це викликано впровадженням інформаційних технологій навчання, що реалізуються із застосуванням комп'ютерів. Застосування комп'ютерного тестування, його вплив на структурні зміни у форми подання та збереження інформації є першим кроком на цьому шляху і виявляються у більш чіткому визначенні предметних понять та їх структурованості, алгоритмованості зв'язків предметних понять, зменшенні фонові інформації з малою інформативністю, змінах у формі фіксації та збереженні вимірних результатів навчальної успішності тощо.

Отже, в усіх підсистемах системи навчання мають місце певні зміни. Іншими словами вона є адаптивною системою, яка змінює своє функціонування в залежності від особливостей керуючого елемента. За класифікацією — це *параметрична адаптація*, при якій змінюються характеристики підсистем: *мотиваційний елемент навчання у того, хто навчається; активізація професійної діяльності викладача; збільшення інформаційної значущості, обсягу інформації, яку повинні засвоювати спочатку викладач, а потім студент; формою збереження результатів.*

У роботі за ієрархією прийняття рішень, ступенем відповідальності детально проаналізовані та конкретизовані завдання, що постають при впровадженні комп'ютерних технологій взагалі та комп'ютерного тестування зокрема, визначений пріоритет їх виконання. А саме досліджені: роль та дії адміністрації щодо забезпечення цього процесу, технічне оснащення, підготовка викладацького складу, розробка програмного забезпечення, дидактична та методична підтримка, розробка тестів з різних дисциплін, організація процесу стандартизації тестів.

Підсумкові висновки

1. Впровадження новітніх технологій навчання, що ґрунтуються на нових підходах щодо подання та засвоєння знань, потребує нових сучасних методів їх вимірювання та оцінювання. Застосування традиційних методів вимірювання знань: усного опитування та чотирибальної системи їх оцінювання гальмує не лише інтегрування випускників медичних вузів у міжнародну спільноту практикуючих лікарів, а й процес інформатизації освіти взагалі.

2. Вимірювання якісних характеристик, до яких належить рівень знань, має відповідати загальним вимогам теорії вимірів, найважливішими з яких є критерії: *об'єктивності, надійності, валідності й точності*. Порівняльний аналіз методів вимірювання рівня знань, що використовуються в сучасній педагогіці, свідчить про те, що метод тестування найбільше відповідає критеріям якості при визначенні рівня теоретичних знань. Тоді як усна форма контролю знань, як педагогічного методу діагностики, ефективно може використовуватися при вимірюванні знань, пов'язаних з мовним розвитком або мовними умовами спілкування. Форма письмового контролю має використовуватися при вимірюванні знань, пов'язаних із визначенням когнітивного розвитку та проблемного мислення. Співбесіда у формі інтерв'ю — для одержання інформації щодо особистості респондента: його зовнішнього вигляду, поведінки, мови тощо.

3. Статистичний аналіз результатів іспитів з 43-х дисциплін, як результат вимірювання рівня підсумкових знань, доводить, що оцінка, конвертована за чотирибальною шкалою, як випадкова величина, розподілена за нормальним законом. Отже, використання

процесу тестування, орієнтованого на норми, як методу вимірювання й оцінювання рівня вихідних та підсумкових знань, є науково обґрунтованим. Такі вимірювання є підґрунтям створення нормативної бази — нормативних інтервалів — для впровадження вимірювань, орієнтованих на критерії, і прийняття рішення щодо досягнення навчальної мети.

4. Категоріальний аналіз процесу вимірювання рівня знань свідчить, що основними категоріями процесу комп'ютерного тестування є: *інструментом вимірювання* — тест, складений із тестових завдань; *процедурою вимірювання* — тестування на комп'ютері; *процедурою оцінювання* — шкалювання. Аналіз усіх категорій процесу вимірювання рівня знань ґрунтується на визначальному понятті — «валідність», яке на основі системного аналізу диференційовано за категоріальною та функціональною ознаками на: валідність методу, валідність інструментарію, валідність процедури тестування, валідність процедури оцінювання.

5. Аналіз валідності інструментарію пропонується проводити на двох рівнях: апіорному та апостеріорному. На апіорному рівні проводиться експертиза тестових завдань. На апостеріорному — аналіз валідності результатів тестування, що забезпечує експертизу тесту на основі теорії тестів.

6. Згідно з розробленою теорією рівень підсумкових знань X , виміряний при комп'ютерному тестуванні, розподілений за нормальним законом. Відхилення результатів тестування від нормального закону розподілу, що визначаються за коефіцієнтами асиметрії та ексцесу, пов'язані з порушенням валідності основних категорій процесу тестування. Цей висновок покладено в основу сформульованих автором *критеріїв об'єктивності результатів тестування та критеріїв якісного аналізу тестів*.

7. Розроблена математична модель тестів, в основі якої лежить

визначення поняття характеристичної кривої тесту як інтегральної функції розподілу, надала можливість ввести статистичний сертифікат тесту — інтегровані параметри категорій процесу тестування, що є інструментарієм кількісного аналізу результатів тестування. За його допомогою можна на кількісному рівні проводити валідизацію процесу тестування, використовуючи оригінальний алгоритм процедури покрокової валідизації тесту шляхом валідизації тестових завдань.

8. Оцінювання одержаних при тестуванні результатів потрібно проводити з урахуванням основних принципів теорії шкалювання. Де визначено, що тільки з величинами, конвертованими за шкалами інтервалів та відношень, можна проводити адитивні операції, визначати середнє значення. Чотирибальна шкала може бути інтервальною, якщо результати тестування конвертувати за визначеною формулою для введеної О-шкали. Для неї статистично обгрунтовані початок відліку та масштаб. Запропоновано та статистично обгрунтовано введення норми, яка може бути покладена в основу критерію визначення мінімально необхідного рівня знань для прийняття рішення щодо досягнення навчальної мети.

9. Перехід на багатобальну шкалу оцінок повинен проходити з урахуванням основних принципів шкалювання та можливістю визначення кінцевого результату як у балах, так і у перцентілях. Проводити цей перехід потрібно поступово в декілька етапів:

- 1 етап — перехід на О-шкалу з введенням кількісного визначення нижнього граничного значення початку відліку,
- 2 етап — перехід на багатобальну Т-шкалу або її похідні.

10. Порівняння кінцевих результатів, одержаних при вимірюванні та оцінюванні навчальної успішності може, проводитися на трьох рівнях: соціальна відносна норма, індивідуальна відносна норма, предметна відносна норма. Порівняння на перших

двох рівнях здійснюються за результатами нормативно-групових вимірювань за розробленою у дисертації методикою порівняльного аналізу.

На рівні предметно відносної норми відбувається порівняння щодо виконання особою тестування певних критеріїв, в нашому випадку — досягнення навчальної мети. Воно здійснюється за результатами критеріальних вимірювань. Чисельне значення критеріїв може бути визначено на основі загальної теорії вимірювання успішності навчання, розробка якої планується у подальших дослідженнях.

11. Використання комп'ютерного тестового контролю призводить до *параметричної адаптації* системи навчання, викликаючи зміни характеристик її підсистем: мотиваційного елемента навчання у того, хто навчається; активізації та вдосконалення професійної діяльності у того, хто навчає; збільшення інформаційної значущості та обсягу інформації, яку повинні засвоювати спочатку викладач, а потім студент.

12. Процес впровадження комп'ютерного тестування можна поділити на такі етапи: *підготовчий, апробаційний, етап проведення тестування, етап після тестування*. Впровадження комп'ютерних технологій у навчальний процес це кропітка, довготривала робота усього викладацького колективу навчального закладу.

Основні матеріали дисертації
викладено у роботах:

I. Монографія.

1. Комп'ютерна діагностика навчальної успішності.— К.: ЦМК МОЗ України, 1995.— 14,1 др.арк.

II. Брошури.

2. Теорія комп'ютерного тестування.— К.: ЦМК МОЗ України, 1994.— 2,58 др.арк.

3. Історія розвитку та сучасний стан педагогічної тестології.— К.: ЦМК МОЗ України, 1994. — 0,86 др.арк.

4. Тестологія як методологічна основа вимірювання знань.— К.: ЦМК МОЗ України, 1994.— 1,12 др.арк.

5. Кількісна модель валідазації процесу вимірювання знань.— К.: ЦМК МОЗ України, 1994.— 2,75 др.арк.

6. Загальні принципи оцінювання знань.— К.: ЦМК МОЗ України, 1994.— 1,21 др.арк.

III. Методичні рекомендації.

7. Методические указания по внедрению автоматизированного контроля уровня знаний студентов с применением программируемых микрокалькуляторов.— К.: МЗ УССР КМИ, 1987.— 0,86 др.арк.— У співавт.

8. Применение методов математической статистики при решении медико-биологических задач.— К.: МЗ УССР КМИ, 1988.— 2,83 др.арк.— У співавт.

9. Использование программируемых микрокалькуляторов для математической обработки медико-биологической информации: Методические разработки для самостоятельной работы студентов и слушателей ФПК медицинских институтов.— К.: МЗ УССР КМИ, 1987.— 2,79 др.арк.— У співавт.

10. Использование программируемых микрокалькуляторов в учебном процессе: Пакет обучающих и контролирующих программ.— К.: МЗ УССР КМИ, 1987.— 1,86 др.арк.— У співавт.

11. Учебные задания для программированного контроля знаний с использованием микро-ЭВМ по теме "Физические методы лечения".— К.: МЗ УССР КМИ, 1988.— 1,86 др.арк.— У співавт.

IV. Авторське свідотство.

12. Способ прогнозирования кровотечения после удаления зубов /Математическая модель: Положительное решение по заявке № 5044805 от 31.05.1993.— У співавт.

V. Наукові статті, матеріали та тези конференцій.

13. Критерий оценки функциональной взаимосвязи между

отдельными признаками заболевания //Врачебное дело.— 1989.— № 4.— 0,25 др.арк.— У співавт.

14. Комп'ютеризована методика вимірювання рівня знань.— Київ, 1995.— 0,95 др.арк.— Деп. у ГНТБ України, 28.02.1995, №541.

15. Новітні комп'ютерні технології навчання та передумови їх впровадження.— Київ, 1995.— 0,67 др.арк.— Деп. у ГНТБ України, 28.02.1995, №542.

16. Метод контроля уровня знаний студентов с применением программируемых микрокалькуляторов типа "Электроника БЗ-34" //Материалы координационного совещания заведующих кафедрами медицинской и биологической физики медвузов Украины.— Ворошиловград, 1987.— 0,1 др.арк.— У співавт.

17. Оптимизация контроля уровня знаний студентов с помощью программируемых микрокалькуляторов //Социально-психологические и педагогические аспекты формирования здорового образа жизни: Материалы региональной научно-практической конференции.— Новосибирск, 1988.— 0,1 др.арк.— У співавт.

18. Метод вычисления клинических показателей оценки тяжести заболевания //Социально-психологические и педагогические аспекты формирования здорового образа жизни: Материалы региональной научно-практической конференции.— Новосибирск, 1988.— 0,1 др.арк.— У співавт.

19. О вероятностных алгоритмах диагностики заболеваний //Социально-психологические и педагогические аспекты формирования здорового образа жизни: Материалы региональной научно-практической конференции.— Новосибирск, 1988.— 0,1 др.арк.— У співавт.

20. О новом методе оценки взаимосвязи между признаками заболеваний //V Всесоюзный съезд геронтологов и гериатров: Тезисы докладов.— Тбилиси, 1988. (співавтор - 0,1 др.арк.)

21. Статистический анализ параметров возраста в геронтологии //V Всесоюзный съезд геронтологов и гериатров: Тезисы докладов.— Тбилиси, 1988.— 0,1 др.арк.— У співавт.

22. Априорный прогноз динамики заболевания //Новые физические методы в медицине.— Ворошиловоград, 1990.— 0,1 др.арк.— У співавт.

23. Структурная фармакинетическая модель со случайными параметрами //Новые физические методы в медицине. — Ворошиловоград, 1990.— 0,1 др.арк.— У співавт.

24. Математическое прогнозирование заболеваний //Новые физические методы в медицине.— Ворошиловоград, 1990.— 0,1 др.арк.— У співавт.

25. ЭВМ в оптимизации учебного процесса в медвузах //Актуальные психолого-педагогические проблемы медицинского образования в Украинской ССР: Тезисы докладов 2-ой Республиканской конференции.— Ивано-Франковск, 1990.— 0,1 др.арк.— У співавт.

26. Об опыте преподавания основ вычислительной техники и математического моделирования аспирантам и соискателям медицинского профиля //Актуальные психолого-педагогические проблемы медицинского образования в Украинской ССР: Тезисы докладов 2-ой Республиканской конференции.— Ивано-Франковск, 1990.— 0,1 др.арк.— У співавт.

27. Компьютеризация экзаменационного контроля //Актуальные психолого-педагогические проблемы медицинского образования в Украинской ССР: Тезисы докладов 2-ой Республиканской конференции.— Ивано-Франковск, 1990.— 0,1 др.арк.— У співавт.

28. Опыт компьютеризации экзаменационного контроля //Современные проблемы подготовки медицинских кадров: Тезисы научно-методической конферен.— Москва, 1990, Часть III.— 0,1 др.арк.— У співавт.

29. Перспективи нових форм навчання у медичному вузі //Застосування комп'ютерної техніки в учбовому процесі медичних вузів.— Дніпропетровськ, 1991.— 0,1 др.арк.— У співавт.

30. Про організацію комп'ютерного іспиту у Київському медичному інституті //Застосування комп'ютерної техніки в учбовому процесі медичних вузів.— Дніпропетровськ, 1991.— 0,1 др.арк.— У співавт.

31. Проблеми впровадження тестового контролю знань студентів у КМІ //Критерії, методи та засоби об'єктивізації контролю якості підготовки спеціалістів-лікарів медико-профілактичних профілей.— Київ, 1992.— 0,1 др.арк.— У співавт.

32. Про комп'ютерне тестування слухачів ФПКв //Застосування комп'ютерної техніки в процесі підготовки медичних кадрів: Матеріали конференції.— Вінниця, 1992.— 0,1 др.арк.— У співавт.

33. Про комп'ютерні технології у навчальному процесі медичних вузів //Розробка та впровадження в процес підготовки медичних кадрів сучасних технологій навчання: Тези доповідей наукової учбово-методичної конферен.— Тернопіль, 1992.— 0,1 др.арк.— У співавт.

34. Використання експертних систем у підготовці медичних кадрів //Розробка та впровадження в процес підготовки медичних кадрів сучасних технологій навчання: Тези доповідей наукової учбово-методичної конферен.— Тернопіль, 1992.— 0,1 др.арк.— У співавт.

35. Про новітні технології навчання //Апробація та впровадження педагогічних інновацій в процесі підготовки медичних кадрів.— Київ, 1993.— 0,1 др.арк.— У спіавт.

36. ЕОМ — як засіб активізації та оптимізації навчального процесу у підготовці фахівців медиків УДМУ //Застосування комп'ютерної техніки в учбовому процесі медичних та фармацевтичного вузів МОЗ України: Тези республіканської науково-методичної конференції.— Полтава, 1993.— 0,1 др.арк.— У спіавт.

37. Аналіз якості тестів та результатів тестування //Застосування комп'ютерної техніки в учбовому процесі медичних та фармацевтичного вузів МОЗ України: Тези республіканської науково-методичної конференції.— Донецьк, 1994.— 0,1 др.арк.— У спіавт.

38. Управління навчальним процесом і комп'ютерний тестовий контроль //Модернізація системи освіти в Україні на засадах національних традицій, гуманізації та демократизації, світового досвіду: Матеріали Всеукраїнської наук.-практ. конф., Ч.5.— Ів.-Франківськ: "Плай", 1995.— 0,2 др.арк.

Загальна кількість наукових праць, у яких викладений основний зміст роботи, 44.

Булах И.Е. Теория и методика компьютерного тестирования успеваемости (на материалах медицинских учебных заведений).

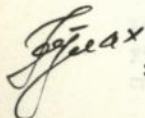
Диссертация в виде рукописи на соискание ученой степени доктора педагогических наук по специальности 13.00.01. — теория и история педагогики. Институт педагогики и психологии профессионального образования Академии педагогических наук Украины, Киев, 1995.

Диссертация содержит теоретическое исследование основных закономерностей процесса измерения уровня знаний методом компьютерного тестирования. Установлена взаимозависимость результатов от характеристик основных категорий процесса измерения. Создан инструментарий количественного анализа результатов тестирования — статистический сертификат тестов, который основывается на разработанных: теории компьютерного тестирования успеваемости, критериях количественного анализа основных категорий процесса измерения, математической модели тестов. Разработаны принципы оценивания уровня знаний, введены нормативные шкалы оценивания и критерии принятия решения о достижении учебной цели. Разработаны практические рекомендации по введению компьютерного тестирования в учебный процес, компьютерное программное обеспечение.

Bulakh I.E. Theory and Methods of Computer Testing of Study Progress (Based on the Materials of Medical Educational Institutions).

The Thesis for conferring a scientific degree of Doctor of Pedagogical Science, speciality 13.00.01 — Theory and History of Pedagogy. Institute of Pedagogy and Psychology of Professional Education of the Academy of Pedagogical Science of Ukraine, Kiev, 1995.

The thesis contains the theoretical research of the main regularities of the progress of knowledge level measurement by computer testing. The interdependence of results of the measurement on characteristics of its main categories has been founded. The instrumentation of quantitative analysis of test results has been created - the statistical certificate of a test which based on: the theory of computer testing; the criteria of quantitative analysis of main categories of measuring process; the mathematical model of tests. The principles of knowledge level estimation have been developed; the normative estimation scales and criteria of decision making on the learning aim achieved are introducing. The practical recommendation on introduction of the computer testing in a teaching process and computer software have been developed.



Ключові слова: вимірювання, оцінювання, тестування, рівень знань, валідизація, стандартизація тестів, нормовані шкали.

ДЛЯ НОТАТОК

Имя, по которому выданы эти документы

№ документа

О.О.О. "АВИА", 1980г. №100, стр. 100

4154616

AB 32.983

AB 32.983

