

**П.В. МИКИТЕНКО  
О.В. ГАЛИЦЬКИЙ  
І.І. КУЧЕРЕНКО**

**КОМП'ЮТЕРНО-ОРІЄНТОВАНІ ТЕХНОЛОГІЇ ОСВІТНІХ  
ВИМІРЮВАНЬ**

**Київ 2019**

**УДК 378.016:[004:37]**

Монографію присвячено проблемам методики організації педагогічної діагностики та контролю знань з використанням засобів комп'ютерно-орієнтованих технологій. Проаналізовано роль комп'ютерно-орієнтованих технологій освітніх вимірювань у навчально-виховному процесі, розглянуто передумови комп'ютерного супроводу контролю якості системи навчання у ЗВО. Розкрито концептуальні положення методів реалізації комп'ютерної педагогічної діагностики якості знань студентів та запропоновано педагогічну модель комплексу комп'ютерно-орієнтованих технологій освітніх вимірювань. Детально розглянуто етапи визначення основних психометричних характеристик тестових завдань, а також описані методи опрацювання та інтерпретації результатів тестування на основі оцінювання законів розподілу ймовірностей випадкових величин.

Для аспірантів, викладачів та наукових співробітників закладів вищої освіти.

Рецензенти:

доктор педагогічних наук, професор кафедри медичної і біологічної фізики та інформатики Національного медичного університету імені О.О. Богомольця Н.В. Стучинська,

кандидат фізико-математичних наук, старший науковий співробітник відділу Обчислювальної математики НАН України Д.О. Ситник

Науково-видавничий відділ фізико-математичної та технічної літератури

Редактор А.О. Микитенко

ISBN 978-966-00-1706-1

© П.В. Микитенко,  
О.В. Галицький,  
І.І. Кучеренко, 2019

## ЗМІСТ

<b>ПЕРЕДМОВА.....</b>	<b>5</b>
<b>РОЗДІЛ 1. ТЕОРЕТИЧНІ ЗАСАДИ ВИКОРИСТАННЯ КОМП'ЮТЕРНО-ОРІЄНТОВАНИХ ТЕХНОЛОГІЙ ОСВІТНІХ ВИМІРЮВАНЬ .....</b>	<b>6</b>
1.1. Освітні вимірювання як засіб діагностики та прогнозування якості знань .....	6
1.1.1. Історико-генезисні проблеми педагогічної діагностики знань .....	8
1.1.2. Психолого-педагогічні засади діагностики та контролю знань .....	10
1.1.3. Функції та завдання педагогічної діагностики.....	17
1.2. Комп'ютерно-орієнтовані технології освітніх вимірювань як педагогічна проблема.....	20
1.2.1. Передумови комп'ютерного супроводу контролю якості знань у вищому навчальному закладі .....	22
1.2.2. Загальна характеристика комп'ютерно-орієнтованих технологій освітніх вимірювань .....	27
<b>РОЗДІЛ 2. МЕТОДИЧНІ ЗАСАДИ ЗАСТОСУВАННЯ КОМП'ЮТЕРНО-ОРІЄНТОВАНИХ ТЕХНОЛОГІЙ ОСВІТНІХ ВИМІРЮВАНЬ .....</b>	<b>33</b>
2.1. Сутність і специфіка методів і методик комп'ютерної педагогічної діагностики знань .....	33
2.2. Комплекс комп'ютерно-орієнтованих технологій освітніх вимірювань.....	42
2.2.1. Основи використання комп'ютерно-орієнтованих технологій освітніх вимірювань .....	45
2.2.2. Реалізація адаптивного тестування засобами комп'ютерно-орієнтованих технологій освітніх вимірювань .....	69
2.2.3. Аналіз якості показників тестових матеріалів з використанням комп'ютерно-орієнтованих технологій.....	75
2.3. Педагогічна модель комплексу комп'ютерно-орієнтованих технологій освітніх вимірювань .....	88
<b>РОЗДІЛ 3. АНАЛІЗ РЕЗУЛЬТАТІВ ТЕСТОВИХ МАТЕРІАЛІВ .....</b>	<b>96</b>
3.1. Первинний аналіз результатів тестування .....	96
3.1. Розподіл балів результатів тестування .....	104
3.3.1. Перевірка нормальності розподілу балів результатів тестування .....	104
3.3.2. Критерій $\chi^2$ (хі-квадрат) .....	106
3.3.3. Критерій Шапіро-Вілка.....	109
3.4. Основні характеристики тестових завдань.....	111
3.4.1. Дискримінативність тестових завдань.....	111
3.4.2. Складність тестових завдань.....	117
3.5. Валідність результатів тестування .....	119
3.5.1. Перевірка валідності результатів тестуванням за допомогою .....	

кореляційного аналізу .....	123
3.6. Надійність результатів тестування .....	133
3.7. Експертне дослідження використання вступного тесту з іноземної мови до магістратури .....	137
3.7.1. Розробка інструментарію для тесту .....	137
3.7.2. Апробація та оцінка тестових завдань .....	141
<b>СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ.....</b>	<b>149</b>
<b>ДОДАТКИ .....</b>	<b>154</b>
<b>Додаток 1 .....</b>	<b>154</b>
<b>Додаток 2 .....</b>	<b>156</b>

## ПЕРЕДМОВА

Однією з найважливіших передумов, яка впливає на якісну підготовку майбутнього фахівця в закладі вищої освіти (ЗВО), є організація і управління повноцінною навчально-пізнавальною діяльністю студентів, направленою на засвоєння системи знань, умінь і навичок, оволодіння досвідом самостійної діяльності. У сукупності засобів, що забезпечують функціонування системи управління якістю підготовки фахівців, важливу роль відіграє науково обґрунтована, ретельно спланована та раціонально організована діагностика навчального процесу та результатів навчання студентів. Це все спонукає до пошуку нових педагогічних підходів удосконалення фахової підготовки студентів, а також розроблення систем діагностики якості знань і коригування освітнього процесу.

Вимоги до підвищення якості підготовки випускників педагогічних ЗВО зумовлюють необхідність продовження пошуку нетрадиційних підходів до організації діагностики та контролю знань, умінь і навичок студентів з використанням сучасних комп'ютерно-орієнтованих технологій.

Згідно з Законом України "Про національну програму інформатизації" вагома роль у розвитку сучасного суспільства належить інформатизації всіх сфер діяльності людей, зокрема і в системі освіти [32].

Сучасний стан системи освіти зобов'язує педагогів шукати нові підходи до оцінювання навчальних досягнень студентів. Автоматизація систем оцінювання освітнього процесу дає змогу детальніше та глибше визначити ті ланки в навчальному процесі, які потребують якісних змін і дистанційного керування діяльністю всіх учасників освітнього процесу. На базі інформатизації мають формуватися та розвиватися інтелектуальний потенціал нації, вдосконалюватися форми та зміст навчання, а також впроваджуватися комп'ютерно-орієнтовані методичні системи навчання. Комп'ютерно-орієнтовані системи навчання повинні ґрунтуватись на новітніх педагогічних та інформаційно-комунікаційних технологіях, використання яких може забезпечити створення в навчальному закладі єдиного освітньо-наукового інформаційного середовища.

## РОЗДІЛ 1. ТЕОРЕТИЧНІ ЗАСАДИ ВИКОРИСТАННЯ КОМП'ЮТЕРНО- ОРІЄНТОВАНИХ ТЕХНОЛОГІЙ ОСВІТНІХ ВИМІРЮВАНЬ

### 1.1. Освітні вимірювання як засіб діагностики та прогнозування якості знань

Освітні вимірювання - це складова освітньої діяльності, сутність якої полягає у визначенні рівня відповідності результатів навчально-виховної роботи встановленим освітнім стандартам за допомогою тієї чи іншої системи вимірювання. У Фізичному енциклопедичному словнику надається таке визначення поняття вимірювання: "... вимірювання – це послідовність експериментальних та обчислювальних операцій, здійснюваних з метою знаходження значення фізичної величини, що характеризує певний об'єкт або явище". "Вимірювання навчальних досягнень" учасників освітнього процесу можна назвати освітніми вимірюваннями. Останні передбачають використання засобів педагогічної діагностики (педагогічних тестів) не лише в навчально-виховному процесі, а й для порівняння та прогнозування якості знань цілих освітніх систем.

Теорія і методологія освітніх вимірювань розвивалися в руслі загальної теорії вимірювань, а також під впливом досліджень у психології та педагогіці. Умовно періодом виникнення теорії освітніх вимірювань вважають кінець XIX ст. - початок XX ст., коли в центрі уваги вчених знаходилися філософські та математичні проблеми вимірювань.

Серед найвагоміших праць цього періоду можна назвати теорію градації інтелектуальних здібностей, створену А. Біне, а також дослідження Ф. Гальтона та Дж. Кеттела, які вже в 1890 році використовували термін "mental test" (тест на інтелект) і ввели в науковий обіг спеціальну невід'ємну характеристику тесту, що вказує на ступінь диференціації піддослідних за інтелектуальними здібностями. Основи освітніх вимірювань заклав Ч. Спірмен, поєднавши елементи теорії фізичних вимірювань, кореляційний аналіз і накопичений психологами досвід з оцінювання здібностей студентів. З початку 1920-х років інтенсивного розвитку набула класична теорія освітніх вимірювань. З точки зору методологічних властивостей найефективнішим засобом освітніх вимірювань є тестування. Під засобом вимірювання традиційно розуміють технічний засіб, який використовується в ході вимірювань і має нормовані метрологічні властивості. Тестування – це метод вимірювання певних властивостей особистості за допомогою тесту [66]. З англійської мови "test" перекладається як випробування, перевірка. У Великому тлумачному словнику сучасної української мови поняття тест в освіті трактується як система формалізованих завдань, призначених для встановлення кваліфікаційного рівня особи. Педагогічне тестування як

форма оцінювання знань студентів базується на застосуванні педагогічних тестів, які є стандартними засобами вимірювання рівня навчальних досягнень.

Істотним є впровадження тестування в освітній процес на шляху розвитку методики діагностики та контролю рівня засвоєних знань студентів. Це дає змогу здійснити плавний перехід від суб'єктивного оцінювання до об'єктивних обґрунтованих методів оцінювання результатів навчання. Однак, як і будь-яке інше педагогічне нововведення, цей крок повинен здійснюватися на строго науковій базі, спираючись на результати педагогічних експериментів і наукових досліджень. Тестування не повинно замінити традиційні методи педагогічного контролю та діагностики, а має лише деякою мірою доповнити їх. Питання теорії та методології педагогічного контролю й діагностики на базі тестової технології розглядали такі науковці, як В.С. Аванесов [16], В.П. Безпалько [18], І.Є. Булах [23], О.А. Раков [63], В.П. Сергієнко [65] тощо.

В.С. Аванесов виділяє такі основні переваги тестового методу [15]:

1. Висока наукова обґрунтованість тесту, що дає змогу отримувати об'єктивні оцінки рівня підготовленості випробуваних.
2. Технологічність тестових методів.
3. Точність вимірювань.
4. Наявність однакових для всіх користувачів правил проведення педагогічного контролю й адекватної інтерпретації тестових результатів.
5. Сполучуваність тестової технології з іншими сучасними освітніми технологіями.

Діагностичний тест – це інструмент оцінювання, який виявляє рівень підготовки, а також "сильні" й "слабкі" сторони окремих груп студентів. Тестування є одним з інструментів педагогічної діагностики та прогнозування якості знань.

Якість знань як зазначено в [29] – це співвіднесення видів знань (закони, теорії, прикладні, методологічні, оцінювальні знання) з елементами змісту освіти та з рівнями засвоєння.

Якості знань властиві такі характеристики:

- повнота – кількість знань про об'єкт вивчення;
- глибина – сукупність осмислених студентами зв'язків і відношень між знаннями;
- систематичність – осмислення складу певної сукупності знань у їхніх ієрархічних і послідовних зв'язках;
- системність – осмислення студентом місця знання в структурі наукової теорії;
- оперативність – вміння користуватися знаннями в подібних ситуаціях;

- гнучкість – вміння самостійно знаходити варіативні способи застосування знань у змінених умовах;
- конкретність – вміння розкласти знання на елементи;
- узагальненість – вміння виразити конкретне знання в узагальненій формі.

З розвитком психології та педагогіки з'явилась необхідність у введенні не тільки якісних, а й кількісних оцінок для величин, що відрізняються за ступенем прояву тієї чи іншої властивості [46]. Головні компоненти педагогічної системи, які використовуються в освітніх вимірюваннях, – це педагогічна діагностика та контроль (вони є комплексом взаємопов'язаних складових управління освітнім процесом), ці поняття хоч і близькі, однак різні за своєю метою та спрямованістю отриманих результатів.

### **1.1.1. Історико-генезисні проблеми педагогічної діагностики знань**

Педагогічна діагностика – це один із важливих елементів управління та забезпечення педагогічного процесу як в цілому у сфері освіти, так і безпосередньо у вищому педагогічному навчальному закладі. Її роль зумовлена тим, що якість проведеної роботи з виявлення та розпізнавання індивідуальних, групових, педагогічних і психологічних особливостей студентів, вивчення, аналізу й оцінювання педагогічних явищ і фактів прямо та безпосередньо впливає на ефективність підготовки майбутніх учителів.

У вітчизняній та зарубіжній педагогіках розглядаються питання, які стосуються як вищих навчальних закладів, так і загальноосвітніх шкіл, вони є важливими для теорії та практики педагогічної діагностики у вищому педагогічному навчальному закладі.

У цілому результати аналізу свідчать, що елементи педагогічного діагностування застосовувалися в усіх педагогічних системах з найдавніших часів до наших днів для виявлення і врахування особливостей учасників педагогічного процесу, професійного добору студентів, оцінювання їх діяльності, професійного та особистісного розвитку, поточного та підсумкового контролю результатів навчання і виховання. Саме такий сенс буде вкладатися надалі в поняття *педагогічної діагностики*.

Термін "діагностика" (від грецьких слів *dia* – між, через, та *gnosis* – знання) означає процес розпізнавання чи визначення стану. У науковій літературі можна зустріти різні тлумачення цього поняття. Але в усіх випадках під словом "діагностика" розуміють таку діяльність, яка здійснюється на підставі отриманого знання про людину, колектив, явища, процеси і містить в собі відповідний аналіз і рекомендації. Методи педагогічної діагностики емпірично формуються в ході розвитку



педагогічної практики людства. Водночас педагогічна діагностика протягом тривалого часу мала досить суб'єктивний і несистематизований характер. Поняття "педагогічна діагностика" в літературних джерелах з'явилося в 1960-ті роки, хоча по суті педагогічна діагностика як практична, емпірична діяльність має такий вік, як і вся педагогічна діяльність. Вже на зорі історії люди на власному досвіді переконалися в тому, що не можна успішно виховувати і навчати молодь, не намагаючись визначити результати своїх зусиль і рівень вихованості та навченості учнів. Кінцева мета діагнозу мала важливе соціальне значення і полягала в стабілізації суспільного життя та забезпеченні наступності поколінь [38]. У міру розвитку педагогічної практики і теорії розширювався методологічний та методичний арсенал пошуку шляхів і засобів оцінювання та аналізу навчально-виховного процесу. Значним внеском у становлення і розвиток наукової педагогічної діагностики була діяльність Я.А. Коменського, Дж. Локка, Д. Дідро, Ж-Ж. Руссо, Ф. Дістервега, М.В. Ломоносова, В.Ф. Одоєвського, К.Д. Ушинського та інші. У XVI – середині XIX ст. сформувались методи педагогічного оцінювання, базовані на перевірці знань учнів у усній та письмовій формах. Застосовувані при цьому традиційні методи (опитування, контрольні роботи, іспити) потребували від учнів відтворення навчального матеріалу (як правило, у вигляді його фрагментів) або виконання певних завдань.

Чеський педагог Я.А. Коменський (XVII ст.) вважається основоположником дидактики і досі діючої системи навчання. На його думку, необхідно ретельно стежити за успішністю кожного учня. У праці "Закони добре організованої школи" проблеми педагогічної діагностики розглядаються в окремому розділі "Закони для випробувань". У ній рекомендується кілька типів випробувань: годинні, денні, тижневі, місячні, триместрові та річні. Годинні випробування проводив учитель за освоєним домашнім матеріалом і за новим після пояснення. Денні вів діскуріон (десятник), який щодня оцінював, як учні засвоїли урок. Суть тижневих випробувань полягала в тому, що один школяр міг викликати на змагання іншого – на звання кращого учня.

К.Д. Ушинський цілісно обґрунтував концепцію вивчення особистості в педагогічному процесі. До найбільш реалізованих у сучасній педагогічній діагностиці можна віднести такі проблеми та ідеї:

- місце психології в педагогічній практиці;
- важливість прогнозування вихованості та навченості;
- мета виховання як важливий об'єкт вивчення під час оцінювання;
- роль фактів поведінки та особливостей їх вивчення в психолого-педагогічному розпізнаванні.

Починаючи з кінця XIX ст. рекомендації, пов'язані з діагностикою в

педагогічному процесі, стають дедалі більш науковими. Особливо цьому сприяли активний розвиток педагогічної психології на початку ХХ ст. і використання психодіагностики в педагогічному процесі. У першій половині ХХ ст. ідеї, пов'язані з проблематикою педагогічної діагностики в науці активно розвивали О.П. Нечасв, П.П. Блонський, А.С. Макаренко.

Внесок А.С. Макаренка в теорію і практику вивчення та оцінювання ефективності педагогічного процесу був вагомим. З його ідей слід особливо виділити такі положення, що стосуються до педагогічної діагностики:

- 1) про усвідомлену і конкретну мету діагностики – забезпечення перетворення особистості;
- 2) про характер взаємозв'язку педагогічного впливу та психологічного дослідження;
- 3) про методик у вивчення вихованців, про характер використання результатів досліджень;
- 4) про допоміжну роль психології, фізіології та інших наук у педагогічному вивченні особистості [26].

Тестові та інші експериментальні методики оцінювання рівня навченості, та й у цілому дослідження в галузі педагогіки й психотехніки, що мають діагностичну спрямованість, плідно розвивалися в першій чверті ХХ ст. Однак наприкінці 1920-х – початку 1930-х років вони зазнали різкої критики як ті, що не відповідають марксистському напрямку в педагогіці й психології. В цілому до 1970-х років у вітчизняній педагогічній літературі проблеми діагностичного плану розглядалися як елементи теорії навчання.

У період 1970 – 80-х років проблеми, які стосувалися предметного поля педагогічної діагностики, досліджувалися в працях таких видатних психологів, як А.Н. Леонтьєв, А.Р. Лурія, Д.Б. Ельконін, А.Л. Венгер. А в другій половині 1980-х і в 1990-ті роки педагогічна діагностика безпосередньо стає об'єктом спеціальних педагогічних досліджень. Наприкінці 1990-х років термінологія і проблематика педагогічної діагностики дедалі більше входять у науковий обіг.

В останні роки нагальною стає потреба в дослідницькій діяльності педагога, оскільки реформування всієї системи освіти, пов'язане з утворенням незалежної Української держави та прагненням України ввійти до Європейського освітнього простору, передбачає забезпечення її технологічної модернізації, що можливе лише за умови проведення наукових досліджень у педагогічній галузі.

### **1.1.2. Психолого-педагогічні засади діагностики та контролю знань**

Складовими частинами навчально-виховного процесу є діагностика, контроль та оцінювання результатів навчання, саме завдяки яким відбувається педагогічна взаємодія між викладачем і студентом.

М.О. Сорокін, розкриваючи зміст поняття контролю знань, зауважує, що контроль знань означає перевірку. Якщо контроль знань побудований правильно, то він сприяє своєчасному виявленню прогалин у знаннях і вміннях студентів, систематизації та встановленню рівня готовності до засвоєння студентом нового навчального матеріалу, формування вміння самоконтролю. І.П. Підласий [60], Н.В. Морзе [57] і Л.П. Крившенко [41] розглядають педагогічний контроль як складову педагогічної діагностики. На нашу думку, педагогічний контроль і педагогічна діагностика розрізняються за метою і спрямованістю результатів, тому це різні, хоча й дуже близькі поняття. У табл. 1.1 наведено чинники за О.Г. Колганіним, які зумовлюють диференціацію понять педагогічна діагностика та педагогічний контроль. Зауважимо, що на результатах контролю ґрунтується оцінювання навчальних досягнень студентів, що характеризує рівень оволодіння студентами знаннями, уміннями і навичками. Контроль знань, здебільшого спрямований на виявлення рівня засвоєння студентами вже вивченого матеріалу.

Таблиця 1. 1

### Диференціація понять "педагогічна діагностика" та "педагогічний контроль"

<b>Педагогічна діагностика</b>	<b>Педагогічний контроль</b>
Компоненти педагогічної системи	
Контроль є однією з функцій діагностики	
Педагогічний контроль застосовує деякі методи діагностики	
Педагогічній діагностиці підлягають не тільки навчальні досягнення студента, а і його початкова підготовка, мотиви, деякі психофізіологічні властивості, що впливають на ефективність навчання	Контролювати можна винятково те, що свідомо формується (навчальні досягнення, поведінку тощо)
Результати діагностики емоційно нейтральні	Контроль передбачає емоційно-ціннісне оцінювання особистості

Педагогічна діагностика	Педагогічний контроль
Діагностика передбачає опрацювання даних: інтерпретацію, класифікацію, формування рекомендацій щодо коригування навчання, прогнозування	Контроль включає в себе інтегральну оцінку навчальної діяльності студента та його досягнень
Діагностика завжди передбачає детальний аналіз за елементами, що формують очікуваний результат навчання, а також передбачає виявлення причин труднощів	Контроль може здійснюватися як на підставі діагностики узагальненням результатів, так і за допомогою інтегральних методів, що перевіряють сформованість професійної компетентності через виконання комплексних практичних завдань
Діагностика передбачає непрямий вплив на студента через рекомендації щодо вибору методу навчання	Контроль передбачає прямий вплив на студента: заохочення або покарання
Інтерпретація результатів діагностики здійснюється після накопичення необхідного обсягу даних	Контроль передбачає оперативну реакцію на виявлені порушення або успіхи
<b>Мета діагностики</b> – інформаційне забезпечення системи управління навчальним процесом щодо вибору змісту і методів навчання	<b>Мета контролю</b> – перевірка, оцінювання результатів навчання, спостереження за навчальним процесом

Проаналізувавши педагогічні та методичні праці з'ясували, що Є.І. Петровський, М.О. Архангельський, Т.С. Панфілов термін "перевірка знань студентів" ототожнювали чи замінювали терміном "контроль знань", тоді як перевірка – це структурний елемент контролю знань.

Сутність і роль контролю знань відображено в багатьох педагогічних працях. Ідеї контролю, зокрема, висвітлені в працях К.Д. Ушинського, Я.А. Коменського, А.В. Дістервега, М.І. Пирогова та інших. Ідеї контролю, висунуті Я.А. Коменським та сформульовані у вигляді коротких правил, детально встановлюють порядок контролю знань студентів для сучасного вищого навчального закладу [35].

Функції контролю в педагогіці розглядаються багатьма авторами, серед яких Д.О. Лордкіпанідзе, М.М. Покровська, Н.Є. Анкудінова,

М.Т. Калинчук, М.В. Поха, В.О. Уметський та інші. Як свідчать результати аналізу педагогічних джерел, різні автори по-різному трактують функції контролю знань студентів. Нижче наведено деякі основні, на нашу думку, функції контролю:

1. Діагностична функція полягає у тому, що через систему завдань викладач визначає можливості подальшого індивідуального просування кожного студента. Вона полягає у тому, що викладач діагностує ефективність прийомів і методів навчання, які він використовує в навчальному процесі (виявлення знань, умінь і навичок, утруднень, недоліків, забезпечення зворотного зв'язку: "студент – викладач" і "студент – студент").

2. Навчальна функція направлена на вироблення таких завдань для студента, які сприяють узагальненню, поглибленню та систематизації знань, розвитку логічного мислення у студентів (сприяння поглибленню, розширенню, удосконаленню знань студентів, уточненню та систематизації навчального матеріалу з дисципліни).

3. Виховна функція виявляється в методиці проведення викладачем контролю, в наступному коментуванні й оцінюванні робіт. Істотне значення для здійснення цієї функції має спрямування її на розвиток інтелектуальних здібностей (спрямовану на поліпшення особистої дисципліни, розвиток волі, характеру, навичок систематичної самостійної праці).

4. Розвивальна функція тісно пов'язана з іншими функціями контролю знань. Більшість психологів і дидактів вважають, що цю функцію використовують у всіх типах контролю (сприяння розвитку психічних процесів особистості – уваги, пам'яті, мислення, інтересів, пізнавальної активності, мовленнєвої культури студентів).

5. Контролювальна функція передбачає встановлення рівня навчальних досягнень окремих студентів і групи в цілому, вона дає викладачу змогу оцінювати якість засвоєння знань, своєчасно планувати роботу коригування й методики вивчення наступного матеріалу (визначення рівня знань, умінь і навичок студентів, підготовленості до засвоєння нового матеріалу, виставлення оцінок студентам).

6. Орієнтовна функція впливає на розумову роботу, сприяє усвідомленню студентом процесу цієї роботи і розумінню власних знань.

7. Стимулювальна функція впливає на вольову сферу через переживання успіху чи невдачі (стимулювання студентів до поліпшення навчальної діяльності, розвитку особистої відповідальності, формування мотивів навчання).

У межах системи навчального процесу визначають об'єкти контролю знань. Перед вивченням нового матеріалу застосовується попередній контроль, мета якого виявити рівень знань, умінь, які мають

стати фундаментом для засвоєння нової теми. У ході вивчення матеріалу, його закріплення і актуалізації, використовуються поточний контроль якості засвоєного матеріалу, що дає викладачу можливість оперативного впливати на прогалини в знаннях студентів або вибрати інші методи роботи. Тематичний контроль має на меті виявити рівень засвоєння обов'язкових знань з теми. У кінці року здійснюється підсумковий контроль.

Основними завданнями контролю знань є:

- оцінювання рівня засвоєння студентами навчального матеріалу дисципліни;
- інформування студентів щодо якості їхньої роботи з вивчення дисципліни;
- мотивація студентів до активної систематичної роботи протягом семестру;
- аналіз успішності й вплив викладача на процес самостійної роботи студентів та ефективність навчального процесу в цілому.

Якщо до організації контролю підходить традиційно, використовуючи його для отримання показників досягнутого рівня знань студентів, то він буде давати уявлення лише про кінцевий результат виконаної роботи, при цьому дані про сам процес (чи відповідає форма дій даному етапу засвоєння) будуть відсутні. Правильно організований контроль навчальної діяльності дає змогу викладачу оцінювати одержувані студентами знання, уміння, навички, вчасно надати необхідну допомогу і досягати поставлених цілей навчання. Усе це в сукупності створює сприятливі умови для розвитку пізнавальних здібностей студентів і активізації їхньої самостійної роботи.

Контроль залежно від специфіки його організаційних форм проведення поділяють на фронтальний, індивідуальний, груповий і комбінований, а також самоконтроль студентів (Рис. 1.1).

Мета контролю знань визначає вибір методів, при цьому слід враховувати, що такі методи можуть застосовуватися в усіх типах контролю. Необхідно пам'ятати, що лише комплексне їх застосування дає можливість регулярно та об'єктивно виявляти динаміку формування системи знань та умінь студентів. Кожний метод контролю має свої переваги і недоліки, жоден з них не може бути єдиним, здатним діагностувати усі аспекти процесу навчання. Тільки правильне та педагогічно доцільне поєднання усіх методів сприяє підвищенню якості навчально-виховного процесу [71].

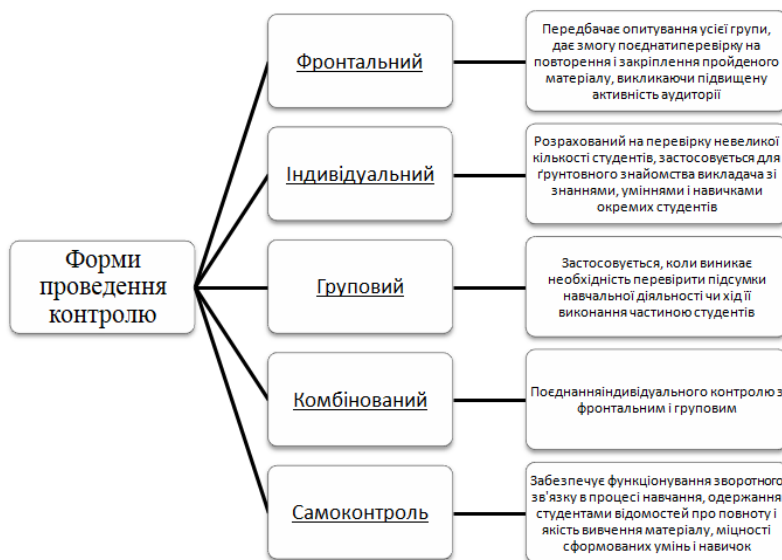


Рис. 1.1. Форми проведення контролю

За місцем, яке посідає контроль у навчальному процесі, розрізняють попередній (вхідний), поточний, рубіжний і підсумковий контроль.

*Попередній контроль* застосовується як передумова для успішного планування і керування навчальним процесом. Він дає змогу визначити наявний рівень знань для використання їх викладачем, щоб орієнтуватися щодо складності завдань. Він проводиться на попередньому етапі вивчення нової дисципліни чи курсу, щоб з'ясувати реальність оцінок, отриманих на вступних іспитах. Попередній контроль у вигляді перевірки і оцінювання залишкових знань також проводять після підсумкового іспиту з метою як оцінювання міцності знань, так і визначення рівня знань для з'ясування сприйняття нових навчальних дисциплін.

*Поточний контроль* знань є засобом виявлення ступеня засвоєння навчального матеріалу. Керування навчальним процесом можливе тільки на підставі даних поточного контролю. Поточний контроль має такі завдання:

- 1) виявити обсяг, глибину і якість сприйняття навчального матеріалу;
- 2) визначити прогалини у знаннях і спланувати способи їх усунення;
- 3) виявити ступінь відповідальності студентів і їх ставлення до роботи, встановивши причини, які перешкоджають роботі;
- 4) визначити рівень опанування навичок самостійної роботи і

спланувати способи і засоби їх розвитку;

5) стимулювати інтерес студентів до вивчення дисципліни і їхню активність у пізнанні.

Головне завдання поточного контролю – допомога студентам організувати роботу, навчитись самостійно, відповідально та систематично вивчати навчальний матеріал.

*Рубіжний контроль* (тематичний, модульний, блоковий) знань є показником якості вивчення окремих розділів, тем, а також пов'язаних з цим пізнавальних, методичних, психологічних і організаційних якостей студентів. Його завдання – сигналізувати про стан процесу навчання студентів для прийняття заходів щодо оптимального його регулювання. Такий контроль можна проводити усно й письмово, у вигляді контрольної роботи, індивідуально або у групі.

*Підсумковий контроль* виконують переважно у формі іспиту студентів з метою оцінювання їхніх знань і навичок відповідно до моделі фахівця. До підсумкового контролю належать семестрові, курсові та державні іспити, а також заліки перед іспитом. Природно, що підсумковий контроль більшою мірою, ніж інші типи контролю, здійснює контролюючу функцію, потребує систематизації й узагальнення знань і певним чином реалізує навчальну, розвивальну та виховну функції контролю [71].

Контроль навчання студентів у ЗВО нині не відповідає вимогам підготовки фахівців сучасного рівня. Головний недолік у тому, що контроль не повністю виконує такі основні функції, як навчальну, діагностичну, стимулювальну, виховну тощо. Наявна практика оцінювання навчальних досягнень перетворює навчання в гонитву за оцінками та має істотний елемент випадковості.

Особливе місце серед методів оцінювання успішності студентів посідає такий метод, як тестування. Відповідно до Постанови Кабінету Міністрів України Про Державну Національну програму "Освіта" ("Україна ХХІ століття"), розділу "Вища освіта", одним з пріоритетних напрямів реформування вищої освіти є розроблення системи діагностики якості освіти та системи тестів для визначення відповідності рівня освіти державним стандартам [30].

Тестовий контроль здійснюється за набором стандартизованих завдань, які дають можливість за порівняно короткий термін часу перевірити засвоєння навчального матеріалу, виміряти обсяг і рівень конкретних знань.

Найважливішими принципами діагностування та контролю успішності учнів є об'єктивність, систематичність і наочність [59].

Об'єктивність полягає в науково обґрунтованому змісті діагностичних тестів, діагностичних процедур, рівному ставленні педагога до всіх, точному й адекватному встановленню критеріям оцінюванню



знань і умінь. Об'єктивність діагностування означає, що виставлені оцінки є однаковими незалежно від методів і засобів контролю, а також педагогів, які здійснюють діагностування.

Систематичність полягає в необхідності проведення діагностичного контролю на всіх етапах дидактичного процесу – від початкового сприйняття знань до їх практичного застосування. Контроль необхідно здійснювати з такою частотою, щоб надійно перевірити все те важливе, що студентам варто знати і вміти. Принцип систематичності потребує комплексного підходу до проведення діагностування, за якого різноманітні форми, методи і засоби контролю, перевірки, оцінювання використовуються у тісному взаємозв'язку і єдності, а також підпорядковуються єдиній меті. Такий підхід відкидає універсальність окремих методів і засобів діагностування.

Наочність полягає насамперед у проведенні відкритих іспитів усіх студентів за одними і тими самими критеріями. Рейтинг кожного учня чи студента, який встановлюється в процесі діагностування, має наочний порівняльний характер. Принцип наочності потребує також оголошення і мотивації оцінок. Оцінка – це орієнтир, за яким учні та студенти роблять висновки щодо еталонів вимог до них і об'єктивності викладача. Необхідною умовою реалізації принципу є оголошення результатів діагностичних зрізів, обговорення й аналіз їх за участю зацікавлених, складання перспективних планів ліквідації прогалин тощо.

Розроблення системи контролю навчальних досягнень студентів, якості знань та й загалом якості навчальних послуг – одна з актуальних проблем сучасної педагогічної науки. Обставини зумовлюють важливість і доцільність вирішення цієї проблеми, вказують на необхідність її детального дослідження та впровадження результатів у навчальний процес вищого навчального закладу.

### **1.1.3. Функції та завдання педагогічної діагностики**

Педагогічна діагностика відрізняється від традиційних процедур контролю та оцінювання. Ці відмінності полягають у такому:

– змістовій, організаційній та методичній цілісності діагностичної діяльності;

– комплексності її компонентів;

– відносній самостійності в рамках педагогічного процесу;

– великому впливові на зміст педагогічного процесу.

Аналіз наукових джерел свідчить [45; 47; 44], що педагоги-науковці однакостайні у визнанні необхідності діагностики особистісних якостей студентів, а також систематичного глибокого аналізу їхніх досягнень під час навчально-виховного процесу з метою його оптимізації. Відповідно, виділимо такі головні об'єкти педагогічної діагностики:

- вихованість і освіченість особистості, сформованість інтегрованих якостей;

- поведінка і діяльність студентів;

- різноманітність впливу зовнішнього середовища;

- особливості загальношкільного і класного колективів, їх педагогічна характеристика;

- зміст і ефективність педагогічної діяльності.

З розвитком педагогіки спектр застосування педагогічної діагностики істотно розширюється. Він включає в себе такі функції:

- Аналітична функція – це психолого-педагогічний аналіз освітнього процесу на всіх рівнях його структурної організації як педагогічної системи.

- Діагностична функція – це психолого-педагогічне вивчення засвоєних знань, вихованості та розвитку студента, а також рівня професійної компетентності викладача.

- Функція контролю – педагогічна діагностика дає змогу виконувати оперативний контроль за навчально-виховним процесом, оскільки містить відомості про його стан.

- Оцінювальна функція – це оцінювання роботи адміністрації навчального закладу, кожного викладача чи студента.

- Коригувальна функція – це дидактичне коригування навчально-виховного процесу і психолого-педагогічне коригування власної активності викладача в бік саморозвитку.

- Орієнтувальна функція – це орієнтація педагогічного колективу на виконання цілей і завдань навчального закладу, а також на усунення недоліків, які мають місце в житті педагогічного колективу та окремих викладачів.

- Інформаційна функція – це постійне інформування всіх учасників педагогічного процесу про результати педагогічної діагностики.

- Функція мотивації і стимулювання – педагогічна діагностика дає змогу проводити диференційовану оплату праці, адекватніше використовувати зовнішні стимули, підвищує рівень особистої мотивації.

Отже, виходячи зі змісту функцій педагогічної діагностики, можна сформулювати конкретні завдання, які стоять перед нею:

- виявлення особистісних якостей викладача, які впливають на навчально-виховний процес;

- виявлення позитивних і негативних рис викладача;

- розроблення критеріїв ефективності роботи викладача;

- фіксація професійно-необхідного рівня знань і умінь викладача.

У цілому на підставі аналізу наявних підходів до визначення сутності та змісту педагогічної діагностики як своєрідного явища педагогічної діяльності можна сформулювати низку положень, на яких, на

нашу думку, має ґрунтуватися теоретичний розгляд і практичне вирішення проблем педагогічної діагностики:

– діагностика – цілеспрямований, самостійний тип пізнавальної діяльності викладача;

– мета діагностики – забезпечити ефективний позитивний вплив на студентів та навчально-виховний процес. Вона має прикладний характер і забезпечує досягнення кінцевого результату;

– основою діагностики є дані про сутність і зміст об'єктів вивчення, що визначає специфіку принципів, методів, прийомів і способів діагностування;

– під педагогічною діагностикою слід розуміти пізнавально-перетворювальну діяльність із розпізнавання навченості, вихованості, індивідуальних і групових особливостей учасників педагогічного процесу та його компонентів, спрямовану на забезпечення результативності даного процесу і досягнення педагогічних цілей;

– основна мета педагогічної діагностики – забезпечити ефективну підготовку фахівця, особистісний та професійний розвиток із урахуванням отриманих у ході діагностики характеристик індивідуальності та груп студентів;

– головні завдання педагогічної діагностики – розпізнавання, аналіз і оцінювання навченості, вихованості, індивідуальних і групових психологічних особливостей, особистісного розвитку учнів та їх підготовки до професійної діяльності;

– педагогічна діагностика проводиться на всіх етапах освітнього процесу і передбачає отримання та використання в педагогічних цілях достовірних відомостей про учасників, умови, зміст і результати навчального процесу.

У цілому результати досліджень і практика свідчать про необхідність визначення в педагогічному процесі індивідуальності студента у взаємозв'язку як психологічних, так і педагогічних характеристик особистості. Водночас, оскільки педагогічна діагностика як специфічна діяльність у рамках педагогічного процесу вивчається порівняно недавно, розробленість теоретичних положень, рівень практичного вирішення проблем педагогічної діагностики не відповідає потребам педагогічної науки і практики. Недостатньо вивченими та вирішеними є питання педагогічної діагностики у ЗВО. Результати досліджень свідчать, що концептуальні технології педагогічної діагностики потребують подальшого обґрунтування, формування та розвитку. Сучасний період розвитку теорії та практики освітніх вимірювань (педагогічної діагностики та контролю) характеризується інтенсивним впровадженням у освітній процес комп'ютерної техніки та інформаційних технологій.

## 1.2. Комп'ютерно-орієнтовані технології освітніх вимірювань як педагогічна проблема

Використання комп'ютерно-орієнтованих технологій у навчальному процесі ЗВО сприяє усвідомленню необхідності наукового узагальнення накопиченого досвіду та розробленню ефективних способів організації освітніх вимірювань із залученням технологій, що мають відповідати цілям і змісту фахової підготовки, зокрема, студентів. Процес змін у системі освіти включає в себе, зокрема, введення нових механізмів і процедур забезпечення якості підготовки студентів. Оскільки успішність навчання багато в чому зумовлюється оперативністю і достовірністю даних про навчальні досягнення, виконання завдання вдосконалення процесів керування якістю вищої освіти передбачає підвищення ефективності діагностики та контролю якості підготовки студентів ЗВО на кожному з рівнів.

У науково-педагогічній літературі відсутнє чітко визначене тлумачення поняття "комп'ютерно-орієнтовані технології освітніх вимірювань". Під комп'ютерними технологіями навчання розуміють сукупність комп'ютерних методів, засобів і прийомів, що базується на сучасних інформаційно-комунікаційних технологіях і використовуються для вирішення навчальних проблем. Комп'ютерно-орієнтовані технології освітніх вимірювань можна розглядати як одну з категорій комп'ютерних технологій навчання. Використання комп'ютерно-орієнтованих технологій освітніх вимірювань зорієнтоване на забезпечення та супровід педагогічного контролю та діагностики, коли передбачається не тільки організація та спосіб проведення тестування, а й подальше прогнозування стану навчальних досягнень, аналіз результатів тестування та якості критеріїв педагогічного тесту для його коригування та вдосконалення.

Освітні вимірювання – це складова частина освітньої діяльності, сутність якої полягає у визначенні рівня відповідності результатів навчально-виховної роботи встановленим освітнім стандартам за допомогою тієї чи іншої системи вимірювання. Застосування комп'ютерно-орієнтованих технологій освітніх вимірювань передбачає автоматизоване рейтингове оцінювання навчальної та виробничої діяльності учасників освітнього процесу. Освітні вимірювання ґрунтуються на використанні засобів педагогічної діагностики (тестів) не лише в навчально-виховному процесі, а й для порівняння та прогнозування якості знань майбутніх фахівців.

Дослідивши специфіку застосування та функціональні характеристики комп'ютерно-орієнтованих технологій освітніх вимірювань, можна стверджувати, що *це технології, використання яких має забезпечити процедуру проведення педагогічної діагностики та контролю якості знань майбутніх фахівців, набутих у процесі навчання.*

Застосовуючи комп'ютерно-орієнтовані технології освітніх вимірювань у процесі підготовки вчителів, можна значно зменшити обсяги використання людських ресурсів і розширити можливості діагностики та контролю засвоєних знань студентами. Вони не тільки забезпечують значну економію часу викладача, а й дають змогу швидко й об'єктивно оцінити реальні знання студента, тобто студент може їх ефективно використати під час самопідготовки до іспитів і заліків. Проте існуючих можливостей систем тестування явно недостатньо для повноцінного контролю знань [21].

Як зазначено в праці [57], можна виділити такі переваги формування компетентностей студентів із застосуванням комп'ютерно-орієнтованих технологій:

- зменшення часу перевірки великого обсягу навчального матеріалу у численній групі опитуваних;
- можливість регулювання заздалегідь визначеного рівня вимог (допускаючи автоматизовану зміну ступеня складності запитань);
- можливість самоконтролю на попередньому етапі з метою самооцінювання результатів підготовки перед офіційним тестуванням;
- отримання об'єктивної оцінки без людського чинника;
- можливість організаційного зворотного зв'язку між студентом і викладачем з використанням мережі Інтернет;
- можливість формування узагальнених статистичних оцінок результатів контролю, а отже, й самого процесу навчання.

Однак, як наголошував А.В. Романов [64], застосування таких технологій має і свої недоліки:

- наявність тільки однієї правильної відповіді;
- можливість вибору правильної відповіді навамання;
- відсутність можливості самостійного формування відповіді;
- у розширених запитаннях існує можливість оцінити лише кінцевий результат;
- гранична заформалізованість відповідей, складність перевірки змісту досліджуваних явищ і фактів, а також їхніх закономірностей;
- потреба у висококваліфікованих фахівцях і експертах, які формують тестові завдання.

Сучасний стан розвитку комп'ютерно-орієнтованих технологій освітніх вимірювань дає змогу уникнути як не всіх цих недоліків, то більшої їх частини.

Зауважимо, що в сфері вирішення проблем контролю знань студентів наразі багато ЗВО розробляють власні комплексні комп'ютеризовані системи, призначені для ефективного контролю й оцінювання знань студентів. Основною складовою цифрового університету поряд з його типовими системами: корпоративним та

студентським Інтернет-порталом, цифровою бібліотекою, публічним веб-сайтом, системою електронного навчання, має бути інформаційно-аналітична система комп'ютерної діагностики навчальних досягнень студентів.

Це зумовлено тим, що метою діяльності ЗВО є забезпечення підготовки фахівців з вищою освітою, наукових кадрів, проведення наукових досліджень та ефективного використання наукового потенціалу університету. На жаль, наявна система контролю діяльності учасників освітнього процесу є значною мірою суб'єктивною, оскільки викладач як суб'єкт керування переважно сам оцінює якість знань студентів, результати їх навчально-пізнавальної діяльності, а отже, і якість власної педагогічної роботи [67]. Інтегруючи комп'ютерно-орієнтовані технології в керування освітнім процесом та безпосередньо залучаючи їх до супроводу педагогічної діагностики та контролю якості фахової підготовки, слід розглянути та визначити передумови такого процесу.

### **1.2.1. Передумови комп'ютерного супроводу контролю якості знань у вищому навчальному закладі**

Освітня парадигма визначає процес навчання як керувану навчальну діяльність, націлену не стільки на збільшення багажу знань, умінь і навичок, скільки на більш важливі зміни особистості: підвищення інтелекту, виховання певних рис особистості, психологічні зміни особистості в напрямі суттєвої самореалізації. Особливо це стосується навчання в педагогічному університеті, де студент поступово в міру самоусвідомлення своєї професійної та соціальної значущості перетворюється з об'єкта на суб'єкт керування іншими людьми та собою.

Поняття "комп'ютерна технологія навчання" вперше зустрічається в наукових публікаціях М.І. Жалдака. В педагогічному словнику [29] термін "комп'ютеризація навчання" визначається як застосування комп'ютера з різною метою. Система комп'ютерного навчання включає технічне, програмне й навчальне забезпечення. Інтенсивна комп'ютеризація освіти сприяє появі та впровадженню в навчальний процес нових комп'ютерних технологій навчання. Комп'ютерні технології навчання – це технологія навчання, що базується на принципах технологій та реалізована за допомогою комп'ютерів, сукупність засобів (програмного і технічного забезпечення, методичних прийомів і теоретичних знань) та способів їх застосування для ефективної діяльності учнів і викладачів під час самостійної роботи, на лекціях, практичних і лабораторних роботах.

Що стосується поняття "комп'ютерні технології навчання", то науковці трактують його по-різному. І.І. Мархель визначає їх як комплекс уніфікованих методологічних, психолого-педагогічних, програмно-технічних та організаційних засобів, призначених для інтенсифікації

самостійної пізнавальної діяльності, навчання чи керування навчанням [49]. І.Ю. Сліпчук під комп'ютерними технологіями навчання розуміє сучасні технології навчання, що створюються з метою підготовки підростаючого покоління до життя в інформаційному суспільстві, формування в нього необхідного світогляду й інформаційної культури на основі підвищення ефективності навчального процесу завдяки впровадженню комп'ютерних засобів навчання [68].

Традиційно під комп'ютерними технологіями навчання розуміють – технології навчання, базовані на використанні комп'ютера та програмного забезпечення, що сприяють виконанню завдань, вирішення яких без використання комп'ютера недостатньо ефективне. Поєднуючи в собі якості різноманітних традиційних інформаційних технологій, комп'ютерні технології дають змогу істотно оптимізувати процес навчання. Використовуючи комп'ютерні технології навчання, можна здійснювати діагностичне цілеспрямоване навчання, виявляти результативність навчальної діяльності, виконувати алгоритмізацію діяльності, посилити її керованість, відтворюваність, коригування, економічність, проводити постійний оперативний зворотний зв'язок.

Недоліки традиційної системи контролю очевидні. Простота та універсальність усного опитування дуже приваблива, однак воно є досить вибіркоким, не дає змоги опитувати всіх студентів групи, а в разі слабких відповідей зацікавленість "сильних" студентів до теми обговорення повністю зникає. Іспити, наприклад, дають враження рівного творчого змагання між студентом і викладачем, але через особливі умови їх проведення знижується можливість реалізації високого рівня знань багатьма студентами. Система оцінок слабо диференціює студентів, зазвичай в одну групу потрапляють люди з дуже різним рівнем знань. Нарешті, всі педагоги знають, як урізноманітнює навчальний процес і підвищує мотивацію студентів до навчання використання нових елементів у викладанні, тому до наявної системи контролю постійно потрібно шукати ефективне доповнення.

Прикладом такого доповнення є тестовий контроль, який останнім часом привертає дедалі більшу увагу педагогів. Перевага тестового контролю полягає в тому, що він є науково обґрунтованим методом емпіричного дослідження, а також певною мірою дає змогу подолати суб'єктивність оцінювання знань студентів.

Результати аналізу зарубіжного і вітчизняного досвіду свідчать, що за допомогою тестування можливе успішне виконання таких педагогічних завдань:

- визначення рівня отриманих знань, умінь і навичок студентів;
- отримання даних відповідності фактичного рівня навченості освітнім вимогам;

– отримання порівняльного оцінювання рівня навченості в різних групах студентів.

Водночас зазначимо, що часто тестові завдання, які використовуються викладачами не відповідають методичним вимогам з конструювання тестів.

Засобами тестового контролю викладач може оперативно перевірити базові знання, внаслідок чого залишиться більше часу для спілкування зі студентами на рівні концепцій і висновків, а також перевірити розуміння проблематики тієї чи іншої навчальної дисципліни.

Комп'ютерний супровід контролю результатів навчання надає широкі можливості для індивідуалізації процесу засвоєння знань студентами. З'являється реальна можливість ефективного виконання низки завдань із підвищення якості навчання не за рахунок збільшення навантаження на викладача, а впровадженням у навчальний процес комп'ютерно-орієнтованих технологій і тестування [40]. Впровадження комп'ютерно-орієнтованих технологій, які об'єднують діагностику, контроль і навчання, спочатку дає змогу виявити ступінь розуміння навчального матеріалу за допомогою вхідних тестів, а потім проаналізувати ті труднощі, які трапляються студентам при виконанні завдань у режимі навчання. На підставі результатів аналізу в рамках системи проводиться своєчасне коригування процесу засвоєння нових знань. Метою комп'ютерного тестування знань студентів є оцінювання відповідності якості підготовки фахівців вимогам Державних освітніх стандартів, створення внутрішньої системи якості освіти на основі незалежного контролю.

У процесі комп'ютеризованого контролю спеціальні інструментальні засоби дають змогу здійснювати систематичну покрокову діагностику поточних результатів, яка важлива для диференціації стимулювальних впливів з боку викладача та своєчасного коригування процесу засвоєння нових знань. Коригування має диференційований характер і зводиться, як правило, до допомоги студентові. Залежно від помилок коригування може мати характер підказки, надання зразка завдання, поглибленого розгляду базового теоретичного матеріалу, необхідного для виконання завдання, посилання на відповідний розділ підручника або рекомендації звернутися за консультацією до викладача. Таким чином, участь викладача в процесі коригування є вкрай незначною. Він діє на етапі виявлення прогалин у підготовці за допомогою комп'ютера, лише коли виникають суттєві труднощі і студент потребує його розгорнутої та ґрунтовної консультації. У ході розроблення програмного забезпечення системи тестового контролю весь навчальний матеріал поділяється на окремі модулі, які на всіх етапах супроводжуються тестуванням. Окрім того, комп'ютерно орієнтовані технології



забезпечують можливість самоконтролю з боку студента в процесі виконання завдань і дають змогу за потреби звернутися за допомогою для виконання окремих найважчих завдань. Можливість звернення за допомогою до викладача в процесі самоконтролю створює у студентів додаткові мотиви активної навчальної діяльності, забезпечує прагнення отримати нові знання у співпраці з викладачем.

Комп'ютерний контроль має ще одну перевагу. Без особливих витрат часу він дає змогу опитувати всіх студентів з усіх розділів навчального курсу. Сума оцінок може сформувати рейтинг знань, який, на розсуд викладача, може бути підставою для звільнення студента від складання частини завдань, а в деяких випадках – і всього курсу. Такі завдання зацікавлюють студентів своєю незвичайністю порівняно з традиційними формами контролю, спонукають до систематичних занять з дисципліни, створюють додаткову мотивацію навчання [25].

І.Є. Булах зазначає, що комп'ютерне тестування успішності дає можливість реалізувати основні дидактичні принципи контролю навчання [24]:

- принцип індивідуального характеру перевірки й оцінювання знань;

- принцип системності перевірки й оцінювання знань;

- принцип тематичності;

- принцип диференційованого оцінювання успішності навчання;

- принцип однаковості вимог викладачів до студентів.

У наукових джерелах виділяють такі типи комп'ютерного тестування студентів:

- вхідне тестування навчальної дисципліни;

- поточне тестування навчальної дисципліни;

- контрольне модульне тестування навчальної дисципліни;

- контрольне підсумкове тестування навчальної дисципліни;

- контрольне тестування з державної атестації.

З досвіду проведення комп'ютерного тестування, можна зробити висновки, що його впровадження сприяє:

- систематичному відстеженню якості та динаміки навчальних досягнень студентів;

- отриманню статистично достовірної картини індивідуального прогресу кожного студента;

- створенню регіонального комп'ютерного банку даних навчальних досягнень студентів із дисциплін за тривалий час навчання;

- інтенсифікації навчального процесу завдяки збільшенню обсягу навчального матеріалу на занятті;

- підвищенню зацікавленості студентів навчально-виховним процесом;

– можливості творчого та практичного застосування знань, умінь і навичок;

– можливості виконувати завдання не лише під контролем викладача, а й здійснювати самоконтроль навчальної діяльності [36].

У процесі дослідження взаємодії між викладачем і студентом побудовано алгоритм, за яким між ними має здійснюватися взаємодія під час комп'ютерного тестування (Рис. 1.2). За такої діяльності можна виділити два типи зворотних зв'язків - внутрішній та зовнішній.

Внутрішній зв'язок надає відомості, які надходить від автоматизованої системи контролю до студента у відповідь на виконанні тестові завдання. Ці відомості використовуються для коригування власних знань студентів, вони спонукають студента до рефлексії, є стимулом до подальших дій, допомагають оцінити і скоригувати результати навчання, тобто свого роду – це самоконтроль. Проаналізувавши літературні джерела [33; 43; 50], з'ясували, що існують консультативний і результативний внутрішні зворотні зв'язки:

– Консультативний зворотний зв'язок може бути різним: допомога, роз'яснення, підказка тощо.

– Результативний зворотний зв'язок також може бути різним: від "правильно-неправильно" до демонстрації правильного результату або способу дії.

Зовнішній зворотний зв'язок надає викладачеві відомості від автоматизованої системи контролю для проведення моніторингових процедур, педагогічної діагностики, коригування знань, аналізу тестів і тестових завдань.

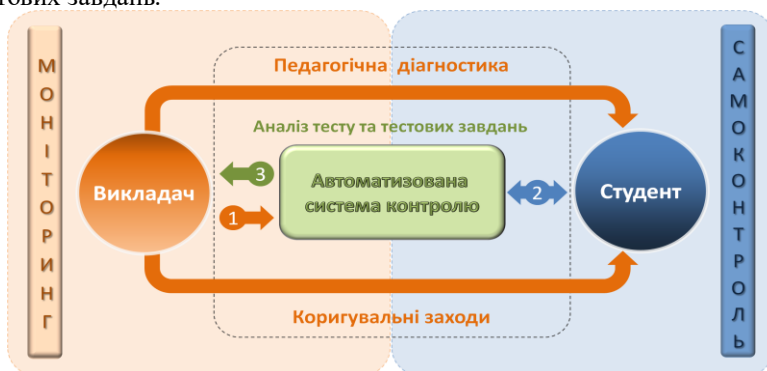


Рис. 1.2. Взаємодія "викладач – автоматизована система контролю – студент": 1 – конструювання тесту, 2 – тестування, 3 – результати контролю.

Проаналізувавши досвід використання та методичні надбання в сфері застосування автоматизованих систем контролю можна виділити

такі його головні функції:

- дослідницька – автоматизоване опрацювання даних про досліджувану систему з метою вироблення рекомендацій щодо підвищення якості знань;
- інформативна – надання відомостей про педагогічну систему;
- інтегративна – формування уявлень про науку як єдиний комплекс знань і методів;
- оцінювальна – зіставлення результатів контролю з еталоном;
- організаційна – інтерпретація якісного і своєчасного виконання навчальних планів;
- контролювальна – здійснення контролю якості навчального процесу;
- прогностична – прогноз успішності діяльності студентів;
- реорганізовувальна – реорганізація праці викладачів і навчального процесу (усунення рутинної праці);
- стимулювальна – заохочення творчої діяльності, спонукання студентів та викладачів до самоосвіти, пошуку нових форм і шляхів навчання;
- регулятивна – надання рекомендацій з регулювання навчального процесу.

Використання комп'ютерно-орієнтованих технологій навчання, зокрема комп'ютерного тестування результатів, розширює можливості контролю та оцінювання рівня навчальних досягнень студентів і є альтернативою традиційним методам діагностики. Такий метод оцінювання швидко, об'єктивно й ефективно діагностує результати підготовки майбутніх фахівців. Таким чином, можна стверджувати, що застосування комп'ютерного тестування як компонента контролю навчальних досягнень є ефективною і перспективною формою.

### **1.2.2. Загальна характеристика комп'ютерно-орієнтованих технологій освітніх вимірювань**

Для того щоб використання системи комп'ютерного тестування було доцільним, необхідне дотримання декількох умов [22]:

- комп'ютерне тестування має надавати результати не гірші, ніж його не автоматизований аналог;
- результат комп'ютерного тестування має якісно наближатися до результату іспиту;
- для отримання додаткової переваги над іспитом тестування має максимально абстрагуватися від людського чинника під час перевірки робіт і визначення оцінки.

У науково-педагогічній літературі [72] відображено різні погляди, щодо комп'ютерного тестування та самої термінології, пов'язаної з ним.

Розглянемо деякі з них:

- комп'ютерно-опосередковане тестування (чітко не визначено як саме застосовується комп'ютер – його можна використовувати лише для аналізу результатів тестування чи аналізу якості педагогічних тестів);

- тестування із застосуванням комп'ютера (комп'ютер є головним засобом для проведення тестування, однак спосіб не визначено);

- тестування, що ґрунтується на використанні комп'ютера (охоплює програмне забезпечення для організації, проведення тестування та аналізу результатів тестування);

- веб-орієнтоване тестування (передбачає проведення тестування через глобальну мережу Інтернет в реальному часі).

Всі ці технології в чомусь подібні, однак у них є свої функціональні відмінності.

У стандарті ISO 9126 (International Organization for Standardization – міжнародна організація мета якої – вироблення стандартів для оцінювання якості програмного забезпечення) визначено шість характеристик, які описують якість програмного забезпечення. Цей стандарт призначено для характеристик, пов'язаних з придбанням, експлуатацією, підтримкою, розробкою, супроводом чи перевіркою програмного забезпечення. У березні 2011 року міжнародною організацією зі стандартизації в межах роботи над проектом SQuaRE (серія стандартів ISO/IEC 25000), призначеним для комплексного оцінювання якості програмного забезпечення, прийнято стандарт ISO/IEC 25010, який повинен замінити ISO 9126-1. Принципових змін у стандарті не відбулося, лише розширено набір базових характеристик до восьми [39]. Як зазначає В.В. Котяк, "...відповідних державних стандартів в Україні не прийнято", тому основні характеристики наведемо на підставі стандарту ISO 9126 [5]:

1. Функціональні характеристики (набір атрибутів, що стосуються функцій програмного забезпечення та їхніх особливостей):

- придатність – наявність та відповідність функціональних характеристик конкретним завданням;

- точність – забезпечення відповідності результатів істинному значенню;

- функціональність – відповідність стандартам, угодам, положенням законів і підзаконних актів.

2. Надійність (набір атрибутів, що стосуються характеристик програмного забезпечення, придатного для підтримування найефективнішої роботи в заданих умовах протягом певного періоду часу):

- стабільність – частота відмов чи помилок у програмному забезпеченні;

- стійкість до помилок – підтримка необхідного рівня якості

функціонування, якщо допущено програмні помилки чи порушено порядок дій;

- відновлення – властивість програмного продукту до відновлення певного рівня продуктивності та даних у разі виходу з ладу;

- дотримання надійності – властивість до відновлення роботи та даних після відмови, а також часові та людські затрати на відновлення роботи.

3. Практичність (набір атрибутів, що стосуються зусиль, необхідних для використання програмного забезпечення, та індивідуального оцінювання такого використання):

- зрозумілість – зусилля користувачів, необхідні для розуміння загальної логічної концепції та застосування;

- здатність до засвоєння – зусилля користувачів, необхідні для навчання роботі з програмним продуктом;

- придатність до використання – зусилля користувачів, необхідні для керування і контролю над функціонуванням програмного засобу;

- привабливість – властивість програмного продукту бути привабливим для користувача;

- простота використання – зусилля користувачів, необхідні для експлуатації й оперативного керування програмним засобом.

4. Ефективність (набір атрибутів, що стосуються взаємозв'язку між рівнем продуктивності програмного забезпечення і кількістю використовуваних ресурсів у заданих умовах):

- швидкодія – властивість програмного продукту до забезпечення належного часу відгуку;

- ресурсемність – властивість програмного продукту, пов'язана з витратами відповідної кількості ресурсів, коли програмне забезпечення використовується за призначенням у певних умовах;

- відповідність вимогам до ККД (коефіцієнт корисної дії).

5. Розширення (набір атрибутів, що стосуються зусиль, необхідних, щоб зробити певні зміни в структурі програмного засобу):

- легкість аналізу – зусилля користувачів, необхідні для діагностики недоліків чи відмов;

- легкість змін – зусилля користувачів, необхідні для модифікації чи зміни функціонала;

- стабільність – ризики непередбачуваних ефектів від модифікації;

- легкість тестування – зусилля користувачів, необхідні для перевірки модифікацій.

6. Мобільність (набір атрибутів, що стосуються придатності для переходу програмного забезпечення від одного середовища до іншого (операційні системи, пристрої)):

- придатність до адаптації – можливість адаптації до конкретних

умов;

- встановлення – властивість програмного продукту до інсталяції в певне середовище;

- простота введення в експлуатацію;

- відповідність стандартам мобільності програмного забезпечення.

Прийнятий у 2011 році стандарт ISO/IEC 25010 має деякі відмінності від попереднього, фактично зміст шести характеристик не змінився, однак до них додано ще дві, а саме:

7. Сумісність – ця характеристика раніше входила до функціональності та мобільності (набір атрибутів, що стосуються придатності двох або більше компонентів програмного забезпечення для забезпечення обміну даними):

- можливість заміни, за якої програмний продукт можна використовувати замість іншого для тих самих цілей і в тому самому середовищі;

- сумісність, за якої є можливість створення єдиного простору для обміну даними без будь-яких втрат;

- сумісність, за якої програмний продукт можна використовувати разом з одним (або більше) іншим програмним продуктом;

– відповідність стандартам сумісності, коли програмне забезпечення відповідає стандартам, конвенціям і правилам, які стосуються сумісності.

8. Безпека (набір атрибутів, що стосуються властивості захисту системних елементів від випадкового або зловмисного проникнення, використання, модифікації та знищення):

- конфіденційність, придатність до захисту від несанкціонованого проникнення до даних;

- цілісність – властивість зберігати точність і повноту даних;

- неспростовність – властивість доведення дії окремого користувача;

- звітність – придатність до відслідковування дій користувача;

- справжність – властивість ідентифікації користувача;

- відповідність стандартам безпеки – програмний продукт відповідає стандартам, конвенціям і правилам безпеки.

Комп'ютерно-орієнтовані технології освітніх вимірювань класифікують за різними характеристиками:

*За рівнем їх автономії:*

- комплекс модулів;

- окремий модуль;

- у складі системи керування навчанням.

*За ступенем універсальності:*

- універсальні – орієнтується на створення тестових завдань у будь-яких галузях;

– спеціалізовані – орієнтується на певну галузь, наприклад, фізику, інформатику, математику.

*За характером проведення тестування:*

- на основі інструментальних, програмних оболонок;
- ґрунтуються на використанні браузера.

Безкоштовне розповсюдження поділяється на таке, що має відкритий код, та таке, що має закритий програмний код.

Існує й інша класифікація за А.П. Кудіним [42], згідно з якою комп'ютерно-орієнтовані технології освітніх вимірювань можна поділити на такі типи:

- за можливістю поповнення бази завдань і внесення змін до програмної оболонки: відкриті і закриті;

- за схемою взаємодії з користувачем: гнучкі та формалізовані;

- за середовищем застосування: локальні (паритетні, "клієнт – сервер") і глобальні;

- за предметно-галузевим спрямуванням: профільні та універсальні.

На думку А.П. Кудіна, яка є прийнятною для нас, складовими типової системи комп'ютерного тестування є:

- підсистема тестування;

- підсистема створення, підготовки і редагування тестів;

- підсистема статистичних даних;

- підсистема керування користувачами;

- база даних тестових завдань і пов'язаних з ними статистичних даних;

- база даних користувачів.

У ході впровадження та використання комп'ютерно-орієнтованих технологій освітніх вимірювань потрібно передбачити:

- здатність до адаптації системи оцінювання результатів тестування;

- можливість використання як для внутрішнього, так і відкритого тестування;

- аналіз тестів з метою оцінювання якості, придатності окремих тестових завдань і виявлення некоректно сформованих чи просто неправильних тестових завдань, що здійснюється шляхом відправлення спеціальних повідомлень адміністратору системи від користувача;

- легкість створення і модифікації тестів;

- відсутність твердої прив'язки до дисципліни, галузі знань;

- збір, збереження і подання в зручній формі різноманітних статистичних відомостей щодо процесу тестування;

- можливість одночасного тестування необмеженої кількості користувачів;

- відсутність твердої прив'язки до конкретного апаратного та програмного забезпечення;

- висока безпека, захищеність і стабільність;
- необмежена кількість тестів, тем, тестових завдань і дистракторів;
- зручний і ефективний імпорт та експорт тестів;
- контроль часу тестування і збереження відомостей щодо перерваного сеансу тестування;
- підтримка безпечного й універсального механізму керування користувачами з розподілом прав доступу;
- підтримка режиму модульного контролю, у процесі якого потрібно дати відповідь на низку тестових завдань протягом певного часу [22].

На рис 1. наведено класифікацію типів систем комп'ютерного контролю знань, яку зроблено на основі опрацьованих джерел.



Рис. 1.3. Класифікація типів комп'ютерного тестового контролю

Впровадження комп'ютерно-орієнтованих технологій в навчальний процес дає можливість здійснювати у ЗВО максимально ефективні навчально-виховні впливи та приймати рішення щодо підвищення якості навчального процесу. Однак досі не вирішено проблему теоретико-методологічного обґрунтування використання комп'ютерно-орієнтованих технологій для забезпечення діагностики та контролю рівня знань студентів, а саме: підтримки створення засобів, орієнтованих на автоматизовану підтримку процесу діагностики фахової підготовки, залишається відкритим.



## РОЗДІЛ 2. МЕТОДИЧНІ ЗАСАДИ ЗАСТОСУВАННЯ КОМП'ЮТЕРНО-ОРІЄНТОВАНИХ ТЕХНОЛОГІЙ ОСВІТНІХ ВИМІРЮВАНЬ

### 2.1. Сутність і специфіка методів і методик комп'ютерної педагогічної діагностики знань

Мета комп'ютерної педагогічної діагностики та контролю якості знань полягає в реалізації моделі гнучкої траєкторії індивідуального професійного самовдосконалення студентів та в інформаційному забезпеченні системи керування навчальним процесом щодо вибору змісту і методів навчання.

Відповідно до поставленої мети можна виділити такі завдання педагогічної діагностики:

- отримання достовірних відомостей про процес і поточний стан фахової підготовки студентів;
- вироблення коригувальних заходів, що сприяють оптимальному засвоєнню системи знань на підставі прогнозу можливих змін;
- подальше відстеження реальних змін у розвитку професійної компетентності на базі коригувальної програми.

Проаналізувавши організацію діагностики та контролю якості знань, можна виділити системо-утворювальні компоненти комплексу комп'ютерно-орієнтованих технологій освітніх вимірювань як засобу удосконалення фахової підготовки студентів, а саме:

- добір фактичного матеріалу;
- оцінювання стану фахової підготовки студентів, можливостей та ресурсів її розвитку;
- прогнозування: пошуковий прогноз (прогноз варіантів станів фахової підготовки студентів); нормативний прогноз (способи педагогічного впливу на всі компоненти фахової підготовки студентів).

Загальна логіка проведення діагностики та контролю якості знань, виявлена нами в ході аналізу праць відомих учених, дала змогу визначити основні методи (засоби) та етапність проведення цих процедур.

Вибір методів залежить від особливостей досліджуваного об'єкта. Тому для комп'ютерної педагогічної діагностики якості знань, набутих у процесі фахової підготовки, слід визначити методи і методики, які будуть забезпечувати вирішення завдань педагогічної діагностики.

Під *методом* слід розуміти спосіб організації або сукупність прийомів освітньої діяльності, спосіб дослідження або пізнання, спосіб організації практичного та теоретичного освоєння дійсності, зумовлений закономірностями розглядуваного об'єкта, спосіб досягнення мети, упорядкована певним чином діяльність. Метод комп'ютерної педагогічної діагностики якості знань, набутих у процесі фахової підготовки студентів

слід визначати як сукупність підходів різних методик, які володіють спорідненістю технологічних прийомів або спорідненістю теоретичних систем уявлень, призначених для визначення індивідуальних і групових особливостей, станів і властивостей учасників освітнього процесу та його компонентів засобами комп'ютерно-орієнтованих технологій.

Під *методикою* слід розуміти технічні прийоми реалізації методу на основі конкретного матеріалу і конкретної процедури, вироблення відповідно до завдань і змісту навчання методів, методичних засобів та організаційних форм навчання. Методику комп'ютерної педагогічної діагностики якості знань, набутих в процесі фахової підготовки студентів, слід розуміти як конкретну процедуру, сукупність технічних засобів, методів і прийомів визначення індивідуальних і групових особливостей, станів і властивостей учасників освітнього процесу та його компонентів.

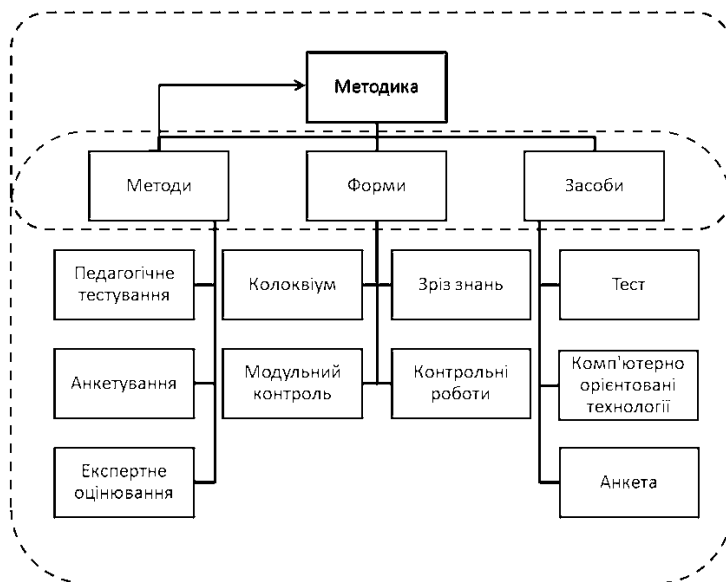


Рис. 2.1. Методи, форми та засоби комп'ютерної педагогічної діагностики

Методика на відміну від методу містить конкретні інструкції з проведення педагогічної діагностики, опрацювання даних та інтерпретації результатів. Для вирішення певних дидактичних проблем відповідно до цілей, особливостей груп, можливості залучення тих чи інших методів і засобів педагогічної діагностики розробляється та впроваджується в навчальний процес конкретна методика. В рамках одного методу може бути використано деякі аспекти різних методик (Рис. 2.1).

Для вибору методів і методик комп'ютерної педагогічної діагностики необхідно визначити загальні та конкретні можливості їх

застосування в процесі фахової підготовки, їхнє місце в системі педагогічної діяльності.

Виділимо групи методів педагогічної діагностики, зокрема, використання яких ґрунтувалось на залученні комп'ютерно-орієнтованих технологій:

1. Методи збору діагностичних відомостей. Вхідними даними для педагогічної діагностики є результат спостереження за навчально-виховним процесом, дані педагогічного контролю та вимоги щодо потрібного рівня якості знань. Методи які слід віднести до цієї групи – це спостереження та опитування (основними традиційними формами опитування є бесіда або інтерв'ю. Однак виходячи з головної мети, а це визначення таких методів, які б ґрунтувались на використанні комп'ютерно-орієнтованих технологій, виділимо анкетування та педагогічне тестування), а також аналіз результатів педагогічної діяльності.

2. До методів опрацювання та оцінювання діагностичних відомостей належать методи шкалювання, математико-статистичного опрацювання, рейтингового оцінювання та кореляційного аналізу.

3. Методи інтерпретації та накопичення результатів діагностики – інтерпретація узагальнених даних педагогічної діагностики, які подано у вигляді рекомендації щодо коригування процесу підготовки фахівця, уявлення про реальний стан, особливості групи або окремої особи. Це можуть бути результати опрацювання тестів та тестових завдань, бази даних, статистичні таблиці, діаграми, моделі, графіки.

4. Методи використання результатів діагностики – психолого-педагогічний вплив, координація та планування психолого-педагогічних дій, прогнозування, рекомендації, побажання, вимоги, управлінські рішення.

5. Методи оцінювання достовірності результатів діагностики – педагогічний контроль, аналіз результатів навчальної діяльності, експертні методи, спостереження, статистичний аналіз.

Одним із ефективних методів реалізації комп'ютерної педагогічної діагностики для раннього виявлення негативних тенденцій в процесі фахової підготовки з метою їх попередження є тестування. Як головний засіб комп'ютерної педагогічної діагностики було обрано тест.

Умовно засоби можна розділити на дві групи: традиційні та комп'ютерно-орієнтовані. Ці дві групи потрібно гармонійно поєднувати та взаємодоповнювати в процесі навчально-пізнавальної діяльності.

*Засоби навчання* – це матеріальні та ідеальні об'єкти, що використовуються в освітньому процесі як носії даних та інструменти діяльності викладача й студентів і застосовуються ними як окремо, так і сумісно. До них належать природне і соціальне оточення, обладнання,

підручники, книги, наукові видання, довідники, енциклопедії, комп'ютери з відповідним програмним забезпеченням, комп'ютерні мережі з відповідним інформаційним забезпеченням та інформаційними ресурсами, електронні посібники та підручники, електронні словники, електронні бібліотеки і т. д.

У педагогічній діагностиці тест – це стандартизована, обмежена в часі система формалізованих завдань, призначених для встановлення освітнього (фахового) рівня студента. Тести слід розглядати як основний інструмент комп'ютерної педагогічної діагностики, оскільки його вимірвальні можливості значно вищі, ніж інших методів.

До класифікації педагогічних тестів немає єдиного підходу, її здійснюють за різними критеріями. Найпоширеніша – узагальнена класифікація, за якою класи тестів згруповано за певними ознаками [44]:

#### 1. За процедурою створення:

– стандартизовані – вимірюється стандартний набір широко визначених результатів навчання, використовується стандартні інструкції та стандартна методика оцінювання, також можна порівнювати результат окремого студента з результатами інших студентів, які виконували той самий тест за аналогічних умов;

– не стандартизовані – складаються на базі конкретної теми для перевірки рівня сформованості певних навичок або вмінь, потребують визначення всіх кількісних показників якості.

#### 2. За метою та змістом:

– тести здібностей (оцінювання здатності студента до оволодіння різного типу діяльністю);

– тести інтелекту (аналіз рівня розвитку пізнавальних процесів і функцій мислення у студентів);

– тести особистості (оцінювання емоційно-вольових якостей студента);

– тести досягнень (оцінювання розвитку знань, вмінь та навичок після закінчення навчання).

#### 3. За методологією інтерпретації результатів:

– нормативно-орієнтовані (порівняння індивідуальних навчальних результатів з результатами інших студентів групи);

– критеріально-орієнтовані (грунтується на порівнянні індивідуальних результатів тестування з загальним обсягом знань, який має засвоїти студент на певному етапі навчання).

#### 4. За цілями використання:

– попередні (на початку вивчення дисципліни або окремого змістового модуля);

– поточні (у процесі вивчення дисципліни);

– підсумкові (після завершення вивчення дисципліни).

## 5. За формою:

- закритого типу;
- відкритого типу;
- на встановлення відповідності;
- на послідовність.

## 6. За конструкцією:

- однорідні (всі тестові завдання однакової складності);
- наростаючі (тестові завдання розташовані в міру зростання складності);
- адаптивні (тест є системою завдань, в якій вибір чергового завдання залежить від відповіді на попереднє).

## 7. За призначенням:

- навчальні.
- прогнозувальні;
- контролювальні;
- класифікаційні;
- інтегративні (моделюють професійну діяльність);
- гомогенні (для вимірювання знань за однією дисципліною);
- гетерогенні (для вимірювання знань за декількома дисциплінами);

## 8. За способом надання:

- бланкові;
- апаратурні;
- практичні;
- комп'ютерні.

## 9. За характером дій:

- вербальні;
- невербальні.

## 10. За формою тестування:

- індивідуальні;
- групові;
- масові.

## 11. Відповідно до мети тестування:

- тести на визначення сформованості загальних умінь;
- тести успішності;
- діагностичні тести;
- тести визначення здібностей.

Наприклад, у праці Р.Б. Каттелла [2] наведено таку класифікацію педагогічних тестів: анкета; тест на сприйняття; тест здібностей; тест умінь і навичок; естетичний тест (уподобання); проєктивний тест; ситуаційний тести; соціально-психологічний тренінг; організаційно-діяльна гра; фізіологічний тест; фізичний тест; випадкове спостереження.

Тест є інструментальним засобом освітніх вимірювань. Власне

вимірювання – це метод, за допомогою якого здійснюється надання кількісного значення досліджуваному критерію. Методом оцінювання тесту є шкалювання, за допомогою якого визначається, яким чином вимірювана величина набуває кількісного значення.

Конструювання тестів для педагогічної діагностики якості знань набутих у процесі фахової підготовки потребує високого рівня психолого-педагогічної компетентності викладача, а у випадках застосування комп'ютерно-орієнтованих технологій – ще й спеціальної підготовки. Ця робота пов'язана з добором завдань, забезпеченням надійності та обґрунтованості даних. Слід враховувати вікові характеристики, склад груп, етапи дослідження, умови проведення та мотиваційний настрій.

При конструюванні тестів для розроблення якісних тестових завдань, які відповідають вимогам надійності, валідності, заданої складності, необхідно дотримуватися певної послідовності. Уточнимо етапи створення педагогічного тесту:

*Eman 1.* Визначення мети тестування. Метою проведення тестування може бути: навчання студентів (самостійне вивчення матеріалу); поточний контроль (діагностика засвоєння окремих тем і розділів); рубіжний контроль; підсумковий контроль знань студентів (з усієї навчальної дисципліни); контроль залишкових знань (з однієї дисципліни, циклу дисциплін).

*Eman 2.* Добір змісту навчального матеріалу. Для забезпечення короткого запису розроблених завдань весь матеріал навчального курсу поділяється на достатньо великі розділи (модулі). Кожен з розділів (модулів) у свою чергу поділяється на теми. За потреби тема може поділятися на блоки і ще дрібніші дидактичні одиниці змісту навчального матеріалу.

*Eman 3.* Проектування матриці тесту. Технологічна матриця задає зміст навчального матеріалу, який буде дібрано для перевірки, і важливість того чи іншого елемента змісту. Вона може містити рівні досягнень, які будуть перевірені, їх співвідношення, відповідність стандарту та деякі інші компоненти. При складанні матриці тесту для будь-якої дисципліни розробник зобов'язаний переконатися, що весь навчальний матеріал охоплено пропонованими завданнями. Зміст дисципліни повинен повністю покриватися елементами матриці за усіма темами. Якщо має місце тестування за окремими підтемами, то і в цьому випадку необхідно, щоб вся підтема була охоплено завданнями тесту. У випадку, коли питання або частина завдань не відповідає темі чи не повністю зрозумілі в рамках даної теми, від них слід утриматися. Таким чином, фіксується вимога широти тесту, повного охоплення всіх розділів дисципліни, що відображається в елементах матриці.

*Eman 4.* Формування структури банку тестових завдань і

конструювання тесту відповідно до рівнів пізнавальної діяльності. Структура банку тестових завдань формується відповідно до структури дисципліни чи навчального курсу. Чіткому визначанню цілей навчання сприяє їх класифікація. Таксономія цілей навчання є описом і навчанням у термінах професійної діяльності.

Класифікація пізнавальної сфери містить шість класів цілей, а саме:

– *Знання* – це здатність запам'ятовувати факти, принципи, процеси в різних предметних галузях.

– *Розуміння* – здатність студентів розуміти матеріал, який вивчають.

– *Застосування* – здатність використовувати матеріал у нових ситуаціях.

– *Аналіз* – здатність структурувати навчальний матеріал так, щоб була зрозумілою загальна організаційна структура.

– *Синтез* – здатність поєднати окремі частини для отримання цілого, що набуває нової якості.

– *Оцінювання* – здатність судити про цінність даного навчального матеріалу в рамках поставленої мети.

*Етап 5.* Проведення тестування. Щоб звести до мінімуму вплив суб'єктивних чинників на результати тестування, необхідна максимальна стандартизація умов його проведення та аналіз результатів.

Д. Уїлмс [31] виділяє такі головні підходи до розроблення принципів педагогічного тестування:

1. Чітке визначення результатів навчання, які відповідають цілям навчання (потрібно визначити цілі навчання, конкретизувати їх, для кожної цілі визначити точні результати навчання).

2. Тести мають містити такі типи завдань, які найбільше підходять для визначення якості знань, що вимагаються від студента на певному етапі навчання (потрібно максимально охопити навчальний матеріал; дотримуватися правил конструювання тестових завдань, підібрати найефективніші і виключити малоінформативні тестові завдання, а також включати в тест різні форми тестових завдань).

3. Тести мають відповідати вимогам щодо результатів. Тестування використовується на різних етапах навчання: для визначення рівня знань на початку навчання (попереднє тестування); для оцінювання знань під час навчання (поточне тестування); для виявлення труднощів у навчанні та їхні причин (діагностичне тестування); для оцінювання рівня сформованості знань (підсумкове тестування).

4. Тести мають відповідати поставленим завданням навчання.

5. За допомогою тестів перевіряється ступінь виконання завдань навчання, тобто за результатами тестування слід встановити рівень навчальних досягнень студентів.

6. Тести мають допомагати в навчанні. Завдяки зворотному зв'язку можна отримати уявлення про результати навчання, виявити труднощі та прогалини в знаннях, які потребують коригування.

7. Мають бути рекомендації щодо стандарту й критеріїв тестування. Тести використовуються для розподілу студентів за рівнем знань і для оцінювання рівня підготовленості за критеріями засвоєння знань. Отже, ці підходи розрізняються за певними характеристиками (за метою використання, за задачами і об'ємом матеріалу, за способами добору завдань і визначенням рівня знань).

8. Результати тестування мають бути обґрунтованими і надійними.

*Етап 6.* Оцінювання та аналіз результатів тестування.

Цей етап включає в себе такі процедури:

1. Формування зведеної таблиці показників якості тестових матеріалів.

2. Побудова комплексних показників якості тестових матеріалів.

3. Нормування показників якості тестових матеріалів.

4. Визначення довірчих інтервалів комплексних оцінок.

5. Формування числової шкали якості тестових матеріалів.

6. Формування підсумкового висновку у вигляді рекомендацій:

- щодо застосування стандартизованих тестових матеріалів для контролю якості навчальних досягнень;

- для авторів з коригування та опрацювання тестових матеріалів із зазначенням причини невідповідності вимогам якості.

Проведення комплексної експертизи якості тестів дасть змогу поліпшити якість тестових матеріалів на етапі їх розроблення і сформувані банк стандартизованих тестових завдань і тестів, що забезпечить контроль-оцінювальні процедури надійними і валідними тестовими вимірювачами.

Аналізуючи результати тестування групи студентів, викладач може здобути різні відомості, зокрема:

- визначити рейтинг студентів для порівняння їхньої успішності та діагностики реального стану;

- виявити структуру знань кожного учасника тестування на діяльнісному рівні;

- виявити, порівнюючи результати тестування та поточну успішність з дисципліни, особливості конкретного студента, пов'язані з недостатньою сформованістю певних навичок роботи з тестовим матеріалом, психологічною неготовністю до незалежного оцінювання навчальних досягнень, відсутністю правильної мотивації;

- виявити недостатність засвоєння теми чи розділу;

- визначити коректність тесту та повноту виконання ним своїх завдань;



– диференціювати студентів за рівнями підготовленості для оптимізації індивідуального підходу у побудові процесу навчання.

Впровадження тестових технологій в навчальний процес істотно змінює роль викладача: від функції передачі знань і подальшого контролю здійснюється плавний перехід до співпраці зі студентами.

Поряд з педагогічним тестом одним з важливих методів педагогічної діагностики є анкетування – один із способів діагностики знанневих якостей особистості студента. Анкета – це структурно організований набір запитань, кожне з яких логічно пов'язане з центральним завданням діагностики. Запитання анкети можуть стосуватися фахової спрямованості, моральних і психологічних якостей особистості та ін. Анкетування ґрунтується на самооцінюванні опитуваними своїх властивостей і якостей. Форма відповіді – оцінка в балах тієї чи іншої особистісної властивості. Оцінювання соціально-психологічних та особистісних якостей може проводитися групою експертів [56].

На нашу думку, метод експертного оцінювання – один з ефективних методів педагогічної діагностики, оскільки дає змогу провести кількісне або якісне оцінювання тих чи інших якостей майбутніх фахівців, без суб'єктивного самооцінювання. Він ґрунтується на судженнях професорсько-викладацького складу і самих студентів про рівень розвитку певних якостей студента.

Мету комп'ютерної педагогічної діагностики можна визначити як розроблення методів всебічного дослідження ознак, проблем педагогічного процесу, також забезпечення валідності, надійності та достовірності їх результатів з використання комп'ютерно-орієнтованих технологій освітніх вимірювань. Таким чином, комп'ютерна педагогічна діагностика через використання певних методів і методик стає інструментом для здійснення коригувальних дій, спрямованих на підвищення ефективності навчального процесу та якості навчання. Комп'ютерна педагогічна діагностика сприяє оперативному здобуттю даних із високою точністю їх вимірювання. Однак зазначимо, що вона не має повністю замінити традиційні методи, а навпаки, для отримання достатньо повних та максимально вірогідних відомостей про суб'єкт діагностування вона повинна доповнювати традиційні методи.

Завдання комп'ютерної педагогічної діагностики полягає у визначенні розроблених методів розпізнавання стану групи чи окремої особистості шляхом фіксації її визначальних характеристик автоматизованими системами освітніх вимірювань, їх зіставлення з тенденціями педагогіки для прогнозування поведінки досліджуваного об'єкта, інтерпретації результатів з метою самоконтролю та впливу на поведінку студентів та прийняття коригувальних рішень.

Автоматизація процесу педагогічної діагностики потребує не тільки знання психолого-педагогічних засад навчання студентів педагогічних вищих навчальних закладів, а й додаткової обізнаності в сфері комп'ютерно-орієнтованих технологій та математичної статистики.

## 2.2. Комплекс комп'ютерно-орієнтованих технологій освітніх вимірювань

Ефективне застосування сучасних комп'ютерно-орієнтованих технологій у навчальному процесі можливе лише у тому випадку, коли відповідні технології не є певною надбудовою до існуючої системи навчання, а обґрунтовано й гармонійно інтегруються в даний процес, забезпечуючи нові можливості як викладачам, так і студентам.

Комплекс комп'ютерно-орієнтованих технологій освітніх вимірювань характеризується його компонентами (Рис. 2.). На рис 2.3. наведемо схему виконання дій з використанням комплексу комп'ютерно-орієнтованих технологій освітніх вимірювань і відображено основні етапи дій. *Мета створення* комплексу комп'ютерно-орієнтованих технологій освітніх вимірювань – удосконалення фахової підготовки студентів.

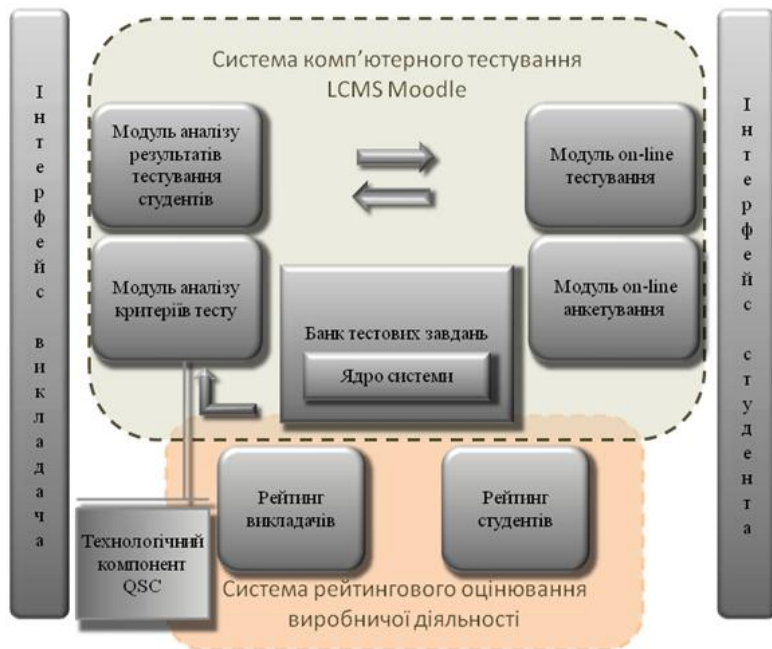


Рис. 2.2. Структура комплексу комп'ютерно-орієнтованих технологій освітніх вимірювань

*Завдання створення* комплексу комп'ютерно-орієнтованих технологій освітніх вимірювань: забезпечити інтенсифікацію процесу навчання та всебічну педагогічну діагностику якості знань; підвищити навчально-пізнавальну активність студентів; забезпечити якість навчання студентів на рівні вимог інформаційного суспільства; створити умови для інтелектуального розвитку студентів і розкриття їх творчого потенціалу; підвищити рівень підготовки майбутніх фахівців.

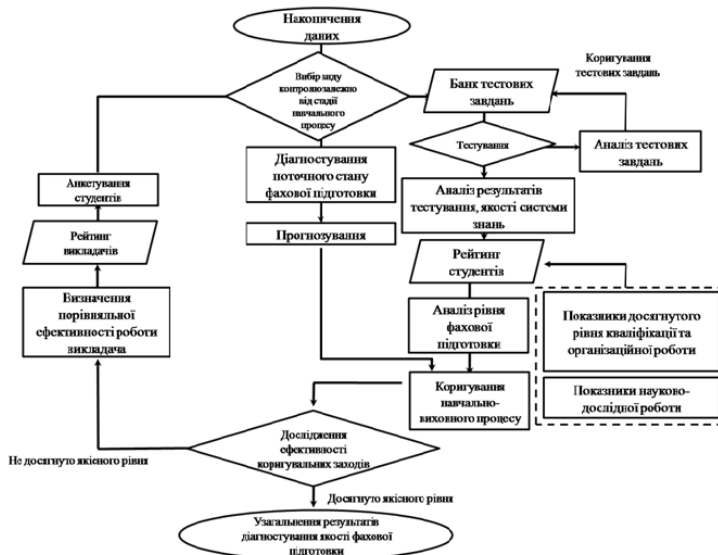


Рис. 2. 3. Схема виконання дій з використанням комплексу комп'ютерно-орієнтованих технологій освітніх вимірювань

Найрозповсюдженішими засобами організації і проведення контролю є системи комп'ютерного тестування на яких базується запропонований комплекс. З точки зору базових методичних підходів використання системи комп'ютерного тестування дає змогу:

- збільшити обсяг самостійної роботи студентів над навчальним матеріалом дисципліни, що актуально в умовах збільшення частки самостійної роботи в загальному обсязі навчального навантаження студентів, а також впровадження дистанційної форми навчання;
- зменшити навантаження на викладачів у частині, що пов'язана з підготовкою контрольних тестів і завдань, при проведенні контрольних заходів і обробці їх результатів, зокрема при перевірці робіт студентів, формуванні звітної документації, аналізі типових помилок, формуванні рекомендацій щодо їх усунення з урахування індивідуальних результатів кожного студента;
- надати викладачам можливість впроваджувати сучасні гнучкі

технології навчання залежно від контингенту студентів і кількості навчального часу, що виділяється на вивчення дисциплін;

- Підвищити мотивацію і зацікавленість студентів у навчанні дисципліни протягом семестру, а не тільки перед підсумковим контролем;
- Зробити процес оцінювання навчальних досягнень студентів об'єктивнішим.

При організації педагогічної діагностики та контролю якості знань із використанням запропонованого комплексу у вищому навчальному закладі доцільно застосовувати такі засоби, форми та методи, які наведено в табл. 2.1.

Таблиця 2. 1

### Комп'ютерно-орієнтовані методи, засоби і форми організації контролю

Компоненти	Зміст компонентів
Засоби	Апаратне забезпечення (комп'ютер, засоби телекомунікацій)
	Програмне забезпечення (операційні системи, табличні процесори, системи керування базами даних, експертні системи, педагогічні програмні засоби, проблемно-орієнтовані програми)
Методи	Наочні методи навчання (робота з навчальними та навчально-контролюючими програмами)
	Практичні методи навчання (дослідницька робота в комп'ютерних лабораторіях, обчислювальні експерименти, телекомунікаційні проекти)
Форми	Комп'ютерно-орієнтовані практичні і лабораторні заняття, контрольні роботи, комп'ютерно-орієнтована науково-дослідна робота, комп'ютерне тестування, комп'ютерно-орієнтовані екзамени й заліки

До комплексу комп'ютерно-орієнтованих технологій освітніх вимірювань, крім системи комп'ютерного тестування (ядро системи – банк тестових завдань, модуль тестування, модуль он-лайн анкетування, модуль аналізу тестових завдань, модуль аналізу результатів тестування та авторський технологічний компонент QSC), входять також рейтингові системи викладачів та студентів.

### 2.2.1 Основи використання комп'ютерно-орієнтованих технологій освітніх вимірювань

Аналізуючи пропозиції у сфері розроблення програмного забезпечення та досвід використання комп'ютерно-орієнтованих технологій у вищих навчальних закладах, можна констатувати, для обрання ефективного варіанта з метою реалізації цілей освітніх вимірювань існує низка комп'ютерних технологій для підтримки педагогічного контролю та діагностики якості знань студентів. Проте попри значну їх кількість, є проблеми з їх адаптацією у вищих навчальних закладах України та високими витратами для отримання повного пакета послуг, оскільки програмне забезпечення з випробувальним періодом (демоверсія) не має потрібних функціональних характеристик і розроблених методик використання [52].

Розглянемо найпоширеніші комп'ютерно-орієнтовані технології для проведення педагогічного тестування, з метою контролю та діагностики процесу фахової підготовки студентів, які є вільно поширюваними:

*TSExam* [10] – система розроблена італійською компанією Tesnick. TSExam – це веб-інтерфейс, незалежний від платформи та мови, описано система мовою інтерпретатора PHP (Personal Home Page Tools – гіпертекстовий препроцесор, скриптова мова програмування), встановлюється безпосередньо на сервер. Система проста у використанні, користувачі отримують доступ до неї, застосовуючи довільний веб-браузер. Використовувати систему можна в двох режимах: адміністрування та проходження тесту. В режимі адміністрування можна поповнювати групи користувачів, призначати рівні доступу, створювати навчальні групи, а також створювати та редагувати тестові завдання. База даних системи побудована таким чином, що якщо тест було пройдено хоча б одним користувачем, то її (базу даних) видалити вже неможливо. Для кожного студента та навчальної групи ведеться статистика. Тестові завдання та статистичні дані можна зберігати в форматі .pdf (Portable Document Format – відкритий формат файла, створений і підтримуваний компанією Adobe Systems, для подання документів у вигляді, що незалежить від пристрою виведення та роздільних характеристик [9]).

*Система ТАО* [7] – система розроблена підрозділами Centre de Recherche Public Henri Tudor (Громадський науково-дослідний центр Анрі Тудор) в університетах Люксембурга та Франції. Систему створено з використанням мови PHP, XHTML (Extensible Hypertext Markup Language – мова розмітки гіпертексту, з тим самим призначенням, що й HTML, але відповідає синтаксичним правилам XML), CSS (Cascading Style Sheet – каскадні таблиці стилів, спеціальна мова, що використовується для відображення сторінок, описаних мовами розмітки даних. Найчастіше CSS використовують для візуальної презентації сторінок, описаних мовами

HTML та XHTML) та JS (JavaScript – динамічна, об'єктно-орієнтована мова програмування. Найчастіше використовується як частина браузера, що дає можливість взаємодіяти з користувачем, здійснювати керування браузером, асинхронно обмінюватися даними, змінювати структуру та зовнішній вигляд веб-сторінки). Для зберігання даних застосовується база даних MySQL (MySQL – вільно поширювана система управління реляційними базами даних). Система ТАО відома завдяки її використанню для проведення міжнародного тестування студентів PISA (Programme for International Student Assessment – міжнародна програма оцінювання освітніх досягнень учнів у сфері функціональної грамотності). Застосовуючи цю систему, можна створювати тестові завдання різних типів, завантажувати мультимедійні файли, створювати когнітивні тести та імпортувати різні елементи (тестові завдання, бази даних користувачів) у систему. В ній передбачено два режими: адміністративний та користувацький. Принципи роботи в адміністративному режимі подібний до *TCEexam*.

*OpenTEST 2* [61] – це комп'ютерна система для контролю знань, розроблена в Харківському національному університеті радіоелектроніки. Система *OpenTEST 2* функціонує з використанням зв'язків веб-сервера Apache (з англійської *a patchy server* – відкритий веб-сервер Інтернет для UNIX-подібних та інших операційних систем), PHP і MySQL. Система складається з кількох окремих модулів (модуль керування тестами, модуль керування тестуванням, модуль керуванням складом користувачів та модуль проходження тестів). HTML-код всіх модулів системи *OpenTEST 2* оптимізований для роботи в усіх сучасних веб-браузерах.

*Пакет програм для організації тестування MyTest* [7] – система розроблена російською компанією Klyaksa. Програму потрібно індивідуально інсталиувати на кожен ПК. Складається з таких модулів: модуль збору й аналізу результатів, що містить редактор тестів (*MyTestEditor*), модуль тестування (*MyTestStudent*) і журнал результатів тестування (*MyTestServer*). За допомогою *MyTest* можна створювати різні типи тестових завдань: множинний вибір, встановлення порядку проходження та відповідності, правильно/неправильно, числове, вибір місця на зображенні, заповнення пропусків.

*Програмна оболонка для створення тестових завдань Hot Potatoes* [62]. Авторами програми є С. Єрнейл та М. Холмс. Це програмна оболонка для створення тестових завдань. При цьому завдання створюються у форматі веб-сторінок, що дає змогу проводити тестування без застосування спеціальної програмної оболонки. Тому *Hot Potatoes* займає проміжну нішу між програмою, в яку вбудовано модулі для складання тестів, та *LCMS Moodle*. Ресурси, розроблені за допомогою *Hot Potatoes*, побудовані на веб-інтерфейсі.

*Mirax Test* [5] – система для проведення тестування від виробника Mirax Software, в яку крім професійних, логічних та IQ тестів вбудовано редактор, за допомогою якого можна створювати і редагувати власні тести. У тесті може бути необмежене число тестових завдань і є можливість завантаження мультимедійних файлів.

*EasyTest* [3] – система розроблена компанією Easy testmaker, це інтернет-тест генератор з можливістю створення, друку та публікації тестів он-лайн. Його недолік – несумісність із іншими програмними засобами, а також незахищеність даних користувача та результатів тестування, відсутня можливість аналізу якості педагогічного тесту та критеріїв оцінювання тестових завдань, на відміну від інших розглянутих систем.

*EasyQuizzy* [37] – програмний засіб від розробника NetCrate Software, за допомогою якого можна створювати і редагувати тести. Використовуючи підготовлені тести, можна полегшити роботу при проведенні щоденних опитувань студентів і прискорити перевірку результатів. Перевірка відбувається автоматично на базі системи оцінювання, обраної при створенні тесту. В *EasyQuizzy* можна працювати в двох режимах – режим редактора та режим тестування. Можна створювати такі типи тестових завдань: альтернативний вибір, вибір однієї правильної відповіді, вибір кількох правильних відповідей, встановлення послідовності, встановлення відповідності, есе. В програмі є підтримка мультимедійних файлів, експорту тестових завдань та різних шкал оцінювання.

До цікавих програмно-інструментальних і програмно-педагогічних продуктів, створених світовим лідером у комп'ютерному тестуванні Assessment Systems Corporation (ASC), можна віднести такі програми, як IteMan, Xcalibre, Lertap, CATSim, FastTest PC, однак вони мають два суттєвих недоліки: відсутність локалізації українською мовою та висока вартість самого програмного продукту.

*LCMS Moodle* – це модульна об'єктно-орієнтована система керування навчальними ресурсами. Платформу Moodle можна інстальовати в довільну операційну систему (MS Windows, Unix, Linux). Використання цієї платформи забезпечує студентам доступ до численних навчальних ресурсів. За допомогою послуги "Тест" у системі Moodle можна створити 10 різних типів тестових завдань, формувати банк тестових завдань (експорту та імпорту тестових завдань), реалізувати сценарій адаптивного тестування та скористатися системою аналізу тестових завдань і тесту, а також журналом оцінок з підтримкою різних шкал оцінювання. В *LCMS Moodle*, на відміну від попередніх систем, немає окремих модулів для адміністратора чи користувача, доступність до адміністративних функцій визначається встановленими правами користувачів.

На підставі результатів аналізу основних функціональних можливостей та характеристик запропоновано порівняльну характеристику наведених програмних засобів та виокремлено найліпший, для проведення педагогічного контролю та діагностики в процесі фахової підготовки студентів (Додаток ).

Виділимо основні функціональні характеристики комп'ютерно-орієнтованих технологій для реалізації завдань освітніх вимірювань:

1. Стандартні типи тестових завдань.
2. Додаткові типи тестових завдань.
3. Банк тестових завдань.
4. Обмеження в часі на проходження тесту.
5. Вагові коефіцієнти для тестових завдань.
6. Підтримка різних шкал оцінювання.
7. Облікові записи користувачів та груп.
8. Експорт та імпорт даних про користувачів.
9. Експорт та імпорт тестових завдань.
10. Експорт тесту для виконання на папері.
11. Імпорт сканованих бланків тесту.
12. Захист персональних даних.
13. Підтримка мультимедійних даних.
14. Засоби аналізу тесту та тестових завдань.
15. Створення резервних копій.
16. Простота використання.
17. Простота модифікації чи зміни функціонала.
18. Простота встановлення.
19. Зрозумілість.

Відповідно до встановленого рейтингу (Рис. 2.2) визначено, що найбільше задовольняють такі характеристики: LCMS Moodle (на першому місці), TCEam та Open Test2, відповідно на другому та третьому місцях. LCMS Moodle, на відміну від інших комп'ютерно-орієнтованих технологій, найбільше відповідає таким вимогам, в нього вбудовано широкий функціонал для забезпечення підтримки комп'ютерного тестування та аналізу якості педагогічного тесту.



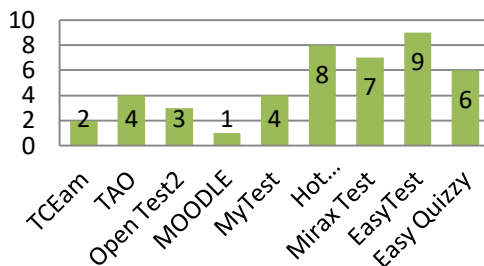


Рис. 2.2. Рейтинг комп'ютерно-орієнтованих технологій освітніх вимірювань

Наведемо переваги використання LCMS Moodle з урахуванням методичних та технічних чинників:

- вільна у розповсюдженні;
- відкритий внутрішній код;
- можливість встановлення на будь-яку операційну систему;
- використовувати можна як локально, так і в глобальній мережі;
- велика можливість модифікації;
- інтеграція з іншими програмними засобами;
- гнучка система організації дистанційного навчання;
- присутність засобів зворотного зв'язку (чат, форум, внутрішня та зовнішня пошта);
- вибір різних систем оцінювання та шкалювання;
- база даних користувачів, з можливістю моніторингу їхньої діяльності;
- локалізація різних мов;
- великий вибір типів тестових завдань;
- внутрішня система аналізу тестів і тестових завдань;
- можливість встановлення різних параметрів при конструюванні тесту.

Програмне забезпечення платформи описано мовою PHP з використанням безкоштовних загальнодоступних баз даних MySQL.

Беручи до уваги необхідність системного удосконалення методики використання комп'ютерних технологій для діагностики й оцінювання якості знань і формування практичних навичок їх використання викладачами (студентами), вважаємо за доцільне проводити тренінги налаштування та використання зазначеного програмного забезпечення за такою системою.

Перед використанням LCMS Moodle спершу її потрібно встановити на веб-сервері. Це може бути виділений сервер або організований на комп'ютері веб-сервер за допомогою програмного комплексу "Денвер" або "Open Server", у деяких версіях разом із інсталятором LCMS Moodle у

комплекті передбачено XAMPP (XAMPP – багатоплатформова збірка веб-сервера, що містить Apache, MySQL, інтерпретатор скриптів PHP, мову програмування Perl та велику кількість додаткових бібліотек, завдяки яким можна запустити повноцінний веб-сервер [13]), що суттєво полегшує встановлення та налаштування LCMS MOODLE. Повний пакет XAMPP містить[14]:

- Веб-сервер Apache з підтримкою SSL (Secure Sockets Layer – криптографічний протокол, який забезпечує встановлення безпечного з'єднання між клієнтом та сервером).
- СУБД MySQL (вільна система управління реляційними базами даних).
- Утиліту phpMyAdmin (веб-додаток з відкритим кодом на мові PHP із графічним веб-інтерфейсом для адміністрування СУБД MySQL).
- Інтерпретатор скриптів PHP.
- FTP-сервер FileZilla (багатоплатформовий FTP клієнт з відкритим кодом).
- Perl (Practical Extraction and Reporting Language – високорівнева, інтерпретована, динамічна мова програмування загального призначення).
- Сервлет-контейнер Apache Tomcat (Apache Tomcat – контейнер сервлетів, розроблений Apache Software Foundation, сервлет – стандартизований прикладний програмний інтерфейс для створення динамічного контенту до веб-сервера з використанням платформи Java).
- POP3/SMTP сервер (протоколи передачі електронної пошти).
- Панель керування XAMPP Control Panel.

Для того щоб встановити систему, потрібно завантажити з офіційного сайту <https://download.moodle.org/> і розархівувати архів "Moodle\_Windows\_Installer". В результаті буде створено директорій під назвою "moodle" з файлами та теками. Даний директорій можна розмістити в кореневий директорій веб-сервера, тоді вузол буде розміщений за адресою *http://адреса\_Вашого\_веб-серверу/moodle*, або скопіювати вміст архіву в кореневий директорій сервера, тоді вузол буде розміщений за адресою *http://адреса\_Вашого\_веб-серверу.com*. Вміст теки "moodle":

- config.php – файл конфігурації системи, що містить основні параметри і створюється під час встановлення системи;
- install.php – файл який містить сценарій, який потрібно запустити, щоб створити config.php;
- version.php – містить відомості щодо поточної версії LCMS Moodle;
- index.php – головна сторінка сайту;
- admin/ – містить файли для адміністрування системи;
- auth/ – містить модулі для реєстрації облікових записів

користувачів;

- blocks/ – містить модулі блоків, які використовують на багатьох сторінках;

- calendar/ – містить файли для подання й керування календарями;

- course/ – містить файли для подання й керування навчальними курсами;

- doc/ – містить файли з довідковими матеріалами для допомоги в роботі з LCMS Moodle;

- files/ – містить файли, що використовуються для перегляду й керування завантаженими файлами;

- lang/ – містить файли, для підтримки багатомовного інтерфейсу;

- lib/ – бібліотеки ядра LCMS Moodle;

- login/ – файли для керування обліковими записами користувачів;

- mod/ – основні модулі для створення навчальних курсів;

- pix/ – містить файли зображень сайту;

- theme/ – пакети тем для зміни зовнішнього вигляду сайту;

- user/ – файли для керування обліковими записами користувачів.

Наступний крок у налаштуванні – це створення бази даних у системі управління базами даних разом із обліковим записом користувача, який має право доступу до цієї бази даних. Це можна зробити, використовуючи веб-інтерфейс для керування базами даних, наприклад, phpMyAdmin. Після створення бази даних необхідно запустити програму інсталяції "install.php". Для цього потрібно у рядку адреси браузера ввести шлях до цього файла *http://адреса\_Вашого\_веб-серверу.com/moodle/install.php*. Потім необхідно встановити необхідні налаштування, які будуть збережені у файл конфігурації "config.php". Це дозволяє автоматичний перехід на сторінку встановлення (Рис. 2.3) та ознайомлення із "Загальною публічною ліцензією", а також підтвердження умови користування.

Цей випуск **Moodle Windows Installer** включає в себе програми для створення середовища, в якому **Moodle** працюватиме, а саме:

- Apache 2.2.17
- PHP 5.3.5 (VC6 X86 32bit) + PEAR
- MySQL 5.5.8 (Community Server)

Цей пакет також включає **Moodle 2.5.9 (Build: 20141110) (2013051409.00)**

Використання всіх частин цього пакету регламентується їх ліцензіями. Повний пакет **Moodle Windows Installer** є **відкритим програмним забезпеченням** і розповсюджується під ліцензією **GPL**.

Наступні сторінки будуть проводити вас через процедуру встановлення та налаштування **Moodle** на вашому комп'ютері. Ви можете прийняти автоматичне налаштування, а потім змінити параметри під себе.

Натисніть кнопку "Далі" для продовження встановлення **Moodle**.

Мова інтерфейсу | Українська (uk)

Далі >

Рис. 2.3. Сторінка встановлення

Оптимальним є такий подальший порядок дій: вихід на сторінку, яка містить форму, де потрібно встановити параметри для створюваного сайту (Рис. 2.4).

**Веб-адреса**

Повна веб-адреса, за якою ваш сайт moodle буде доступним. Moodle може мати тільки одну адресу доступу.

Якщо ваш сайт має кілька публічних адрес, то встановіть в DNS перенаправлення всіх інших адрес на цю.

Якщо до вашого сайту мають одночасний доступ і з Інтернету, і з Інтранету (локальної мережі), то забезпечте доступ з локального середовища за публічною адресою.

Якщо ви почали встановлення з неправильної адреси, видаліть все і почніть встановлення заново.

**Moodle Каталог**

Повний шлях до каталогу встановлення moodle.

**Каталог Даних**

Каталог, де Moodle буде зберігати всі файли, які завантажують користувачі.

Цей каталог повинен бути доступним для читання та запису для користувача, від імені якого запущено веб-сервер (зазвичай 'www-data', 'nobody', або 'apache').

Цей каталог не повинен бути доступним безпосередньо з Інтернету.

Програма встановлення спробує створити цей каталог, якщо його не існує.

---

Веб-адреса

Moodle Каталог

Каталог Даних

« Назад    Далі »

Рис. 2.4. Налаштування шляхів веб-адреси

Для налаштування бази даних необхідно вказати: сервер, ім'я бази даних, логін і пароль користувача бази даних та префікс бази даних (Рис. 2.5). Після запропонованих налаштувань відкриється сторінка з перевіркою елементів сервера (Рис. 2.6). Так запуститься сторінка, на якій буде відображатися процес автоматичного створення таблиць та оновлення бази даних системи.

Сервер бази даних

Ім'я бази даних

Користувач бази даних

Пароль бази даних

Префікс таблиць

« Назад    Далі »

Рис. 2.5. Налаштування бази даних

Після оновлення бази даних та встановлення головних налаштувань пропонується створити обліковий запис користувача – з правами адміністратора (Рис. 2.7). Після автентифікації облікового запису користувача завантажиться домашня сторінка сайту, на якій міститься меню адміністратора, за допомогою якого можна налаштувати головні параметри системи (користувачі, курси, журнал оцінок, локалізація, мова інтерфейсу).

Назва	Інформація	Звіт
unicode		ⓘ повинен бути встановлений і включений
database	mysql	ⓘ рекомендується версія 5.1.33, використовується версія 5.5.8
php		ⓘ рекомендується версія 5.3.3, використовується версія 5.3.5
pcreunicode		ⓘ повинен бути встановлений і включений для кращих результатів
php_extension	iconv	ⓘ повинен бути встановлений і включений
php_extension	mbstring	ⓘ повинен бути встановлений і включений для кращих результатів
php_extension	curl	ⓘ повинен бути встановлений і включений
php_extension	openssl	ⓘ повинен бути встановлений і включений для кращих результатів
php_extension	tokenizer	ⓘ повинен бути встановлений і включений для кращих результатів
php_extension	xmldr	ⓘ повинен бути встановлений і включений для кращих результатів
php_extension	soap	ⓘ повинен бути встановлений і включений для кращих результатів
php_extension	ctype	ⓘ повинен бути встановлений і включений
php_extension	zip	ⓘ повинен бути встановлений і включений
php_extension	gd	ⓘ повинен бути встановлений і включений

Рис. 2.6. Перевірка сервера

▼ Основне

Ім'я входу\*

Оберіть спосіб ідентифікації: ?  
 Ручна реєстрація  
 Ваш пароль повинен мати принаймні 8 символів, принаймні 1 цифр(а), принаймні 1 букв(а) нижнього регістру,  Зняти маску

Новий пароль\* ?

Примусити змінити пароль ?

Ім'я\*

Прізвище\*

Електронна пошта\*

Показувати ел пошту

Формат повідомлень ел пошти

Тип огляду

Автопідписка на форум

Під час редагування тексту

Місто\*

Країна\*

Часовий пояс

Мова, якій віддається перевага

Опис ?

Шлях: p

Рис. 2.7. Налаштування облікового запису адміністратора

Наступний крок це створення навчального курсу, який є головним елементом системи. Курс може складатися з необхідних навчальних

ресурсів та засобів організації навчальної діяльності студентів (елементи курсу). Якщо потрібно створити новий курс, то викладач (студент) має скористатись послугою "Замовити курс" (Рис. 2.8).

Рис. 2.8. Запит курсу

Після здійсненого запиту та підтвердження курсу користувачу (*Викладачу*) потрібно увійти в систему під своїм обліковим записом та налаштувати параметри курсу за своїми вподобаннями. Для цього у блоці "Адміністрування" потрібно вибрати послугу "Редагувати параметри" (Рис. 2.9).

Рис. 2.9. Редагування параметрів курсу

У меню "Редагувати параметри курсу" для керування курсом запропоновано таку методику дій з його параметрами:

- Загальне:

- *Коротка назва курсу* – обов'язкове поле.

- *Категорія курсу* – за допомогою цього параметра визначається категорія, в яку додається курс.

- *Доступність студентам* – за допомогою цього параметра визначається, чи буде курс відображатися у списку курсів. Окрім викладачів та адміністраторів, інші користувачі не зможуть увійти на курс.

- *Дата початку навчання* – за допомогою цього параметра визначається початок першого тижня при потижневому форматі курсу, а також дата початку подій у журналі для діяльностей курсу.

- *Ідентифікатор курсу* – використовується тільки при зіставленні курсу із зовнішніх систем і не відображається ніде на сайті. Якщо для курсу використовується офіційний ідентифікатор, то його можна ввести, інакше поле можна залишити порожнім.

- Опис:

- *Анотація курсу* – резюме курсу, яке відображається в списку курсів. При виведенні результатів пошуку курсів резюме курсу додається до назви курсу.

- *Файли анотації курсу* – такі, як зображення, відображаються у списку курсів разом з анотацією (якщо кількість курсів у категорії не перевищує 10), а також у результатах пошуку курсів.

- Формат курсу:

- *Формат* – визначає, який вигляд буде мати головна сторінка курсу.

- \* *SCORM-формат*: для показу SCORM-пакета в першій секції сторінки курсу (як альтернатива використовується модуль SCORM/AICC).

- \* *Формат-форум*: на першій сторінці курсу виводиться форум.

- \* *Формат-структура*: сторінка курсу організовується у вигляді тематичних секцій.

- \* *Потижневий формат*: сторінка курсу організовується у вигляді тижневих секцій, де перший тиждень відповідає даті початку курсу.

- *Кількість тижнів/тем* – за допомогою цього параметра визначається кількість розділів (секцій) у курсі.

- *Приховані секції* – за допомогою цього параметра визначається, чи будуть приховані секції відображатися для студентів у згорнутому вигляді чи повністю будуть невидимими.

- *Компонування курсу* – за допомогою цього параметра визначається, чи буде увесь курс виводитися на одній сторінці чи буде розділений на кілька сторінок.

- Вигляд:

➤ *Примусова мова* – параметр, за допомогою якого визначається мова курсу.

➤ *Кількість новин для показу* – за допомогою цього параметра визначається, скільки останніх новин буде з'являтися в блоці новин на сторінці курсу. Якщо встановлено значення "0 новин", то такий блок не буде відображатися.

➤ *Показувати студентам журнал оцінок*: параметр визначає, чи матиме можливість студент бачити список усіх своїх оцінок за курс у власному журналі оцінок (через посилання в блоці *Управління курсом*).

➤ *Показувати звіти про діяльність*: звіти про діяльність формується для кожного учасника, який виявляє активність на курсі. Вони виводяться у вигляді списку подій, таких як повідомлення на форумах, виконання завдань, а також включають події доступу до ресурсів. За допомогою цього параметра визначається, чи будуть студенти бачити власні звіти про діяльність на сторінці свого профілю.

• **Файли і завантаження:**

➤ *Файли курсу* – параметр, за допомогою якого можна дозволити завантажувати файли до курсу.

➤ *Максимальний обсяг файла для завантаження* – параметр, за допомогою якого визначається максимальний обсяг файла, який можна завантажити на курс, обмежений на рівні сайта налаштуваннями, встановленими адміністратором, а також налаштування максимального розміру файла, для подальшого обмеження його обсягу.

• **Відстеження виконання:**

➤ *Включити відстеження виконання* – параметр, за допомогою якого можна налаштувати завершення виконання діяльності, що відображається як на сторінці відстеження, так і в налаштуваннях елемента курсу.

• **Групи:**

➤ *Режим роботи з групами*. Цей параметр має три варіанти:

\*Немає груп – всі учасники є членами одного співтовариства.

\*Окремі групи – учасники кожної з груп працюють тільки в межах своєї групи, учасників інших груп вони не бачать.

\*Доступні групи – учасники кожної групи працюють тільки в межах своєї групи, але можуть бачити учасників інших груп. Груповий режим, визначений на рівні курсу, – режим за замовчуванням для всіх елементів, створюваних в курсі. Для кожного елемента, що підтримує груповий режим, можна вказати його власний груповий режим.

➤ *Примусовий режим роботи з групами*: у разі використання примусового групового режиму, його застосовується для всіх елементів курсу. Налаштування групового режиму кожного елемента ігноруються.



➤ *Групування за замовчуванням* – параметр, за допомогою якого можна налаштувати групування в курсі.

• *Перейменування ролі* – параметр, за допомогою якого можна перейменувати назву ролі в курсі. Нова назва ролі буде відображатися на сторінці учасників курсу та скрізь у цьому курсі. Якщо перейменована роль стосується керівних ролей курсу, то нова її назва буде відображатися в списку курсів.

За потреби можна уточнити параметри курсу, скориставшись довідковою системою, для цього потрібно натиснути кнопку допомоги відповідного параметра. Після вибору та налаштування параметрів необхідно натиснути кнопку "Зберегти зміни".

Для редагування курсу потрібно натиснути кнопку "Редагувати" у верхньому правому кутку сторінки. (Якщо така кнопка відсутня, то це означає, що користувач не є *Викладачем* даного курсу).

Процес створення тесту в LCMS Moodle починається з того, що у відповідній темі (як правило, тема охоплює одне заняття або один тиждень у курсі) потрібно за допомогою допоміжного вікна "Додати вид діяльності або ресурс" обрати елемент (вид діяльності) "Тест" (Рис. 2.10).

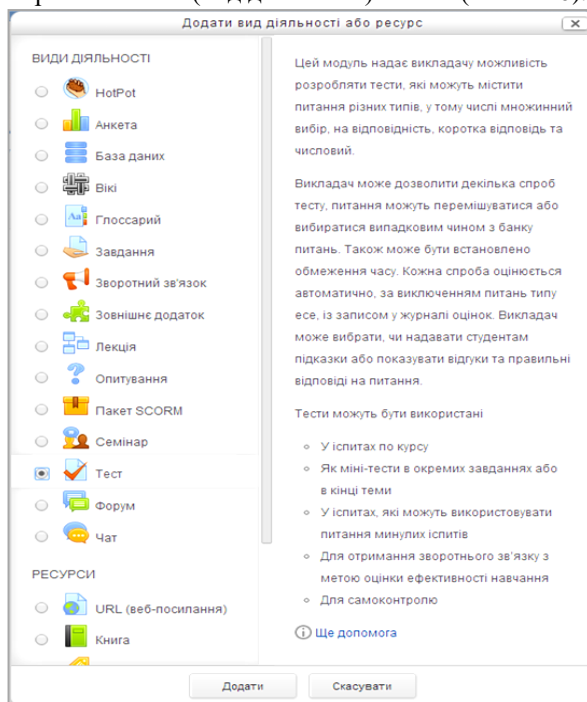


Рис. 2.10. Додавання видів діяльності або ресурсу

У відкритому вікні додається новий тест, в якому потрібно ввести та встановити основні його параметри (Рис. 2.13). Для ефективного функціонування тесту необхідно врахувати налаштування приведених нижче параметрів.

- У полі *Загальне* потрібно вказати *назву* (обов'язкове поле) та *вступ* (опис вступної частини до тесту. За потреби можна скористатися довідкою).

- У полі *Вибір часу* слід вказати дату і час початку та завершення тестування. В разі помилки студенти не зможуть виконати тестування. Студент не може виконати тест до початку тестування чи після його закінчення. Викладач у будь-який час може змінити ці налаштування, задавши обмеження в часі (час виконання тесту може бути обмежений: вказується кількість хвилин, відведених на виконання тесту). За замовчуванням час тестування не обмежується, студентові дається змога використовувати стільки часу, скільки необхідно для завершення тестування. Є можливість налаштувати дію, коли час на тест сплине (параметр "Коли час спливає", визначає, що станеться, коли студент не відправить результати своєї спроби проходження тестування, перш ніж закінчився час відведений на одну спробу. Якщо студент активно працює над тестом у момент завершення часу, то за сигналом таймера зворотного відліку результати спроби завжди будуть автоматично відправлятися на сервер, але якщо студент вийшов із системи, не відправивши результати, то цей параметр визначає, що повинно відбутися коли час відведений на спробу закінчився) та налаштувати *пільговий період подання* (якщо параметр "Коли час спливає" встановлений в значення "Надається пільговий період, коли готові відповіді можуть бути відправлені без можливості відповісти на інші завдання", то за допомогою нього параметра визначається час пільгового періоду).

- В полі *Оцінка* вказувалась категорія оцінки (параметр, за допомогою якого визначається категорія в журналі оцінок, в якій розміщуються оцінки за цей тест), кількість дозволених спроб (вказується кількість спроб для виконання тестових завдань) та метод оцінювання. Якщо студентам надано можливість кілька разів проходити тест, то можна різними способами обчислювати оцінку за тест:

- *Краща оцінка*: остаточною оцінкою вважається краща оцінка з усіх спроб.

- *Середня оцінка*: обчислюється середня оцінка всіх спроб.

- *Перша спроба*: у розрахунок приймається перша спроба.

- *Остання спроба*: результуючою оцінкою вважається оцінка останньої спроби.

Рис. 2.11. Додавання нового тесту

- У полі *Розміщення* налаштовувався *порядок запитань* (порядок виведення тестових завдань тесту під час кожного його виконання студентом змінюється, це запобігає переписуванню відповідей студентами один в одного) та *метод навігації*. Коли послідовну навігацію увімкнено, студент буде вимушений відповідати на тестові завдання послідовно і не зможе повернутися до попередньої сторінки або пропустити сторінку з можливістю повернення.

- У полі *Властивості питання* встановлюється *випадковий порядок відповідей (дистракторів)*. Якщо параметр увімкнений, то відповіді будуть перемішані випадковим чином за кожної спроби тесту. Цей параметр застосовується тільки до тестових завдань, які мають кілька відповідей, таких як тестове завдання з множинним вибором або тестове завдання на встановлення відповідності. Існує можливість налаштувати *режим питань* (наприклад, студенти повинні відповісти на кожне тестове завдання без отримання оцінки та коментаря, а потім переглянути результати всього тесту, - це режим "Відкладеного коментаря". Альтернативно, студенти, відповівши на кожне тестове завдання, одразу отримують коментар і, якщо з першого разу відповідь була неправильною, мають право на повторну спробу з можливістю отримання меншої оцінки, це режим "Інтерактивний, з кількома спробами". Ці два режими найчастіше використовуються для тестових завдань) та налаштовувати залежність кожної наступної спроби від попередньої (якщо дозволено повторні спроби, і цей параметр включений, кожна нова спроба тесту буде

містити результати попередньої спроби. Це дає змогу завершити тест за кілька спроб).

- У полі *Параметри перегляду* вказувалось, коли студенти можуть переглядати весь тест, чи відповідь є правильною, бали за тест, коментарі для окремих тестових завдань, коментарі для всього тесту, правильні відповіді, загальний коментар до результату:

- *протягом спроби;*
- *безпосередньо після спроби;*
- *пізніше, поки тест ще відкритий;*
- *після закриття тесту.*

- У полі *Показувати* налаштовувався показ фото та ім'я студента, кількість десяткових знаків у оцінці за тест і кількість десяткових знаків у оцінках для окремих тестових завдань. А також можна встановити відображення блоків під час проходження тесту.

- У полі *Додаткові обмеження в спробі* вказувалось кодове слово для проходження спроби, адреса комп'ютера/мережі користувача (доступ до тесту може бути обмежений, зокрема, заданою підмережею в локальній мережі або в мережі Інтернет введенням списку неповних або повних IP-адрес, розділених комами), затримка в часі між першою та другою спробами, можна встановити захист браузера. Якщо вибрано опцію "Спливаюче повноекранне вікно з деякими засобами безпеки JavaScript", то:

- тест розпочнеться тільки якщо веб-браузер підтримує JavaScript;
- тест буде виведено на весь екран у допоміжному вікні і не матиме елементів навігації;
- буде заборонено користуватися інструментами копіювання і вставка.

- У полі *Загальний коментар до тесту* викладачу необхідно ввести текст, який відображається після закінчення спроби проходження всього тесту. Вказуючи додаткові межі оцінки (у відсотках або у вигляді числа), можна визначити текст, який буде виведено. Цей текст містить пояснення щодо оцінки відповідей студента. Наприклад,

*Межа оцінки: 100 %.*

*Коментар: Дуже добре.*

*Межа оцінки: 40 %.*

*Коментар: Вам необхідно повторити теми цього тижня.*

*Гранична оцінка: 0 %.*

Студенти, які отримали бали в межах 40 % – 100 %, будуть бачити повідомлення "Дуже добре", а студенти, які отримали бали в межах 39,99 % – 0 %, інше повідомлення. Іншими словами, з урахуванням межі оцінки отримані бали розділяються на інтервали, а коментарі з'являються

залежно від інтервалу, до якого входить оцінка. Межі оцінки можуть бути подані як у відсотках, наприклад, – "1.41 %", тау і числом, наприклад, – 7.

• У полі *Загальні параметри модуля* налаштовується доступність для студентів певних блоків, ідентифікатор який дає можливість визначити діяльність для розрахунку балів та режим роботи з групами – параметр може мати три варіанти:

- *Немає груп*: всі учасники є членами одного великого співтовариства.

- *Окремі групи*: учасники кожної групи працюють тільки в межах своєї групи, інші групи для них невидимі.

- *Видимі групи*: учасники кожної групи працюють тільки в межах своєї групи, але можуть бачити інші групи. Якщо в курсі встановлений примусовий груповий режим, то налаштування групового режиму для будь-якого елемента курсу ігноруються.

Існує можливість налаштувати *групування* – це об'єднання груп в курсі. Якщо обрати групування, то студенти з вибраних груп будуть мати можливість працювати разом.

Після налаштування всіх параметрів (які можна в будь-який час змінити) потрібно натиснути кнопку "Зберегти і повернутися до курсу" або "Зберегти і показати". У разі вибору "Зберегти і показати", то відкриється сторінка редагування тесту із кнопкою "Редагування тесту".

При натисканні кнопки "Редагування тесту" відкриється сторінка з закладками "Редагування тесту" та "Порядок і сторінки". Головні налаштування тесту можна виконати за допомогою блоку *Адміністрування* розділ "Управління тестом".

Технічну поетапність реалізації дій в розділі "Управління тестом" наведено нижче:

• *Редагувати параметри* – послуга, за допомогою якої викладач має можливість налаштувати основні параметри тесту.

• *Поправки для груп*:

- *Поправки для групи*. Потрібно вибрати групу для якої встановлюються додаткові налаштування проходження тесту.

- *Кодове слово для проходження спроби*. Якщо кодове слово (пароль) вказано, то студенту з відповідної групи потрібно ввести його, щоб розпочати спробу тестування.

- *Почати тестування, Завершити тестування* – відповідно час початку та завершення тестування для групи.

- *Обмеження в часі*. Якщо увімкнено, то ліміт часу вказано на початковій сторінці тесту, а таймер зворотного відліку відображається в блоці навігації тесту.

- *Кількість дозволених спроб* – кількість спроб для проходження тесту для вибраної групи.

- *Поправки для користувачів:*

- *Поправки для користувачів.* Потрібно вибрати студентів, для яких встановлюються додаткові налаштування проходження тесту.

- *Кодове слово для проходження спроби.* Якщо кодове слово (пароль) вказано, то студентам потрібно ввести його, щоб розпочати спробу тестування.

- *Почати тестування, Завершити тестування* – відповідно час початку та завершення тестування для обраних користувачів.

- *Обмеження в часі.* Якщо ввімкнено, то ліміт часу вказано на початковій сторінці тесту, а таймер зворотного відліку відображається в блоці навігації тесту.

- *Кількість дозволених спроб* – кількість спроб для проходження тесту для обраних користувачів.

- *Редагування тесту.* Ця послуга містить дві закладки:

- *Редагування тесту.* На закладинці "Редагування тесту" можна змінювати максимальну оцінку за тест, додавати нові тестові завдання до банку тестових завдань, а також додавати до тесту готові тестові завдання з банку тестових завдань.

Як зазначено в методичних рекомендаціях [53], для того щоб заповнити тестовими завданнями блок "Банк тестових завдань", обиралася потрібна категорія тестових завдань, які потрібно додати в блок, і натиснути кнопку "Додати до тесту". Також можна додавати тестові завдання в довільному порядку. Для цього потрібно вибрати за допомогою випадачого списку Додати <число> випадкових тестових завдань кількість тестових завдань з необхідної категорії і натиснути "Додати до тесту".

- *Порядок та сторінки:* Обравши цю закладку можна сортувати тестові завдання на сторінці, переміщувати тестові завдання, робити перерозбиття сторінки, сортувати тестові завдання в тесті та додавати нові сторінки з тестовими завданнями.

- *Перегляд* – послуга, за допомогою якої викладач може переглянути тест.

- *Результати:*

- *Оцінки.* На цій сторінці можна обрати групи, які проходили тестування, та налаштувати параметри, які будуть входити до звіту.

- *Відповіді.* Функції на основі цього параметра збігаються з функціями попереднього параметра, за винятком можливості виводити на екран монітора:

- текст тестового завдання;

- відповідь;

- правильну відповідь.

➤ *Статистика* (див. «Аналіз якості показників тестових матеріалів»).

➤ *Оцінювання вручну* використовується у випадках, коли викладачу необхідно особисто оцінювати відповіді на тестові завдання (типу "Есе").

- *Локально призначені ролі.* Призначаючи роль користувачеві, йому надаються права, визначені в цій ролі, для поточного рівня і всіх "нижчих" рівнів. Наприклад, якщо користувачеві призначається роль студента в курсі, то він також буде мати роль студента в усіх елементах і блоках курсу.

- *Права.* За допомогою цієї послуги додавалися та вилучалися ролі користувачів для роботи з тестами та тестовими завданнями на рівні:

- курс;
- завдання, тест;
- види діяльності;
- звіт з ручного оцінювання;
- статистичний звіт тесту.

- *Перевірити права.* Використовуючи цю послугу, можна перевіряти права для окремих користувачів. Для цього в поле "*Знайти*" потрібно вписати прізвище користувача або вибрати користувача зі списку. Можна скористатися послугою "*Параметри пошуку*" для зручнішого пошуку. Для завершення перевірки прав користувача потрібно натиснути кнопку "*Показати права цих користувачів*", після цього відкриється сторінка на якій будуть відображені права користувача.

- *Фільтри.* На цій сторінці можна включити або відключити фільтри для певної частини сайту. За допомогою деяких фільтрів також можна встановити локальні налаштування, в цьому випадку поряд з його назвою буде посилання на налаштування цього фільтра.

- *Події.* Використовуючи дану послугу, можна переглядати за допомогою відповідної форми звіти про діяльність користувачів у тесті, в курсі та в цілому в LCMS Moodle. Дані виводяться у вигляді таблиці з такими полями: *Курс, Час, IP – адреса, Повне ім'я користувача, Дія та Відомості.*

- *Банк тестових завдань* містить такі пункти:

- *Запитання, де* відображаються всі тестові завдання курсу.

- *Категорії.* За замовчуванням створювалась категорія для всього курсу з назвою "*За замовчуванням "коротка назва курсу"*". В цій категорії можна додати нову категорію, а в створену категорію можна додати ще одну категорію тощо. Можна побудувати ієрархічну структуру категорій, так щоб в них додавати тестові завдання. Тоді останні для тесту можна буде легко обирати з різних категорій. Це корисно, якщо потрібно наприклад зробити заліковий тест, який складається із тестових завдань

усіх категорій, а кожна категорія містить тестові завдання з конкретної тематики. Є також можливість редагувати опис, переміщувати та вилучати категорію.

- *Імпорт*. Можна імпортувати тестові завдання із зовнішнього файла. Для їх імпорту спочатку потрібно обрати формат файла. Підтримується кілька форматів файлів: Aiken, Blackboard V6+, GIFT, Learnwise, Missing word, Moodle XML, WebCT, вбудовані відповіді.

- *Експорт*. Можна виконувати експортування тестових завдань із обраної категорії до текстового файла. При експортуванні деякі дані можуть втратитися. Це зумовлено тим, що у використовуваних форматах експортування не підтримуються усі можливості застосування тестового модуля LCMS Moodle. Деякі завдання не можуть бути експортовані зовсім. Існують також проблеми щодо правильного експортування символів кирилиці. Тому перед використанням експортування для перенесення тестових завдань потрібно протестувати можливість того чи іншого формату експорту. В системі Moodle підтримуються такі формати: GIFT, Moodle XML, XHTML.

Щоб створити нове тестове завдання, потрібно перейти до розділу "Банк тестових завдань", обрати потрібну категорію (якщо потрібної немає, то створити нову) та за допомогою кнопки "Створити нове тестове завдання" в допоміжному вікні, яке відкривається, обрати потрібний тип тестового завдання. Було запропоновано наступний порядок дій при створенні тестових завдань різних типів (Рис. 2.12):

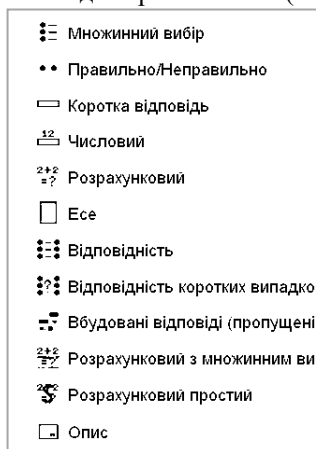


Рис. 2.12. Вибір типу тестового завдання

Основні кроки та параметри налаштування різних типів тестових завдань однакові (Рис. 2.13), однак деякі з них розрізняються залежно від особливості тестового завдання й налаштовуються індивідуально:



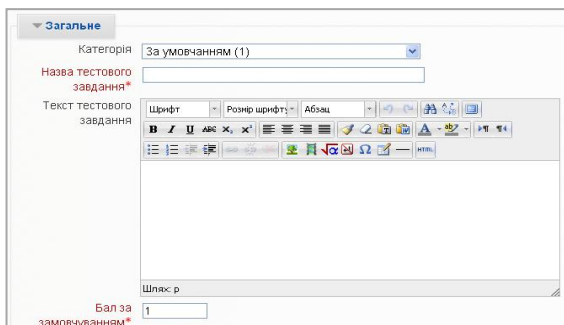


Рис. 2.13. Приклад налаштування загальних параметрів тестового завдання

- Тестове завдання типу *"Множинний вибір"*. За допомогою такого типу завдань можна створювати два типи завдань: вибір однієї правильної відповіді та вибір кількох правильних відповідей.

- Тестове завдання типу *"Правильно/Неправильно"*. За допомогою цього типу завдань можна створювати завдання на вибір з однією правильною відповіддю тільки за двома варіантами (дистракторами) вибору: *"Правильно"* та *"Неправильно"*.

- Тестове завдання типу *"Коротка відповідь"* – завдання, відповідь на яке користувач повинен ввести з клавіатури.

- Тестове завдання типу *"Числове"* – тестове завдання, на яке очікується певна числова відповідь. Цей тип тестового завдання є окремим випадком тестового завдання з короткою відповіддю. Своєрідність тестового завдання у тому, що відповідь повинна бути подана у вигляді числа.

- Тестове завдання типу *"Розрахункове"* – тестове завдання, на яке очікується конкретна числова відповідь (може бути також певна кількість відповідей). Використання розрахункових питань дає змогу створювати індивідуальні числові завдання з використанням символів підстановки (шаблонів) у фігурних дужках, які вводяться з індивідуальними значеннями, коли тест запускається. Наприклад, питання: "Яка площа прямокутника довжиною  $\{L\}$  і шириною  $\{W\}$ ?", має містити коректну формулу для відповіді  $\{=\{L\} * \{W\}\}$  (де \* означає множення).

- Есе (фр. *essai* – спроба, проба, нарис, від лат. *exagium* – зважування) – літературний жанр прозового твору невеликого обсягу й вільної композиції. У тестовому завданні типу "Есе" потрібно описати відповідь (вона може включати зображення) у вигляді есе. Таке тестове завдання оцінюється не автоматично, а особисто викладачем після перегляду відповіді.

▪ Тестові завдання "*На відповідність*" потребують від студента коректного зіставлення назви або речення одного списку з назвами або реченнями іншого списку.

▪ Для студента тестове завдання типу "*Відповідність випадкових коротких відповідей*" має такий самий вигляд, як тестове завдання типу "*Відповідність*". Різниця в тому, що перелік питань для відповідності вибирається випадковим чином з питань типу "*Коротка відповідь*" у даній категорії. В категорії має бути достатня кількість завдань типу "*Коротка відповідь*", інакше буде відображатися повідомлення про помилку.

▪ Тестові завдання типу "*Вкладені відповіді*" складається з тексту, в якому треба вставити пропущені слова за змістом. Кожне пропущене слово може бути відповіддю типу "*Коротка відповідь*", "*Числове*" або "*Множинний вибір*". Структура пропущеного слова має таке подання:

- "{" – початок слова;
- "N" – необов'язкове число, кількість балів за правильну відповідь;
- ":ТИП\_ПИТАННЯ:" – визначає форму подання відповіді;
- "~" – розділяє варіанти відповіді;
- "=" – ставиться перед правильним варіантом відповіді;
- "#" – ставиться перед початком необов'язкового коментаря;
- "}" – кінець слова.

Типи питань можуть бути такими:

- SA або SHORTANSWER – коротка відповідь (регістр не враховується);
- SAC або SHORTANSWER\_C – коротка відповідь (регістр враховується);
- NM або NUMERICAL – числове (відповідь у форматі Число:Похибка);
- MC або MULTICHOICE – множинний вибір (у вигляді списку);
- MCV або MULTICHOICE\_V – множинний вибір (радіокнопки вертикально);
- MCH або MULTICHOICE\_H – множинний вибір (радіокнопки горизонтально).

Наведемо деякі приклади:

1. Відоме місто Цюрих знаходиться у {1:MC:Швеції~Швейцарії#Правильно~Люксембурзі}.
2. В Україні час влітку відрізняється від GMT (часу за Грінвічем) на {1:MCH:1~2~3} години.
3. На вірші поеми "Руслан і Людмила" великого російського поета {1:SAC:%100%Пушкіна~%50%Пушкіна~%50%Пушкіна} композитор {2:SA:=Глінка} написав однойменну оперу.
4. Прискорення вільного падіння у метрах за секунду становить {2:NUMERICAL:=9.8:0.02# Чудово~%50%10:0.18#Близько до істини}.

▪ Тестове завдання *"Розрахункове з множинним вибором"* подібне до тестового завдання типу *"Множинний вибір"*, в якому варіанти відповідей можуть містити розрахункові формули з використанням символів підстановки (шаблонів) у фігурних дужках, які вводяться з індивідуальними значеннями, коли розпочинається тест. Наприклад, якщо є тестове завдання: "Яка площа прямокутника довжиною  $\{L\}$  і шириною  $\{W\}$ ?", то один з варіантів відповідей повинен містити правильну формулу  $\{ = \{L\} * \{W\} \}$  (де  $*$  - множення).

▪ За допомогою тестового завдання типу *"Розрахункове просте"* створюються окремі числові тестові завдання, в яких застосовується символи підстановки (шаблони), що замінюються окремими значеннями, коли розпочинається тест. У тестовому завданні *"Розрахункове просте"* пропонуються найбільш використовувані властивості тестового завдання типу *"Розрахункове"* з простим інтерфейсом для створення таких типів тестових завдань.

▪ Тестове завдання типу *"Опис"* насправді не є тестовим завданням. За його допомогою виводиться деякий текст, на який не потрібно відповідати. Текст даного типу завдання виводиться як під час спроби, так і на сторінці перегляду. Будь-які загальні коментарі відображаються тільки на сторінці перегляду. Використання типу *"Опис"* дає змогу додавати текст до тесту (наприклад, розповідь чи статтю), до якого потім можна поставити запитання чи одразу оцінити. Іншими словами, це засіб структурування тестових завдань при виведенні їх студентам. Застосовуючи його у тесті можна відокремити групу завдань і супроводити це відокремлення текстовими поясненнями чи графічними зображеннями.

На сторінці з оцінками наявні всі відомості про будь-яку спробу проходження тесту кожним студентом (Рис. 2.14).

Для налаштування перегляду оцінок слід використовувати блоки з параметрами показу оцінок:

- *Що включити в звіт.*
- *Параметри показу.*

На цій самій сторінці містяться такі кнопки для оцінювання тесту:

*"Переоцінити все"* – викладач може переоцінити тест, якщо в параметри тесту було внесено зміни вже після проходження його студентами.

*"Виявити усі спроби для переоцінки"* – викладач, скориставшись цією послугою, може виявити спроби студентів, які не були оцінені, та переоцінити їх.

*"Переоцінити виділені спроби"*. Цей засіб використовується для того, щоб заново обчислити оцінки, якщо вже після виконання студентами тесту викладач вніс до нього зміни, що впливають на оцінку. Наприклад,

викладач виправив помилку в позначенні правильної відповіді в одному з тестових завдань, змінив максимальну оцінку за тест або вагу деяких завдань. Доти, поки відповіді студентів зберігаються в базі даних, перерахувати оцінки викладач може коли завгодно, лише натиснувши на посилання "*Переоцінити виділені спроби*". Тоді буде виведено список спроб на тестові завдання. Але потрібно враховувати, що перерахунок оцінок великої кількості студентів може тривати досить довго. Спроби, які змінилися після переоцінювання, подаються у вигляді гіперпосилання, натиснувши на які, можна відкрити сторінку перегляду цих спроб.

Прізвище / Ім'я	Електронна пошта	Стан	Розпочато	Завершено	Витрачений час	Оцінка/100,00	П. 1 /4,76	П. 2 /4,76	П. 3 /4,76	П. 4 /4,76
Собкова Олена Перегляд спроби	sobkova_o@mail.ru	Завершено	27 грудня 2011 7:16	18 січня 2012 10:52	21 днів 15 години	71,43	× 0,00	✓ 4,76	× 0,00	✓ 4,76
Козленко Оля Перегляд спроби	kozlenko_olya@mail.ru	Завершено	2 січня 2012 11:47	8 січня 2012 2:33	6 днів 2 години	74,36	× 0,00	✓ 4,76	✓ 1,50	✓ 4,76
Колеско Юлія Перегляд спроби	kolenko_julia@rambler.ru	Завершено	3 січня 2012 3:38	3 січня 2012 3:51	12 хв 41 сек	36,19	× 0,00	× 0,00	✓ 2,39	× 0,00
Загальна середня						59,11 (115)	3,39 (110)	2,53 (110)	2,73 (110)	2,83 (110)

Рис. 2.14. Перегляд результатів проходження тесту. Оцінки

На рис. 2.16. наведено засіб, за допомогою якого здійснюється фільтрація студентів за групами (якщо в даному курсі передбачено поділ за групами). При цьому виникає можливість переглядати список, починаючи з певної літери прізвища або імені, хоча для кириличної абетки ця функція, як і функція сортування, не завжди є коректною. Дані таблиці можна сортувати за значеннями в будь-якому стовпці. Викладач має змогу вилучити непотрібні рядки таблиці, для цього проти них потрібно поставити на позначки  і натиснути кнопку "*Вилучити виділені спроби*".

Результати проходження тестових завдань доступні для перегляду в графічному поданні (у вигляді гістограми), результати виводяться після таблиці з оцінками (Рис. 2.15).

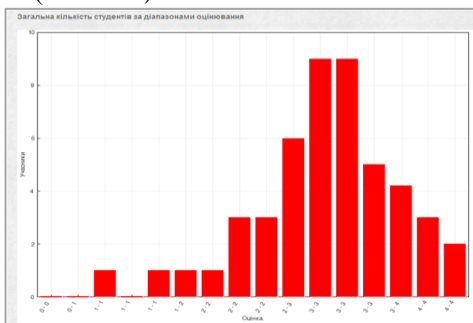


Рис. 2.15. Перегляд результатів проходження тесту (гістограма)

Для оцінювання тестового завдання "Есе" чи "Опис" існує послуга "Оцінювання вручну". На сторінці "Оцінювання вручну" міститься тестове завдання і відповідь студента на нього. Викладач, враховуючи контингент, повинен обрати кількість спроб для оцінювання у випадковому списку "Спроб для оцінювання", встановити кількість тестових завдань на сторінці та відсортувати спроби за датою, ім'ям студента, ID-студента або випадковим чином. За бажання можна написати студенту свій відгук (коментар) на його відповідь у полі "Коментар". У полі *оцінка* слід ввести потрібну оцінку і натиснути кнопку "Зберегти і перейти до наступної сторінки".

Окрім традиційного тесту, виокремлюють адаптивний тест. Нижче розглянемо суть і цілі адаптивного тестування, а також наскільки можливо реалізувати сценарій адаптивного тестування за допомогою комп'ютерно-орієнтованих технологій.

### **2.2.2. Реалізація адаптивного тестування засобами комп'ютерно-орієнтованих технологій освітніх вимірювань**

Для реалізації технологій адаптивного комп'ютерного тестування потрібно створити умови для об'єктивізації інтегральних результатів педагогічного моніторингу, корегування індивідуальних рівнів навченості та керування якістю навчання студентів.

Зважаючи на різні формулювання поняття адаптивного тестування на нашу думку, найкоректнішим є формулювання наведене у працях П.І. Федорука [70], "... адаптивне тестування – це тип педагогічного тестування, за якого порядок подання тестових завдань залежить від відповідей того, хто проходить тестування, на попередні запитання". Головними перевагами використання адаптивних тестів є такі, які дають змогу:

- підлаштуватися під індивідуальні можливості студента;
- підвищити точність оцінювання рівня знань сильних і слабких студентів завдяки використанню більшого за обсягом банку тестових завдань різного рівня складності;
- зменшити тривалість тесту та кількість тестових запитань, необхідних для досягнення достатньої точності оцінки рівня знань студента;
- знизити ступінь втоми студента;
- забезпечити конфіденційність за рахунок надання кожному студенту індивідуального набору тестових завдань, що відповідають його рівню знань;
- спростити процедуру внесення змін у банк тестових завдань, які будуть автоматично враховані адаптивним алгоритмом.

Оцінюючи різні цілі адаптивного тестування вважатимемо, що

головними з них є: спрямування студентів на самостійну роботу за допомогою адаптивних контрольно-навчальних програмних засобів, індивідуалізація навчання та орієнтація в навчальному процесі на адаптивні методи навчання і контролю (Рис. 2.16).



Рис. 2.16. Головні цілі адаптивного тестування

Можливість використання адаптивного тестування в комп'ютерно-орієнтованих технологіях освітніх вимірювань об'єднала б переваги цих двох технологій. На думку авторів, найважливішим напрямом модернізації та інформатизації професійної підготовки майбутнього фахівця є створення комп'ютерного-діагностичного інструментарію як ефективного засобу керування якістю освіти, що інтегрує у собі функцію контролю знань, корекцію траєкторії навчання, діагностичну складову та індивідуальний підхід до організації навчального процесу.

Оскільки адаптивне навчання являє собою педагогічну систему форм і методів, це сприяє ефективному індивідуальному навчанню. Як зазначалося [55], ця система краще за інші враховує рівень і структуру початкової підготовки й оперативно відстежує результати поточної підготовки, що в свою чергу дає змогу раціонально підбирати тестові завдання та вправи.

Існують різні варіанти адаптивного тестування. Виділимо з них три головних:

- Перший називають пірамідальним тестуванням. За відсутності попередніх оцінок усім студентам надається завдання середньої складності, а потім залежно від відповіді кожному тестованому додається легше чи важче тестове завдання.

- Другий варіант (flexilevel) полягає в тому, що початок тестування здійснюється з будь-якого обраного студентом рівня складності, з поступовим наближенням до реального рівня знань.

- Третій варіант (stradaptive) передбачає тестування за допомогою банку завдань, методично розділених за рівнями складності. За правильної

відповіді наступне тестове завдання береться з верхнього рівня, а за неправильної – з нижнього.

Найвідомішими системами реалізації адаптивного тестування є: IBM Lotus Workplace Collaborative Learning [4], Віртуальний університет [1], "Веб-Клас-ХПІ [11]". Усі ці комп'ютерно-орієнтовані засоби дають змогу зберігати дані студента в окремому профілі, який містить персональні відомості, моніторинг його активності, час роботи в системі, що сприяє об'єктивності та відкритості, а також адаптації до групи на базі даних профілю. Коли йдеться про адаптивність систем у цілому, то в більшості випадків вона стосується стратегій навчання і реалізується за рахунок попереднього формування автором сценаріїв для окремих груп студентів. Інакше – керування виконується автоматизовано на основі прогресу у вивченні елементів курсу та організації правил переходу.

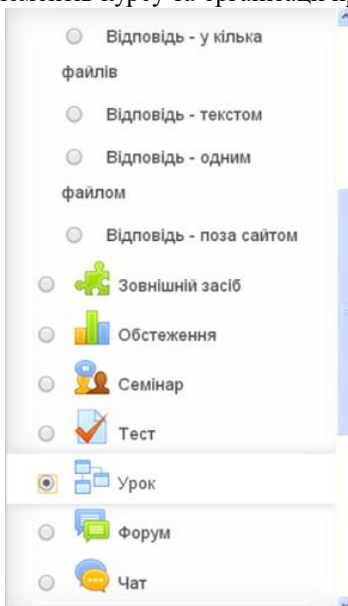


Рис. 2.17. Додавання елементів курсу

Було запропоновано варіант поетапної реалізації адаптивного тестування стандартними елементами LCMS Moodle та окреслено основні кроки створення адаптивного тесту, які ґрунтуються на використанні базового тесту середньої складності. При виконанні тесту на 100 % студент переходить до тесту вищої складності, при певній кількості неправильних відповідей (скажімо, 75 % правильних) відбувається перехід до тесту нижчої складності. Такий механізм можна реалізувати за допомогою елемента системи LCMS Moodle "Урок" (у деяких варіантах перекладу замість назви "Лекція" вживається назва "Заняття" або "Урок").

На першому етапі потрібно додати новий тип діяльності – "Урок" (Рис. 2.17, Рис. 2.18). У полі "В залежності від" потрібно обрати залежність від іншого елемента Moodle "Лекція – Урок".

▼ **Загальне**

Назва\* **Ви повинні вказати значення тут.**

▶ **Вигляд**

▶ **Доступність**

▼ **Prerequisite lesson**

В залежності від

Проведений час

Завершено

Рівень кращий ніж (%)

▶ **Управління обміном даних**

▶ **Оцінка**

▶ **Загальні параметри модуля**

[Зберегти й повернутися до курсу](#) [Зберегти й показати](#) [Скасувати](#)

Рис. 2.18. Додавання нового уроку

Для тесту низької складності обирали базовий тест, пройдений на 75 %, а для тесту високої складності – базовий тест, пройдений на 100 % (Рис. 2.19, Рис. 2.20).

▼ **Загальне**

Назва\*

▶ **Вигляд**

▶ **Доступність**

▼ **Prerequisite lesson**

В залежності від

Проведений час

Завершено

Рівень кращий ніж (%)

Рис. 2.19. Налаштування зв'язку тесту низької складності з базовим тестом



▼ **Загальне**

Назва\*

► **Вигляд**

► **Доступність**

▼ **Prerequisite lesson**

В залежності від

Проведений час

Завершено

Рівень кращий ніж (%)

Рис. 2.20. Налаштування зв'язку тесту високої складності з базовим тестом

Ці три елементи "Лекція – Урок": базовий тест, тест низької складності і тест високої складності, складаюся тільки з тестових завдань (Рис. 2.21). За необхідності інші тести, крім базового можна приховати, натиснувши на кнопку "Сховати" (Рис. 2.22).

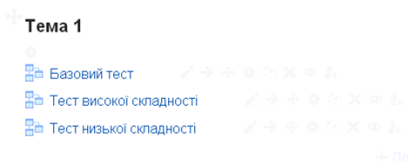


Рис. 2.21. Додавання тестів



Рис. 2.22. Панель швидкого редагування уроку

Розглянемо покроково на прикладі базового тесту додавання тестових завдань до уроку. Щоб створити нові тестові завдання, потрібно перейти в меню редагування, а потім до розділу "Додати сторінку запитань" (Рис. 2.23).



Рис. 2.23. Додавання тестових завдань

Щоб створити тестове завдання, слід обрати з випадаючого списку потрібний тип тестового завдання (Рис. 2.24).

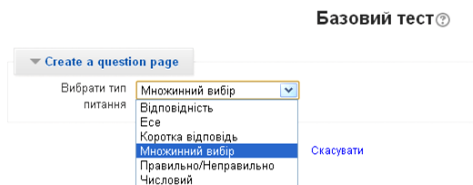


Рис. 2.24. Типи тестових завдань

Далі, обравши потрібний тип тестового завдання, потрібно заповнити та встановити його параметри у відповідних полях (Рис. 2.25).

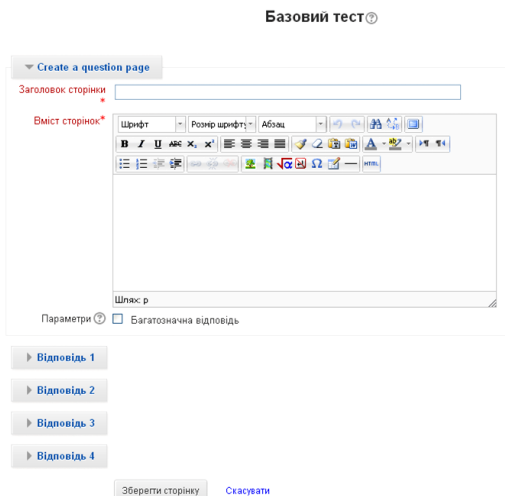


Рис. 2.25. Вікно редагування тестового завдання

Такий механізм доступний як у Moodle 1.9.x, так і Moodle 2.x. Крім того, в Moodle 2.x. можна встановлювати залежність від відсотка проходження попередніх тестів. Для Moodle 1.9.x є спеціальний плагін – QuizPort. Він працює лише з тестами Hot Potatoes і Qedos. Головна характеристика цього модуля – можливість встановлювати залежність від проходження інших тестів.

Однією з головних характеристик адаптивного тестування є складність тестових завдань, і не будь-яка, а математично обґрунтована. Проводячи комп'ютерне адаптивне тестування з використанням модульного об'єктно-орієнтованого динамічного навчального середовища LCMS Moodle, яке спрямоване на пришвидшення статистичного опрацювання результатів тестування і визначення кількісних статистичних

характеристик тестових завдань, можна оперативно визначити прогалини в знаннях студентів. Проаналізувавши найпоширеніші комп'ютерні системи діагностики знань студентів, з'ясували, що лише мала кількість комп'ютерно-орієнтованих технологій освітніх вимірювань здатні реалізувати адаптивне автоматизоване тестування. Крім того, в них не реалізовано автоматичний підбір тестових завдань різної складності та програмний набір методів для адаптації груп.

### **2.2.3. Аналіз якості показників тестових матеріалів з використанням комп'ютерно-орієнтованих технологій**

Під час контролю знань студентів значну увагу приділяють вибору методів і засобів діагностики, а також добору інструментарію. Одним з найуживаніших методів діагностики та контролю якості знань студентів у сучасній системі освіти є тестування. Сучасна індустрія програмних засобів надає великий вибір безкоштовних і комерційних дистрибутивів, які дають змогу здійснювати автоматизований контроль знань. Однак можливість вибору програмних засобів не забезпечує автоматизованого аналізу відповідності встановлених норм тесту та якості тестових завдань. Принцип роботи такої автоматизованої системи ґрунтується на математико-статистичному аналізі, що дає змогу перевіряти тестові завдання на наявність прихованих похибок, які неможливо виявити, застосовуючи експертні методи.

Методика опрацювання результатів тестування в LCMS Moodle ґрунтується на математико-статистичних методах, що дають змогу визначати якість тестових матеріалів. У процесі роботи було проаналізовано вбудовану систему аналізу тестових завдань, використаних в ній математико-статистичних методів, визначено ефективність її застосування під час загальній характеристики тесту [51].

Результати аналізу тесту відображаються в таблиці “Відомості про тест” (Рис. 2.26). Її можна переглянути, перейшовши на вкладку “Статистика”.

На сторінці “Статистика” у випадяючому списку “Доступні групи” викладач самостійно може обирати потрібну групу, яка проходила тестування, а також потрібні спроби студентів у списку “Показувати статистику”, та натиснути кнопку “Показати звіт”.

Назва тесту	Комп'ютерні мережі та Інтернет. HTML та CSS
Назва курсу	Комп'ютерні мережі та Інтернет. HTML та основи CSS.
Кількість перших повністю оцінених спроб	39
Загальна кількість повністю оцінених спроб	39
Середня оцінка по перших спробах	76%
Середня оцінка по всім спробам	76%
Медіана оцінок (для першої спроби)	81%
Стандартне відхилення (для першої спроби)	17%
Значення асиметрії розподілу (для першої спроби)	-1,51
Значення ексцесу розподілу (для першої спроби)	2,47
Коефіцієнт внутрішньої узгодженості (для першої спроби)	92%
Співвідношення помилок (для першої спроби)	29%
Стандартна помилка (для першої спроби)	5%

Рис. 2.26. Відомості про тест

На сторінці будуть відображені відомості про тест:

– Назва тесту.

– Назва курсу.

– Кількість перших повністю оцінених спроб.

– Загальна кількість повністю оцінених спроб.

– Середня оцінка за першими спробами.

– Середня оцінка за всіма спробами (у випадках коли на проходження тесту надається кілька спроб, ставиться середня оцінка за використаними спробами).

– Медіана оцінок (для першої спроби) – значення ознаки, яке поділяє елементи рангової вибірки, розміщеної в порядку зростання або спадання ознаки, на дві рівні частини. Медіана обчислюється за формулою:

$$M_e = x_{Me} + i_M \frac{\frac{\sum f}{2} - S_{Me-1}}{f_{Me}}, \quad (2.1)$$

де  $x_{Me}$  – нижня границя медіанного інтервалу;  $i_M$  – медіанний інтервал;  $S_{Me}$  – сума спостережень, що була накопичена до початку медіанного інтервалу;  $f_{Me}$  – кількість спостережень у медіанному інтервалі.

– Стандартне відхилення (для першої спроби) або середнє квадратичне відхилення, що позначається  $\sigma$  або  $S$  – показник розсіювання значень випадкової величини відносно її математичного сподівання. Формула для обчислення стандартного відхилення:

$$\sigma = \sqrt{\frac{\sum_{j=1}^n (x_j - \bar{x})^2}{n}}, \quad (2.2)$$

де  $\bar{x}$  – середнє значення вибірки;  $n$  – кількість елементів вибірки.

Середнє квадратична величина використовується для визначення показників варіації досліджуваної ознаки – дисперсії та середнього

квадратичного відхилення. Обчислюється на основі квадратів відхилень індивідуальних значень ознаки (тестових балів) від їх середнього значення.

– Значення асиметрії розподілу (для першої спроби) – числова характеристика розподілу ймовірностей дійсної випадкової величини. Цей критерій застосовують для приблизної перевірки гіпотези про нормальність емпіричного розподілу. Визначення асиметрії дає змогу встановити симетричність розподілу випадкової величини  $x$  відносно математичного сподівання  $M(x) = 1$ . Для цього знаходять третій центральний момент, що характеризує асиметрію закону розподілу випадкової величини. Вважається, що за нормального розподілу вибіркові показники асиметрії дорівнюватимуть нулю, але на практиці така ситуація не спостерігається. Якщо  $\mu_3 = 0$ , то випадкова величина  $x$  симетрично розподілена відносно математичного сподівання  $M(x)$ . Оскільки  $\mu_3$  має розмірність випадкової величини в кубі, то вводять безрозмірну величину – коефіцієнт асиметрії:

$$A_s = \frac{\mu_3}{\sigma^3}. \quad (2.3)$$

- якщо асиметрія від'ємна, то крива нормального розподілу зміщена вліво;

- якщо асиметрія додатна, то крива нормального розподілу зміщена вправо.

– Значення ексцесу розподілу (для першої спроби) – міра розподілу одиниць сукупності учасників тестування за значенням ознаки (тестових балів). Коефіцієнт ексцесу характеризує "крутість", тобто стрімкість зростання кривої розподілу порівняно з нормальною кривою. Для нормального розподілу він дорівнює нулю. Якщо ексцес деякого розподілу відмінний від нуля, то крива його щільності відрізняється від кривої щільності нормального розподілу:

- якщо ексцес додатний, то крива теоретичного розподілу має вищу та "гострішу" вершину, ніж крива нормального розподілу;

- якщо ексцес від'ємний, то крива теоретичного розподілу має нижчу та "плоскішу" вершину, ніж крива нормального розподілу.

При цьому вважається, що нормальний і теоретичний розподіли мають однакові математичні сподівання та дисперсії [28].

Ексцес обчислюють за формулою:

$$E_s = \frac{\mu_4}{\sigma^4} - 3. \quad (2.4)$$

Число три віднімається для порівняння відхилення від центрального закону розподілу (нормального закону), для якого справджується рівність.

$$\frac{\mu_4}{\sigma^4} = 3. \quad (2.5)$$

Отже,  $E_s = 0$  для нормального закону розподілу.

Якщо  $E_s > 0$ , то на графіку функція розподілу має гостру вершину,

а якщо  $E_s < 0$  – більш полого. В такий спосіб можна встановити відхилення заданого закону від нормального.

– Коефіцієнт внутрішньої узгодженості (для першої спроби) – призначений для оцінювання внутрішньої узгодженості (тестових завдань із тестом), має відображати структурну ієрархію моделі підготовки з навчальної дисципліни. Внутрішня узгодженість тесту – характеристика тесту, яка вказує на ступінь однорідності завдань тесту. Внутрішня узгодженість є істотним елементом конструктивної валідності тесту і характеризує, якою мірою завдання спрямовані на вимірювання потрібної ознаки, а також міру внеску кожного завдання в кількісну оцінку всієї методики. Максимальна валідність тесту досягається за рахунок добору таких завдань, які, перебуваючи в значущій кореляції з результатом тесту, водночас мінімально корелюють між собою (за невиправдано високої кореляції окремих завдань, тест буде переобтяжений зайвими практично однозначними пунктами).

– Співвідношення помилок (для першої спроби) – розбіжність значень досліджуваної ознаки визначається за кількома параметрами. Перший вказує на те, на скільки одні студенти кращі ніж інші, другий – деяка випадкова варіація. Співвідношення помилок показує, скільки є випадкових варіацій: чим менше їх тим кращий тест.

– Стандартна помилка (для першої спроби) – це стандартне відхилення оцінок, які будуть отримані при багаторазовій випадковій вибірці певного обсягу з однієї і тієї самої сукупності. Стандартна помилка – це спадна функція обсягу вибірки: чим менша стандартна помилка, тим достовірнішою є оцінка [20]. Обчислюється вона за формулою:

$$m = \frac{s}{\sqrt{n}}, \quad (2.6)$$

де  $s$  – стандартне відхилення;  $n$  – обсяг вибірки.

Над таблицею загальних відомостей про тест міститься випадуючий список “Завантажити повний звіт як”. Обравши потрібний формат, викладач може завантажити звіт на власний комп’ютер. При цьому можна обрати декілька форматів файлів (електронна таблиця Excel \*.xls; електронна таблиця OpenDocument \*.ods; текстовий формат, в якому значення розділені табуляцією – звичайний текстовий файл, в якому рядки – це рядки таблиці, а колонки розділяються табуляцією та ін.).

Проаналізувавши тестові завдання, одержуємо детальнішу структуру тесту, яку можна переглянути в таблиці “Аналіз структури тесту” (Рис. 2.27).

Таблиця аналізу структури тесту містить такі поля:

– № тестового завдання – унікальний номер тестового завдання у системі Moodle. Якщо натиснути на його номер, то відкриється сторінка редагування цього завдання. Під номером є дві піктограми: одна вказує на

тип завдання, а за допомогою іншої можна переглянути це завдання в окремому вікні;

Назва питання	Спроби	Успішність	Станд. відхилення	Оцінка навання	Призначена вага
47	39	73.72%	31.39%	16.67%	2%
46	39	74.36%	30.61%	16.67%	2%
48	39	91.03%	22.57%	16.67%	2%
49	39	80.77%	31.65%	16.67%	2%
50	39	65.38%	28.46%	16.67%	2%
1	39	51.28%	50.64%	25.00%	2%
10	39	51.28%	50.64%	25.00%	2%

Рис. 2.27. Аналіз структури тесту

- *тип тестового завдання* – для кожного типу тестового завдання існує своє графічне позначення, яке відображається в цьому стовпці;
- *переглянути, редагувати* – за допомогою цих вказівок викладач може переглядати або редагувати тестові завдання;
- *назва тестового завдання* – перехід на сторінку з аналізом обраного тестового завдання;
- *спроби* – параметр відображає кількість відповідей на тестове завдання;
- *успішність* (індекс складності (легкості)) тестового завдання (частка правильних відповідей). Цей показник можна записати так:

$$P_{\text{diff}} = \frac{X_{c(i)}}{X_{\text{макс}}}, \quad (2.7)$$

де  $X_{\text{макс}}$  – максимальна кількість балів за виконання  $i$ -го завдання;  $X_{c(i)}$  – середня кількість балів, отримана учасниками тестування за виконання  $i$ -го завдання.

За цим показником можна з'ясувати наскільки конкретне тестове завдання є легким/важким для тих, хто його виконує. Для завдань, які оцінюються в дихотомічній шкалі (0 або 1), ця характеристика показує, яка частка студентів, у %, відповіла правильно на це завдання (табл. 2.2). У випадку, коли  $P_{\text{diff}}=100\%$ , тобто всі студенти відповіли правильно, чи коли  $P_{\text{diff}}=0\%$  – відповіли неправильно, такі завдання слід вилучати з тесту.

Таблиця 2.2

#### Критерії індексу складності тестових завдань

$P_{\text{diff}}$	Висновок
$P_{\text{diff}} \leq 0,20$	Вгадування
$0,20 < P_{\text{diff}} \leq 0,36$	Надто складне
$0,36 < P_{\text{diff}} \leq 0,84$	Середньої складності
$P_{\text{diff}} > 0,84$	Надто легке

– *Стандартне відхилення* (розглядалося вище).  
 – *Оцінка навмання* – це середня оцінка студентів за тестове завдання, якої можна було б очікувати у випадку, коли студент відповідав навмання на конкретне тестове завдання. Випадкові оцінки (припущення) доступні лише для тестових завдань множинного вибору.

– *Призначена вага* – вага тестового завдання, виражена у відсотках загальної оцінки тесту.

– *Ефективна вага* – вага тестового завдання, розрахована для конкретного тестового завдання в тесті. Вона ґрунтується на результатах тестування.

– *Індекс дискримінації* (ІД). Використовуючи цей параметр групи учасників тестування за допомогою конкретного тестового завдання можна розподілити на сильних та слабких. За загальним результатом за тест, вибірка учасників поділяється на три групи: групу з найвищим рівнем підготовки (27 % найкращих результатів), з середнім рівнем підготовки та низьким рівнем підготовки (27 % найгірших результатів). Індекс дискримінації можна подати так:

$$D_j = \frac{(X_{max} - X_{min})}{n}, \quad (2.8)$$

де  $X_{max}$  – кількість балів із першої групи тестованих;  $X_{min}$  – кількість балів із третьої групи;  $n$  – загальна кількість балів за  $i$ -те завдання.

Цей показник змінюється від 0 (усі учасники тестування з сильної та зі слабкої груп відповіли правильно) до 1 (всі учасники з сильної групи відповіли правильно, а учасники із слабкої групи неправильно).

Від'ємний індекс свідчить про те, що учасники тестування із сильної групи відповідають на це тестове завдання гірше, ніж учасники із слабкої групи (табл. 2. 3). Тестові завдання, які дають такі результати, рекомендують вилучати із тесту, адже вони зменшують об'єктивність і точність усієї процедури тестування.

Таблиця 2. 3

<i><b><math>D_j</math></b></i>	<i><b>Висновок</b></i>
$D_j \geq 0,4$	Завдання функціонує задовільно
$0,3 \leq D_j \leq 0,39$	Потрібна невелика корекція завдання
$0,2 \leq D_j \leq 0,29$	Завдання потрібно переглянути
$D_j \leq 0,19$	Завдання потрібно вилучити із тесту або повністю переробити
$D_j = 0$ або $D_j < 0$	Завдання потрібно вилучити з тесту



– *Коефіцієнт дискримінації* (КД). Цей коефіцієнт визначається як коефіцієнт кореляції між множиною відповідей, отриманих при виконанні конкретного тестового завдання, з виконанням тесту в цілому. Значення цього коефіцієнту, можна обчислити за формулою:

$$КД_i = \frac{\sum(x \cdot y)}{(N \cdot S_x \cdot S_y)}, \quad (2.9)$$

де  $x$  – відхилення від середнього значення загальної кількості балів, що отримав випробовуваний, відповідаючи на дане тестове завдання;  $y$  – відхилення від середнього значення загальної кількості балів, що отримав випробовуваний в цілому за тест;  $\sum(x \cdot y)$  – сума добутків відхилень набраних балів для цього питання (тестового завдання) і тесту в цілому;  $N$  – кількість правильних відповідей на тестове завдання;  $S_x$  – середньоквадратичне відхилення множини балів, що отримав учасник тестування, відповідаючи на це тестове завдання;  $S_y$  – середньоквадратичне відхилення множини балів, що отримав учасник тестування за тест у цілому.

Перевагами коефіцієнта дискримінації відносно індексу дискримінації є те, що він використовує усі результати вибірки учасників, а не лише результати сильної і слабкої груп. Таким чином, цей параметр є ефективнішим для визначення якості тестових завдань.

Результати статистичного опрацювання тестових завдань можна переглянути і графічно – у вигляді гістограми (Рис. 2.28).

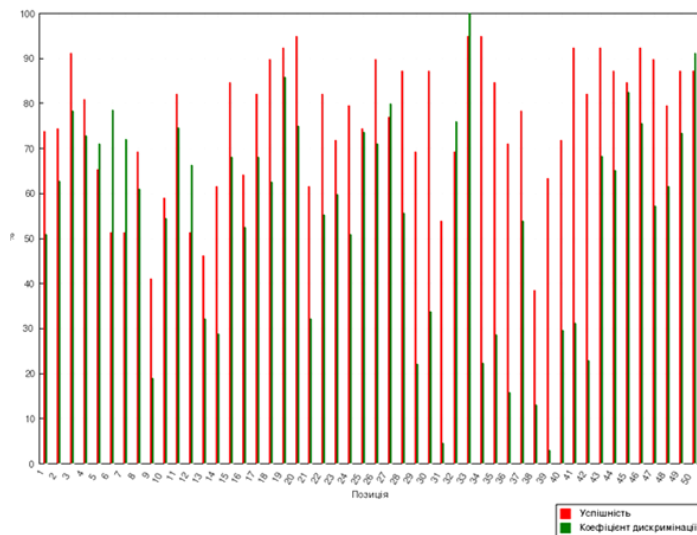


Рис. 2.28. Статистика тестових завдань (гістограма)

Викладач також може детальніше проаналізувати кожне тестове завдання, на яке відповіли учасники. Для цього необхідно "натиснути" на "Назву питання" в таблиці "Аналіз структури тесту", тоді відкриється сторінка з відомостями про обране тестове завдання та статистикою (Рис. 2.29).

Такий детальний аналіз забезпечує дотримання процедур визначення показників якості тесту в цілому й окремих його тестових завдань, зокрема можливе коригування тестових завдань з урахуванням цих показників.

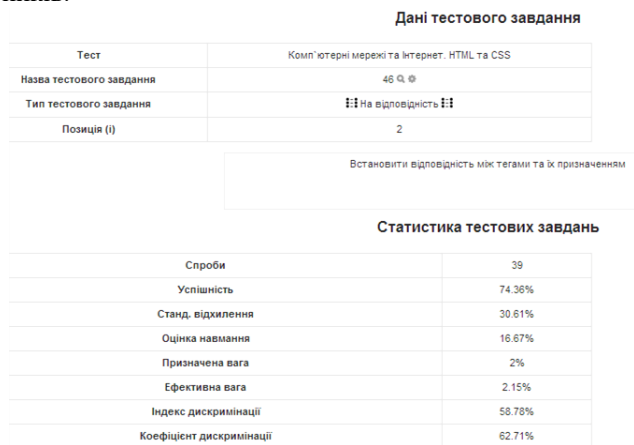


Рис. 2.29. Статистика тестового завдання

Аналіз вбудованої системи аналізу тестових завдань LCMS Moodle дає змогу стверджувати, що показники характеристик тесту є статистично обґрунтованими. Ця система використовує статистичні параметри, які розраховуються з використанням класичної (СТТ – Classical Test Theory) та сучасної теорії тестів (IRT – Item Response Theory). Використання вбудованої системи аналізу тестових завдань дає змогу опрацьовувати результати тестування з можливістю їх аналізу і оцінювання якості кожного тестового завдання та тесту в цілому з точки зору поставленої мети, їх коригування та удосконалення.

Однак система аналізу тестів та тестових завдань LCMS Moodle не дає змоги стандартними засобами розрахувати такі критерії, як валідність та надійність педагогічного тесту. Тому було запропоновано комп'ютерно-орієнтований технологічний компонент визначення зазначених критеріїв "Quiz statistics calculations". Він є веб-сайтом (Рис. 2.30), використання якого дає можливість в он-лайн режимі здійснити обчислення валідності, надійності як стійкості, надійності як внутрішньої узгодженості.

– *Валідність* тесту – це придатність результатів тестування для досягнення мети, заради якої проводилося тестування. Цей критерій

визначає, наскільки тест відображає те, що він повинен оцінювати (табл. 2. 4). Для оцінювання валідності тесту використовують кореляцію між показниками тесту та деяким зовнішнім критерієм (експертне оцінювання). За такого оцінювання важливо обрати значущий зовнішній критерій. Для педагогічних тестів за критерій приймають оцінки експертів або оцінки, які виставляються в результаті проходження комп'ютерного тестування. Валідність (від англ. valid) означає обґрунтований, дійсний, придатний, той, що має силу. Валідність тесту – це поняття, яке визначає, що вимірює тест і наскільки якісно.

Запишемо формулу за якою обраховується валідність:

$$V = \frac{\sum_{i=1}^n (Y_i - y_i) - Y \cdot y}{S_y \cdot S_Y} \cdot \frac{n}{n - 1}, \quad (2.10)$$

де  $n$  – кількість студентів;  $y_i$  – кількість правильних відповідей;  $Y_i$  – експертна оцінка;  $Y$  – середнє арифметичне експертних оцінок:

$$Y = \frac{\sum_{i=1}^n Y_i}{n}, \quad (2.11)$$

$y$  – середнє арифметичне тестових балів:

$$y = \frac{\sum_{i=1}^n y_i}{n}; \quad (2.12)$$

$S_Y$  – стандартне відхилення експертних оцінок:

$$S_Y = \sqrt{\frac{\sum_{i=1}^n (Y_i - Y)^2}{n - 1}}; \quad (2.13)$$

$S_y$  – стандартне відхилення правильних тестових балів:

$$S_y = \sqrt{\frac{\sum_{i=1}^n (y_i - y)^2}{n - 1}}. \quad (2.14)$$

Таблиця 2. 4

**Критерії валідності тестових завдань**

$V$	Висновок
Від 0,6 до 1	Висока валідність тесту
Від 0,3 до 0,6	Середня валідність тесту
Менше за 0,3	Низька валідність тесту

– *Надійність* результатів називається характеристика тесту, яка відображає точність тестових вимірювань, а також стійкість тестових результатів до дії випадкових чинників. Тест вважається надійним, якщо він забезпечує високу точність результатів вимірювань та якщо він дає близькі результати у разі повторного тестування за умови, що підготовка

студентів не змінилася за час до повторного виконання тесту. Таким чином, будемо вважати, що надійність тесту показує, на скільки точно тест вимірює знання студентів. Надійність характеризує точність тесту як вимірювального інструмента. *Надійність як стійкість результатів* вимірюється за допомогою повторного тестування на тій самій вибірці тестованих, зазвичай через два тижні після першого тестування. Для обчислення такої характеристики пропонується використовувати формулу Пірсона.

$$r = \frac{n \sum_{i=1}^n X_i Y_i - (\sum_{i=1}^n X_i) (\sum_{i=1}^n Y_i)}{\sqrt{n \sum_{i=1}^n (X_i)^2 - (\sum_{i=1}^n X_i)^2} \sqrt{n \sum_{i=1}^n (Y_i)^2 - (\sum_{i=1}^n Y_i)^2}}, \quad (2.15)$$

де  $X_i$  – тестовий бал  $i$ -го тестованого при проходженні тесту перший раз;  $Y_i$  – тестовий бал  $i$ -го тестованого при повторному проходженні тесту.

Для перевірки *надійності як внутрішньої узгодженості завдань тесту* використовують метод розщеплення тесту. Під час застосування такого методу матрицю розбивають на дві частини, що складаються з завдань із парними і непарними номерами. Цей коефіцієнт можна обчислити за формулою Рюлона:

$$r = 1 - \frac{S_d^2}{S_z^2}, \quad (2.16)$$

де  $S_d^2$  – дисперсія між тестовими результатами кожного тестованого за обома частинами тесту:

$$S_d^2 = \frac{\sum_{i=1}^n ((X_i - Y_i) - \frac{\sum_{i=1}^n (X_i - Y_i)}{n})^2}{n - 1}; \quad (2.17)$$

$S_z^2$  – дисперсія сумарних балів результату:

$$S_z^2 = \frac{\sum_{i=1}^n (\frac{\sum_{i=1}^n Z_i}{n} - Z_i)^2}{n - 1}; \quad (2.18)$$

де  $X_i$  – кількість правильних відповідей з парним номером;  $Y_i$  – кількість правильних відповідей з непарним номером;  $Z_i$  – кількість правильних відповідей.

Таблиця 2. 5

**Критерії надійності тестових завдань**

$r$	<b>Висновок</b>
Від 0,9 до 1	Дуже висока надійність тесту
Від 0,8 до 0,9	Висока надійність тесту
Від 0,7 до 0,8	Хороша надійність тесту
Менше ніж 0,7	Низька надійність тесту

**Аналіз критеріїв педагогічного тесту**

Головна      Обчислити критерії      Інтерпретація результатів      Контакти

➔ Для початку роботи ознайомтеся з інструкцією

**Оберть критерій**

Валідність      Надійність як стійкість      Надійність як внутрішня узгодженість

Хід виконання роботи:

1. На сторінці "Обчислити критерії" оберть "Валідність".
2. Введіть в поле "Кількість студентів" потрібне значення.
3. Введіть в поле "Кількість тестових завдань" потрібне значення.
4. Натисніть на кнопку "Сформувати таблицю".
5. Заповніть таблицю (у відповідних полях введіть: кількість правильних відповідей та експертну оцінку).
6. Натисніть кнопку "Обчислити", після чого відкриється діалогове вікно із значенням валідності.

Валідність - придатність тестових результатів для тієї мети, заради чого проводилося тестування, визначає, наскільки тест відображає те, що він повинен оцінювати. Для оцінювання валідності тесту використовують кореляцію між показниками тесту і деяким зовнішнім критерієм. При такому оцінюванні важливо обрати значимий зовнішній критерій. Для педагогічних тестів як критерій беруться оцінки експертів, або оцінка яка виставляється в результаті проходження компютерного тестування. Термін валідність походить від англійського "valid" і означає обгрунтований, дійсний, придатний, той, що має силу. Валідність тесту – це поняття, яке визначає, що вимірює тест і наскільки якісно це здійснюється.

**Формула за якою обраховується валідність**

$$V = \frac{\sum_{i=1}^n (Y_i * y_i) - \bar{Y} * \bar{y}}{S_Y * S_y} * \frac{n}{n-1}$$

$Y_i$	кількість правильних відповідей	заповнюється в таблиці
$\bar{Y}$	експертна оцінка	заповнюється в таблиці
$\bar{y}$	середнє арифметичне тестових балів	$= \frac{\sum_{i=1}^n Y_i}{n}$
$\bar{Y}$	середнє арифметичне експертних оцінок	$= \frac{\sum_{i=1}^n Y_i}{n}$

Рис. 2.30. Фрагмент веб-сторінки "Головна"

Сайт складається з таких сторінок:

- *Головна* – на цій сторінці містяться теоретичний матеріал та інструкції щодо виконання роботи з обчислення критеріїв.
- *Обчислити критерії* – розміщені поля та матриці, які потрібно заповнити для обчислення коефіцієнтів (Рис. 2.31).
- *Інтерпретація результатів* – наведені таблиці з еталонними значеннями критеріїв та рекомендації щодо конструювання тестів.
- *Контакти*.

На веб-сайті викладено рекомендації щодо конструювання тестів і тестових завдань з урахуванням особливостей алгоритму обчислення валідності та надійності:

1. Якщо тест розглядати як засіб контролю якості знань, то, насамперед, необхідно удосконалювати форму тестових завдань з метою підвищення якості педагогічного вимірювання, включаючи до тесту не тільки завдання з вибором однієї правильної відповіді, а й з вибором кількох правильних відповідей. Безумовно, тестові завдання з вибором

однієї правильної відповіді зручно використовувати при автоматизованому контролі знань. Однак за допомогою таких завдань не завжди можна якісно перевірити знання випробовуваних, оскільки існує велика ймовірність вгадування правильних відповідей. Такого недоліку позбавлені тестові завдання з вибором кількох правильних відповідей і завдання відкритої форми, а інтенсивний розвиток інформаційних технологій дає змогу включити їх в педагогічний тест і при автоматизованому контролі знань.

2. Якщо розглядати тест як засіб навчання, призначений для формування системи знань студентів, то в педагогічній практиці доцільно використовувати комплекси тематичних тестів із відповідної дисципліни. Ці комплекси повинні складатися не тільки з тестів для перевірки практичних умінь і навичок студентів, а й з тестів для перевірки їх теоретичних знань, які є ядром системи знань студентів.

3. Для підвищення валідності тесту варто дотримуватись таких рекомендацій:

- складати тестові завдання оптимальної складності для забезпечення нормального розподілу балів по тесту;
- проводити експертизу якості змісту тесту;
- правильно розраховувати оптимальний час виконання тесту;
- включати в тест завдання з високою дискримінаційною здатністю.

4. Для підвищення надійності тесту уникайте впливу на тест таких чинників:

- суб'єктивізму при оцінюванні результатів виконання завдань тесту, використовуючи для цього завдання закритого типу;
- вгадування, що істотно знижує надійність тесту, особливо у випадках, коли тестується група слабких студентів, які зазвичай вдаються до вгадування при виконанні найскладніших завдань тесту;
- невинного вибору вагових коефіцієнтів;
- невинної довжини тесту, оскільки його надійність зростає зі збільшенням довжини (зазвичай достатньо 30 тестових завдань);
- відсутності інструкції до тесту, оскільки будь-які неоднозначні відомості в інструкції зумовлюють зниження надійності тесту;
- джерел ненадійності, пов'язаних з тестованими.

5. Іноді під час обчислення коефіцієнта надійності виконується ділення на нуль. Це може статися, коли всі випробувані мають однакову кількість правильних і неправильних відповідей. Таке рідко буває на практиці – швидше за все стався витік даних про правильні відповіді. В цьому випадку слід провести тестування повторно.

6. При знаходженні надійності як стійкості, так само можливий випадок, коли у відповіді отримується невизначеність, тобто відбувається ділення нуля на нуль. Таке може статися, коли випробуваний дав однакову

кількість правильних і неправильних відповідей на першому і повторному тестуваннях. Це означає, що тест складений дуже вдало або, навпаки, дуже невдало. В цьому разі потрібно перевірити інші тестові характеристики та, виходячи з них, зробити висновки.

7. Під розрахунку валідності також можливий випадок, коли відбувається ділення на нуль. Це може відбутися в тому випадку, якщо всі випробувані мають однакову кількість правильних і неправильних відповідей або якщо всі експертні оцінки однакові. Таке рідко може статися на практиці – швидше за все стався витік даних про правильні відповіді, тому результат спотворений [17].

**Аналіз критеріїв педагогічного тесту**

Головна
Обчислити критерії
Інтерпретація результатів
Контакти

**Оберіть потрібні критерії для аналізу якості теста:**

- Обрати все
- Валідність
- Надійність
  - Надійність як стійкість
  - Надійність як внутрішня узгодженість
- 5 Кількість студентів
- 10 Кількість тестових завдань

Сформувати таблицю

---

Заповніть таблицю результатами тестування

#	1	2	3	4	5	6	7	8	9	10	$Y_i$	$Y_i$
1	1	1	1	1	1	1	1	1	1	1	10	100
2	1	1	1	1	1	1	1	1	1	1	10	100
3	1	1	1	1	1	1	1	1	1	1	10	100
4	1	1	1	1	1	1	1	1	1	1	10	100
5	1	1	1	1	1	1	1	1	1	1	10	100

Обчислити

Рис. 2.31. Фрагмент веб-сторінки "Обчислити критерії"

Застосовуючи в комплексі стандартний набір аналізу тестів LCMS Moodle та додатковий компонент "Quiz statistics calculations", можна аргументовано здійснювати коригування інструментаріїв педагогічної діагностики, аналізувати точність даних про стан фахової підготовки студентів, виробити заходи щодо ліквідації прогалин у навчальному процесі та на підставі даних діагностичних процедур обрати подальший

напрямок удосконалення їхньої фахової підготовки.

### **2.3. Педагогічна модель комплексу комп'ютерно-орієнтованих технологій освітніх вимірювань**

Модель – це штучно створений об'єкт у вигляді схеми, фізичних конструкцій, знакових форм або формул, який подібний до досліджуваного об'єкта чи явища, відображає і відтворює в більш загальному вигляді структуру, властивості, взаємозв'язки з елементами цього об'єкта [27].

Педагогічна модель – це мисленева система, яка імітує або відображає певні властивості, ознаки, характеристики об'єкта дослідження, а також принципи його внутрішньої організації та функціонування і подається у формі, притаманній певній соціокультурній практиці. В свою чергу під педагогічним моделюванням розуміють дослідження педагогічних об'єктів із використанням моделювання понятійних, процесуальних, структурно-змістових і концептуальних характеристик та окремих "сторін" навчально-виховного процесу в межах певного соціокультурного простору на загальноосвітньому, професійно орієнтованому або інших рівнях [48].

Педагогічне моделювання дає можливість розв'язати такі завдання: відображення властивостей оригіналу та дослідження цього характеристик із використанням моделі. Перше завдання вирішується під час навчально-виховного процесу, оскільки будь-яка теорія є моделлю реальних об'єктів, їх взаємодії або процесів. Друге завдання вирішується в ході вивчення реальних об'єктів і процесів або при їх конструюванні чи проектуванні.

Створення комплексу комп'ютерно-орієнтованих технологій освітніх вимірювань, а зокрема його педагогічної моделі націлене на удосконалення фахової підготовки студентів. Під удосконаленням у свою чергу ми розуміємо такі зміни фахової підготовки, що зумовлюють поліпшення, тобто формування системи фахових знань та їх якісного рівня в студентів.

Структуро твірним чинником педагогічної моделі є мета, досягнення якої забезпечується низкою компонентів. Один з компонентів розробленої моделі є підсистема педагогічної діагностики якості знань, набутих в процесі фахової підготовки (Рис. 2.32). Вона націлена на забезпечення організації процесу фахової підготовки та керування ним.





Рис. 2.32. Система організації та керування навчальним процесом

При побудові цієї підсистеми першочерговим було визначити мету та завдання педагогічної діагностики. Визначено такі етапи проведення діагностики:

1. *Накопичення даних.* Виконання діагностики потребує використання певної сукупності даних стосовно результатів і перспектив навчання. Частково вхідними даними для проведення діагностики є результати педагогічного контролю. Здобуття певних даних обумовлює вибір методів дослідження, що впливають на коректність діагностичного висновку.

2. *Діагностика поточного стану.* Визначається глибина кризового стану та виконується поглиблений аналіз процесу фахової підготовки. Мета проведення цього етапу – підтвердження висновку щодо якості системи знань студентів. Визначається перелік чинників та оцінюється вплив кожного з них на виникнення негативних тенденцій в процесі фахової підготовки.

3. *Прогнозування.* Завданнями цього етапу діагностичного дослідження є оцінювання вірогідності та можливих наслідків поглиблення негативних тенденцій в процесі фахової підготовки.

4. *Дослідження ефективності коригувальних заходів.* Аналіз усіх чинників впливу та оцінювання можливостей коригування і удосконалення процесу фахової підготовки.

5. *Узагальнення результатів діагностування якості системи знань.* Проведення педагогічної діагностики завершується формуванням загального висновку, в якому узагальнюються результати виконаної

роботи. Загальний звіт за результатами діагностики має в систематизованому вигляді містити всі отримані дані дослідження, а саме: про початкову підготовку студентів, мотиви, деякі психофізіологічні властивості, характер прояву негативних явищ та їхній вплив на навчальний процес; причини та чинники, які їх зумовлюють. Цей етап супроводжується формуванням остаточних рекомендацій щодо удосконалення навчального процесу.

Розглянемо детально розроблений другий компонент педагогічної моделі, а саме систему за якою здійснюється контроль на різних етапах навчально-виховного процесу в рамках педагогічної діагностики (Рис. 2.33). При цьому спочатку доцільно визначити об'єкт дослідження – результати навчання студентів (теоретичні знання, практичні вміння та навички).

Наведемо основні етапи проведення педагогічного контролю:

1. *Визначення цілей і завдань педагогічного контролю.* Головною вимогою цього етапу стало встановлення відповідності цілей контролю нормативним цілям навчання. Принципово важливим для забезпечення комплексної перевірки є визначення повноти цілей контролю. Результати досліджень свідчать, що мета контролю буде задовольняти вимогу повноти, якщо їх сукупність буде спрямована на виявлення у студентів цілісної системи знань і умінь на рівні поточного або підсумкового контролю.

2. *Визначення змісту контрольних завдань.* Відповідність змісту контрольних заходів цілям та змісту навчання.

3. *Вибір організаційних форм контролю, які відповідають цілям та змісту.* Вимогою цього етапу є гнучке та всебічне використання комп'ютерно-орієнтованих технологій.

4. *Визначення порядку інтерпретації студентам контрольних завдань.* Доцільним є гнучке комбінування традиційних форм контролю та комп'ютерно-орієнтованих. Вибір методів контролю зумовлюється специфікою змісту контролю і особливостями його організаційних форм.

5. *Встановлення критеріїв оцінювання результатів контрольних робіт.* Головною вимогою до розроблення критеріїв оцінювання є їхня відповідність цілям і змісту педагогічного контролю.

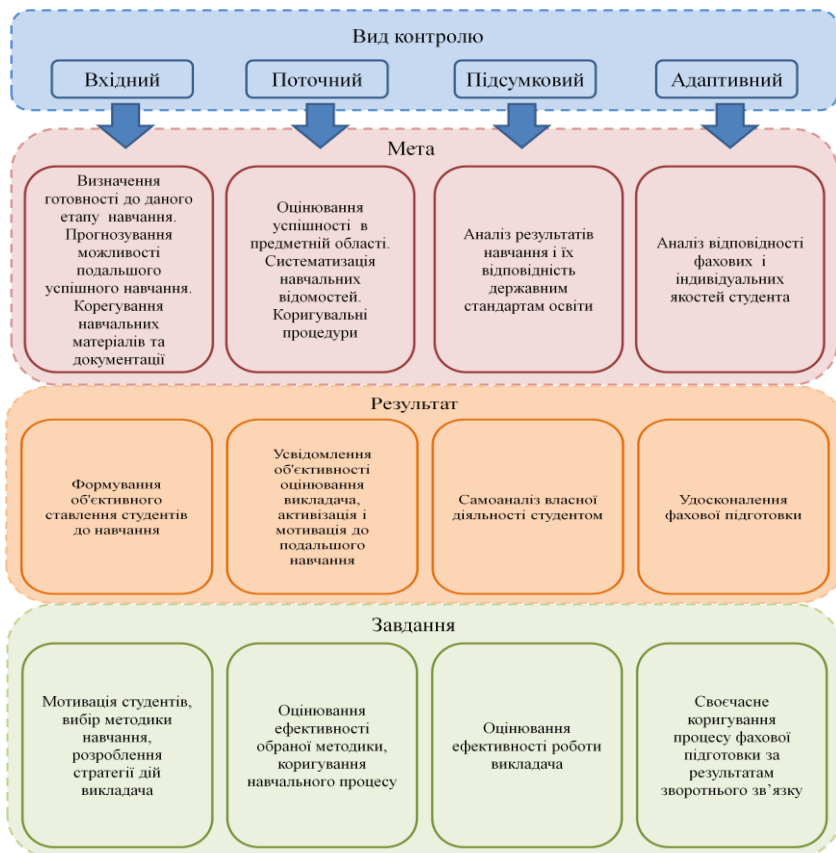


Рис. 2.33. Система педагогічного контролю якості знань

Педагогічний контроль забезпечує організаційний бік процесу підготовки майбутнього фахівця, що стосується управлінських рішень, тому для їх реалізації використовується педагогічна діагностика. У поняття «Педагогічна діагностика» вкладається глибший і ширший зміст, ніж у поняття «Педагогічний контроль»: контроль тільки констатує результати, не пояснюючи їх походження, а педагогічна діагностика включає в себе контроль, оцінювання, накопичення статистичних відомостей, аналіз, розглядає результати з урахуванням способів їх досягнення, виявляє тенденції та динаміку навчально-виховного процесу.

Запропонована педагогічна модель системи комп'ютерно-орієнтованих технологій освітніх вимірювань як засобу удосконалення фахової підготовки враховує основні принципи дидактики і складається з таких визначених блоків: цільовий, методологічний, змістовий,

організаційно-управлінський, оцінювальний та діагностично-результативний, з відкритим змістом навчання. В організаційно-управлінський блок входить підсистема організації та керування навчальним процесом, яка ґрунтується на процедурах педагогічної діагностики та контролю з використанням системи комп'ютерно-орієнтовних технологій освітніх вимірювань. Розкриємо зміст блоків поданих на рис 2.36.

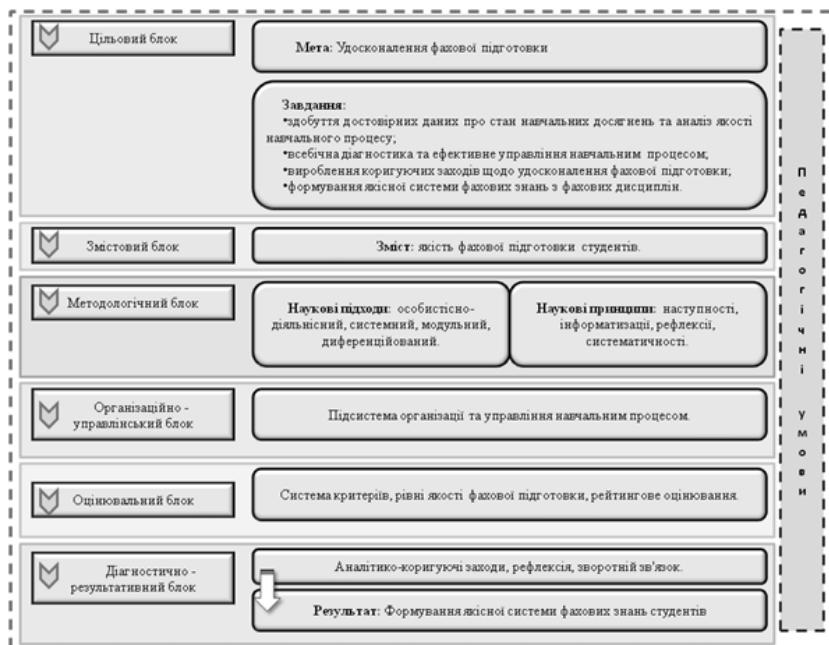


Рис. 2. 34. Педагогічна модель комплексу комп'ютерно-орієнтовних технологій освітніх вимірювань як засобу удосконалення фахової підготовки

*Цільовий блок* – відповідно до системного підходу мета виступає системо-утворювальним чинником, вихідним поняттям для прогнозування й планування діяльності [69]. Мета передбачає майбутній стан, котрий можливо змінити відносно теперішнього та бажано досягнути, тим самим мета є бажаною кінцевою точкою процесу. Мета – це об'єктивно зумовлена, логічна ієрархія завдань, які має реалізувати побудована педагогічна модель системи. Цільовий блок моделі включає в себе стратегічні і тактичні завдання, детерміновані в формах, методах та змісті діяльності суб'єктів навчально-виховного процесу.

*Змістовий блок* – як зміст визначено якість фахової підготовки

студентів. Під цим розуміється якість знань набутих у процесі фахової підготовки, тобто співвіднесення типів знань з елементами змісту освіти та рівнями засвоєння. При цьому враховуються такі характеристики якості знань, як повнота (кількість знань про об'єкт вивчення, визначених навчальною програмою), глибина (сукупність осмислених студентами зв'язків і відношень між знаннями), систематичність (осмислення складу сукупності знань в їхніх ієрархічних і послідовних зв'язках), оперативність (вміння користуватись знаннями в певній ситуації), гнучкість (вміння самостійно знаходити способи застосування знань у змінених умовах), конкретність (вміння диференціювати знання), узагальненість (вміння виразити конкретні знання в узагальненій формі).

У *методологічному блоці* визначено наукові підходи та принципи. Важливі умови при побудові моделі – дотримання наукових підходів, серед них:

– Особистісно-діяльнісний підхід – поєднує та розглядає в єдності особистісний та діяльнісний компоненти. Особистісний компонент передбачає, що в центрі навчання знаходиться студент, його мотиви, цілі, психологічні особливості. За такого наукового підходу організація навчально-виховного процесу орієнтується на самостійне розв'язування студентами конкретних навчальних завдань. Його суть полягає в забезпеченні саморозвитку особистості студента на основі виявлення його індивідуальних характеристик.

Особистісно-діяльнісний підхід у вищому навчальному закладі має такі сутнісні ознаки:

- суб'єкт-суб'єктне гуманне співробітництво всіх учасників навчально-виховного процесу у вищому навчальному закладі;
- діагностично-стимуляційний спосіб організації навчального пізнання студентів;
- діяльнісно-комунікативна активність студентів;
- проектування викладачем (а пізніше й студентами) індивідуальних досягнень студентів у всіх типах діяльності, сенситивних їх розвитку;
- врахування у змісті, методиках, системі оцінювання широкого діапазону особистісних потреб і можливостей студентів у здобутті якісної освіти.

При цьому використання особистісно-діялісного підходу має забезпечити особистісно-смісловий розвиток студентів, підтримку індивідуальності й неповторності кожного з них.

– Системний підхід ґрунтується на формуванні цілісного погляду на світ, що базується на ідеях цілісності складної організації досліджуваних об'єктів. Він передбачає існування чинної у суспільстві освітньої системи, що характеризується:

1. Визначеністю та впорядкованістю її структурних елементів.
2. Наявністю спільної дидактичної мети для всіх структурних елементів системи.

3. Існуванням багаторівневих зв'язків між структурними елементами системи, підпорядкованих спільній дидактичній меті, коли у системі має місце навчальна діяльність, педагогічна діяльність та навчально-педагогічна співпраця. Цей підхід полягає в дослідженні об'єкта навчально-виховного процесу як цілісної множини елементів у сукупності відношень і зв'язків між ними, тобто об'єкт розглядається як системи.

Системний підхід – це створення та функціонування педагогічної системи під час навчально-виховного процесу, що забезпечує високу ефективність підготовки фахівця, озброєного теоретичними сучасними та класичними знаннями, який володіє практичними вміннями, орієнтованого на своє подальше зростання, завдяки розвиненому перспективному мисленню та трудовій діяльності.

– Модульний підхід зумовлюється тенденцією вдосконалення навчально-виховної діяльності студентів, яка базується на самостійному пошуку знань, що оптимально організує навчальний процес за рахунок індивідуалізації і поєднання різних форм і типів навчально-пізнавальної діяльності.

– Диференційований підхід визначається на підставі розроблених у педагогічній теорії та практиці вимог, які ставляться до організації навчальної діяльності студентів, – це особливий підхід до різних груп студентів. Під диференціацією ми розуміємо спосіб організації навчального процесу, для якого характерне врахування індивідуально-типологічних особливостей учнів об'єднанням їх у гомогенні групи, в яких різняться елементи дидактичної системи мета, зміст, методи, форми та результати. Диференційований підхід – вивчення індивідуальних особливостей студентів та, певним чином, пристосування навчального процесу в академічних групах до сильних, середніх і слабких мікрогруп.

Поряд з науковими підходами було визначено принципи: наступності, інформатизації, рефлексії та систематичності.

Реалізація наступності передбачає таку організацію, за якої кожна дія є логічним продовженням попередньої. Під принципом наступності розуміється проходження низки етапів, протягом яких поступово ускладнюються й урізноманітнюються зміст та напрями розвитку тих утворень, які загалом складають цілісну систему характеристик особистості.

Принцип систематичності органічно пов'язаний з науковістю знань. Їх системність зумовлює цілісність уявлень та світогляду. Цей принцип потребує, щоб при засвоєнні фахових знань враховувалися як послідовність та доступність, так і зв'язок дисципліни з іншими

дисциплінами.

Принцип рефлексії передбачає організацію фахової підготовки на базі пізнання суб'єктом самого себе, свого внутрішнього світу, аналізу власних думок, усвідомлення зовнішніх обставин, мотивації, здатності до змін та власних дій з урахуванням певних обставин.

Принцип інформатизації – це постійне вдосконалення програмних і технічних засобів навчання з урахуванням досягнень педагогічних, технічних і психологічних наук, спрямованих на спрощення пошуку необхідних знань, їх засвоєння і практичного застосування. До педагогічних умов було віднесено:

- чітке визначення цілей, завдань та послідовності здійснення діагностичних процедур з використанням комп'ютерно-орієнтованих технологій освітніх вимірювань;
- розроблення необхідного операційного інструментарію;
- реалізація вимог особистісного діяльнісного, модульного, диференційованого та системного підходів;
- формування суб'єктної позиції студента у процесі становлення його як фахівця.

*Організаційно-управлінський блок* містить підсистему організації та керування навчальним процесом, яка ґрунтується на процедурах педагогічної діагностики та контролю з використанням комп'ютерно-орієнтованих технологій освітніх вимірювань.

*Оцінювальний блок* включає в себе систему критеріїв, рівні якості фахової підготовки, рейтингове оцінювання навчальних та наукових досягнень студентів і рейтингове оцінювання виробничої діяльності викладачів.

*Діагностично-результативний блок* складається з процедур, які ґрунтуються на проведенні аналітико-коригувальних заходів, передбачає рефлексію (філософський метод, за якого об'єктом пізнання може бути сам спосіб пізнання або знання, думка, вчинок, рефлексія орієнтована на критику та осмислення теоретичного знання) та зворотний зв'язок "викладач – студент – викладач" [54].

### РОЗДІЛ 3. АНАЛІЗ РЕЗУЛЬТАТІВ ТЕСТОВИХ МАТЕРІАЛІВ

#### 3.1. Первинний аналіз результатів тестування

Необхідність використання різноманітних методів опрацювання та інтерпретації зумовлена прагненням виявити істинні відмінності в рівні знань тестованих – порівняти результати між собою. Це прагнення викликане досить серйозною причиною, пов'язаною з тим, що “сирі” тестові бали не дають відомостей про реальний рівень знань. Інтерпретація результатів і можливість віднести їх до категорії поганих чи хороших залежать від низки чинників. У першу чергу від того, як розподілилися результати решти тестованих, від складності та від кількості завдань у тесті. Методи багатофакторного аналізу результатів тестування є засобом об'єктивної експертизи змістовних компонентів якості освіти та особистісного розвитку: знань, умінь, навичок, творчого досвіду навчальної діяльності, ціннісно-змістового ставлення до навчання, ключових компетностей та ін.

Адекватність інтерпретації результатів тестування досягається шляхом зіставлення результатів досліджуваної вибірки або окремого випробуваного з певними нормами виконання тесту – великою кількістю показників, що встановлюються емпірично в процесі стандартизації тесту. Співвідношення результату випробуваного з нормами виконання дає змогу встановити, чи відповідає даний результат середньому, на скільки він вище або нижче ніж середній результат виконання тесту.

Найбільш зручним для подання статистичних даних є метод графічного аналізу. Парні і множинні порівняння об'єктів спостереження одного і того самого або різних рівнів проводяться за середнім первинним або тестовими балами, за значенням медіани або відсотками виконання завдань. Можуть порівнюватися статистичні дані кількох об'єктів: класи в школі, школи в районі, райони в регіоні тощо.

Первинний аналіз результатів тестування передбачає, крім графічної інтерпретації, роботу з матрицею результатів тестування.

Розглянемо покроково усі етапи первинної обробки результатів тестування:

##### **КРОК 1. Формування матриці результатів тестування**

*Матриця результатів тестування* – таблиця, що містить результати відповідей кожного тестованого на кожне завдання тесту. Номер рядка  $j$  відповідає номеру тестованого, а номер стовпчика  $i$  номеру завдання (Рис. 3.1).

До матриці результатів тестування заносяться дані про успішність виконання тестованими завдань тесту. Результати відповідей тестованих на завдання тесту оцінюються в дихотомічній шкалі: за кожен правильну



відповідь тестованій отримує 1 бал, а за неправильну або відсутню відповідь – 0 балів (Рис. 3.2).

j \ i		Запитання																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Тестовані	1																		
	2																		
	3																		
	4																		
	5																		
	6																		
	7																		
	8																		
...																			

Рис. 3.1. Матриця результатів тестування

Розглянемо на прикладі виконання первинної обробки результатів тестування.

	A	B	C	D	E	F	G	H	I	J	K
1	№	Номер завдання									
2	тестованого	1	2	3	4	5	6	7	8	9	10
3	1	1	1	1	1	1	1	0	0	0	0
4	2	1	1	0	0	0	0	0	0	0	0
5	3	0	0	0	0	0	0	0	1	0	0
6	4	1	1	0	1	1	1	1	1	1	1
7	5	1	0	1	0	1	1	0	0	0	0
8	6	1	1	1	0	0	0	0	1	0	0
9	7	1	1	1	1	0	1	0	0	0	0
10	8	1	1	1	1	0	0	0	0	0	0
11	9	1	1	1	1	1	1	1	1	1	0
12	10	1	1	1	1	1	0	1	0	0	0
13	11	0	0	0	0	0	0	0	0	0	0
14	12	1	1	1	1	1	1	1	1	1	1

Рис. 3.2. Матриця результатів тестування (приклад)

### КРОК 2. Перетворення матриці результатів тестування

На другому кроці з матриці тестових результатів вилучаємо рядки і стовпці, які містять лише нулі або одиниці. У наведеному прикладі таких стовпців немає, а рядків лише два. Один із них – нульовий, відповідає відповідям одинадцятого тестованого, який не зміг правильно виконати жодного завдання тесту.

У цьому випадку висновок такий: даний тест непридатний для оцінювання знань тестованого. Для виявлення його рівня знань тест необхідно полегшити кількома завданнями, які, швидше за все, виконає правильно більшість інших студентів / учнів.

Також непридатним, але з іншої причини, є тест для оцінювання знань 12 тестованого, який виконав правильно усі без винятку завдання

тесту. Причина непридатності полягає в його надмірній легкості (простоті), яка не дає змоги виявити реальний рівень знань 12-го учня / студента (Рис. 3.3).

Таким чином, на даному етапі необхідно видалити із матриці даних 11 та 12 рядки.

	A	B	C	D	E	F	G	H	I	J	K
1	<b>№</b>	<b>Номер завдання</b>									
2	<b>тестованого</b>	1	2	3	4	5	6	7	8	9	10
3	1	1	1	1	1	1	1	0	0	0	0
4	2	1	1	0	0	0	0	0	0	0	0
5	3	0	0	0	0	0	0	0	1	0	0
6	4	1	1	0	1	1	1	1	1	1	1
7	5	1	0	1	0	1	1	0	0	0	0
8	6	1	1	1	0	0	0	0	1	0	0
9	7	1	1	1	1	0	1	0	0	0	0
10	8	1	1	1	1	0	0	0	0	0	0
11	9	1	1	1	1	1	1	1	1	1	0
12	10	1	1	1	1	1	0	1	0	0	0
13											

Рис. 3.3. Матриця результатів тестування після перетворення  
**КРОК 3. Підрахунок індивідуальних балів тестованих та кількість відповідей на кожне завдання тесту**

Індивідуальний бал тестованого отримуємо, додавши усі бали, одержані за виконання завдань тесту (Рис. 3.4).

Для зручності індивідуальні бали (X) розміщуємо в останньому стовпці матриці результатів.

Кількість правильних відповідей на завдання тесту (Y) також отримаємо обчисливши суму балів, розміщених у стовпцях.

	A	B	C	D	E	F	G	H	I	J	K	L
1	<b>№</b>	<b>Номер завдання</b>										<b>Індивідуальні бали (X)</b>
2	<b>тестованого</b>	1	2	3	4	5	6	7	8	9	10	
3	1	1	1	1	1	1	1	0	0	0	0	6
4	2	1	1	0	0	0	0	0	0	0	0	2
5	3	0	0	0	0	0	0	0	1	0	0	1
6	4	1	1	0	1	1	1	1	1	1	1	9
7	5	1	0	1	0	1	1	0	0	0	0	4
8	6	1	1	1	0	0	0	0	1	0	0	4
9	7	1	1	1	1	0	1	0	0	0	0	5
10	8	1	1	1	1	0	0	0	0	0	0	4
11	9	1	1	1	1	1	1	1	1	1	0	9
12	10	1	1	1	1	1	0	1	0	0	0	6
13	<b>Кількість правильних відповідей (Y)</b>	9	8	7	6	5	5	3	4	2	1	50

Рис. 3.4. Матриця результатів із кінцевими сумами балів  
**КРОК 4. Упорядкування матриці результатів тестування**

Значення індивідуальних балів необхідно відсортувати за зростанням. Для цього в MS Excel необхідно:

- ✓ виділити блок комірок, які містять номери тестованих, матрицю результатів та індивідуальні бали (діапазон комірок В3:Л12);
- ✓ в меню «Дані» оберемо команду «Сортування»;
- ✓ у цьому діалоговому вікні необхідно обрати стовпець для сортування (у нашому випадку – «Індивідуальні бали»). Матрицю результатів наведено на (Рис 3.5):

	A	B	C	D	E	F	G	H	I	J	K	L
1	<b>№</b>	<b>Номер завдання</b>										<b>Індивідуальні бали (X)</b>
2	<b>тестованого</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>	
3	3	0	0	0	0	0	0	0	1	0	0	1
4	2	1	1	0	0	0	0	0	0	0	0	2
5	5	1	0	1	0	1	1	0	0	0	0	4
6	6	1	1	1	0	0	0	0	1	0	0	4
7	8	1	1	1	1	0	0	0	0	0	0	4
8	7	1	1	1	1	0	1	0	0	0	0	5
9	1	1	1	1	1	1	1	0	0	0	0	6
10	10	1	1	1	1	1	0	1	0	0	0	6
11	4	1	1	0	1	1	1	1	1	1	1	9
12	9	1	1	1	1	1	1	1	1	1	0	9
13	<b>Кількість правильних відповідей (Y)</b>	9	8	7	6	5	5	3	4	2	1	50

Рис. 3.5. Упорядкована матриця результатів тестування

### **КРОК 5. Графічне представлення даних**

Емпіричні результати тестування можна подати у вигляді полігона частот, гістограми або графіка.

Для зображення розподілу даних у широкому діапазоні значень досліджуваних ознак використовуються гістограми. Вони забезпечують наочність розподілу первинних і тестових балів за 100-бальною або 5-бальною шкалою, дають змогу порівнювати навчальні досягнення з різних масивів тестованих.

У гістограмах кожен стовпчик відповідає числу, відсотку або частці учасників, які набрали відповідний бал або бал у заданому діапазоні. На гістограмі можуть бути наочно подані максимальні і мінімальні значення балів досліджуваної вибірки; найбільша висота стовпчиків відображає результати, які часто зустрічаються; площі окремих ділянок гістограми дають уявлення про число або частку учасників, що отримали бали в тому чи іншому діапазоні балів. Це може бути розподіл первинних балів за весь тест або за частину тесту (Рис. 3.6).

Іноді стовпчики гістограм виділяють кольором або візерунком, позначаючи різні рівні знань тестованих. Залежно від цілей аналізу може бути обраний різний крок балів при побудові гістограм. На гістограмі за площею стовпчиків можна візуально оцінити частку тестованих, які отримали як мінімальні бали (неатестованих) – низький рівень, так і

максимальні – високий рівень, а також частку тестованих із середнім рівнем.

Для графічного подання результатів тестування, впорядкуємо результати експерименту і підрахуємо частоту одержаних балів.

	А	В
1	№ тестованого	Індивідуальні бали (X)
2	3	1
3	2	2
4	5	4
5	6	4
6	8	4
7	7	5
8	1	6
9	10	6
10	4	9
11	9	9
12		

Рис. 3.6. Незгрупований ряд

Для підрахунку рейтингу (рангу) кожного тестованого за індивідуальними балами необхідно скористатися функцією РАНГ, яка визначає ранг числа зі списку (Рис. 3.7).

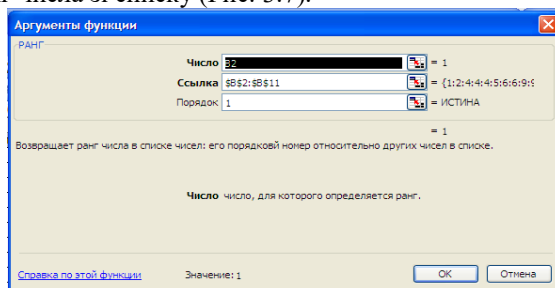


Рис. 3.7. Вікно функції РАНГ для комірки С2

**Ранг числа** – це величина цього числа відносно інших значень у списку (Рис. 3.8-3.9).

№ тестованого	Індивідуальні бали (X)	Ранг
3	1	1
2	2	2
5	4	3
6	4	3
8	4	3
7	5	6
1	6	7
10	6	7
4	9	9
9	9	9

Рис. 3.8. Ранжований ряд

Бал	Частота
1	1
2	1
4	3
5	1
6	2
9	2

Рис. 3.9. Таблиця частотного розподілу

За частотним розподілом можна побудувати гістограму (Рис. 3.10, 3.11).

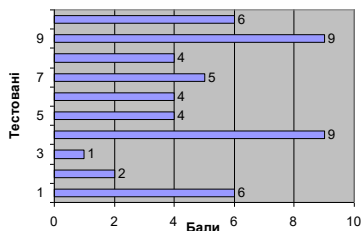
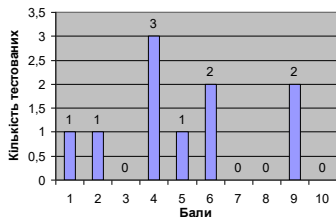


Рис. 3.10. Гістограма частот розподілу балів

Рис. 3.11. Гістограма розподілу індивідуальних балів

При конструюванні тестів необхідно пам'ятати про те, що крива розподілу індивідуальних балів, отриманих із репрезентативної вибірки, є результатом кривої розподілу складності завдань тесту.

#### **КРОК 6. Визначення вибірових характеристик результатів**

На цьому етапі необхідно визначити середнє значення, моду, медіану, дисперсію, стандартне відхилення вибірки.

**Середнім значенням** вибірки називається середнє арифметичне всіх її значень:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} \quad \text{або} \quad \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (3.1)$$

**Мода вибірки** – значення вибірки, яке повторюється найчастіше.

Визначення моди проводять різноманітними способами, і це залежить від того, у вигляді якого ряду – дискретного чи інтервального – подано досліджувану ознаку.

Моду знаходять шляхом звичайного перегляду стовпця частот. Якщо дані розташовані у вигляді дискретного ряду розподілу, то модою буде значення того варіанта, якого найбільша частота. Мода в статистиці застосовується тоді, коли слід охарактеризувати показник, який найчастіше зустрічається в сукупності.

В інтервальному варіаційному ряді розподілу легко відшуковується лише модальний інтервал, а сама мода визначається приблизно так:

$$M_o = X_{M_o} + i_{M_o} \cdot \frac{(f_{M_o} - f_{M_o-1})}{(f_{M_o} - f_{M_o-1}) + (f_{M_o} - f_{M_o+1})}, \quad (3.2)$$

де  $X_{M_o}$  – мінімальна межа модального інтервалу;  $i_{M_o}$  – величина модального інтервалу;  $f_{M_o}$  – частота модального інтервалу;  $f_{M_o-1}$  – частота інтервалу, що передує модальному;  $f_{M_o+1}$  – частота інтервалу, наступного за модальним.

Можуть бути розподіли, де всі варіанти зустрічаються однаково

часто, в цьому випадку моди немає або, можна сказати, що всі варіанти мають однакову модальність. В деяких випадках не один, а два варіанти можуть бути найбільшої частоти. Тоді буде дві моди, а розподіл буде бімодальних. Останній може свідчити про якісну неоднорідність сукупності щодо досліджуваної ознаки.

**Медіана вибірки** – це число, яке "поділяє" навпіл упорядковану сукупність усіх значень вибірки, тобто середня величина досліджуваної ознаки, яка міститься в середині ряду, розміщеного в порядку зростання або спадання ознаки.

У ранжованих рядах знаходження медіани зводиться до пошуку номера медіани. Її можна обчислювати за такою формулою:

$$M_e = x_{Me} + i_M \frac{\frac{\sum f}{2} - S_{Me-1}}{f_{Me}}, \quad (3.3)$$

де  $x_{Me}$  – нижня межа медіанного інтервалу;  $i_M$  – медіанний інтервал;  $S_{Me}$  – сума спостережень накопичених до початку медіанного інтервалу;  $f_{Me}$  – кількість спостережень у медіанному інтервалі.

Наведемо деякі властивості медіани:

1. Медіана не залежить від тих значень ознаки, які розміщені з обох боків від неї.

2. Аналітичні операції з медіаною досить обмежені, тому при об'єднанні двох розподілів з відомими медіанами неможливо завчасно передбачити величину медіани нового розподілу.

3. Медіані притаманна властивість мінімальності. Її зміст полягає в тому, що сума абсолютних відхилень значень  $x$ , від медіани представляє собою мінімальне значення порівняно з відхиленням  $X$  від будь-якої іншої величини.

Для визначення медіани графічним методом використовують накопичені частоти, за якими будується кумулятивна крива. Вершини ординат, які відповідають накопиченим частотам, з'єднують відрізками. Розділивши навпіл останню ординату, яка відповідає загальній сумі частот, і провівши до неї перпендикуляр перетину з кумулятивною кривою, знаходять ординату шуканого значення медіани.

**Дисперсія** – це числові характеристики розсіювання значень випадкової вибірки, що являє собою сукупність результатів незалежних спостережень.

**Дисперсією** називають середнє арифметичне квадратів відхилень індивідуальних значень ознаки. Залежно від вихідних даних дисперсія може обчислюватись за формулами середньої арифметичної простої або зваженої:

$$\sigma^2 = \frac{\sum(x-\bar{x})^2}{n}. \quad (3.4)$$

Дисперсія – це один з найпоширеніших узагальнених показників розміру варіації в сукупності. Її використовують не лише для оцінки

варіації, а й для визначення зв'язків між досліджуваними ознаками; розклад дисперсії на складові дає змогу оцінити вплив різних чинників, які зумовлюють варіацію досліджуваної ознаки.

**Стандартне відхилення** (англ. standard deviation) або *середнє квадратичне відхилення* позначають  $S$  або  $\sigma$  – показник розсіювання значень випадкової величини відносно її математичного сподівання.

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}, \quad (3.5)$$

де  $\bar{x}$  – середнє значення вибірки;  $n$  – кількість елементів вибірки.

Середнє квадратичне значення використовують для визначення показників варіації (коливання) досліджуваної ознаки – дисперсії та середнього квадратичного відхилення. Його обчислюють за квадратичними відхиленнями індивідуальних значень ознаки від їх середнього значення.

Середнє квадратичне відхилення, як і дисперсія, є широко використовуваним узагальнюючим показником варіації.

Значення середнього квадратичного відхилення вказує, на скільки в середньому відхиляються індивідуальні значення ознаки від їх середнього значення. Перевага цього показника порівняно із середнім лінійним відхиленням полягає у відсутності умовної похибки з суми відхилень без урахування їх знаків, бо відхилення підносять до квадрата. Крім зазначеного, перевагою цього показника порівняно з дисперсією є те, що середнє квадратичне відхилення виражають в тих же одиницях вимірювання, що і значення досліджуваної ознаки. Тому цей показник також називають стандартним відхиленням.

У статистичній практиці часто виникає необхідність порівняти варіації різних ознак. При таких порівняннях показники коливання ознак з різними одиницями вимірювання не можна використовувати.

Для здійснення таких порівнянь, а також при зіставленні ознаки у кількох сукупностях із різними середніми арифметичними застосовують відносний показник варіації – коефіцієнт варіації.

**Коефіцієнтом варіації** називають відсоткове відношення середнього квадратичного відхилення до середнього арифметичного значення ознаки:

$$V_{\sigma} = \frac{\sigma}{\bar{x}} \cdot 100\%. \quad (3.6)$$

Чим більший коефіцієнт варіації, тим менш однорідною є сукупність і тим менш типовою є середнє значення для даної сукупності. Встановлено, що сукупність кількісно однорідна, якщо коефіцієнт не перевищує 33%.

Дисперсія посідає особливе місце у статистичному аналізі і є

важливим елементом статистичних методів, зокрема у дисперсному аналізі. Приклад реалізації в Microsoft Excel (Рис 3.12).

	А	В
1	<b>№ тестованого</b>	<b>Індивідуальні бали (X)</b>
2	1	6
3	2	2
4	3	1
5	4	9
6	5	4
7	6	4
8	7	5
9	8	4
10	9	9
11	10	6
12	<i>Середнє значення</i>	5
13	<i>Мода</i>	4
14	<i>Медіана</i>	4.5
15	<i>Дисперсія</i>	6.889
16	<i>Стандартне відхилення</i>	2.625

Рис. 3.12. Вибіркові характеристики результатів тестування

12	<i>Середнє значення</i>	=СРЗНАЧ(В2:В11)
13	<i>Мода</i>	=МОДА(В2:В11)
14	<i>Медіана</i>	=МЕДИАНА(В2:В11)
15	<i>Дисперсія</i>	=ДИСП(В2:В11)
16	<i>Стандартне відхилення</i>	=СТАНДОТКЛОН(В2:В11)

Рис. 3.13. Формули для обчислення статистичних параметрів вибірки

Усі функції, які використовують для визначення вибірових характеристик, належать до категорії Статистичні.

### 3.1. Розподіл балів результатів тестування

#### 3.3.1. Перевірка нормальності розподілу балів результатів тестування

При застосуванні методів математичної статистики дуже важливо знати закон розподілу досліджуваної характеристики чи ознаки. Будь-яка статистична обробка даних починається із спроби визначити закон розподілу. Залежно від того, яким є закон розподілу, для аналізу даних використовують різні статистичні методи, а у випадку коли методи застосовують для розподілів, що відрізняються від гіпотетичного, отримують хибні результати та помилкові висновки.

Критерії перевірки гіпотез щодо законів розподілу називають критеріями згоди, які в свою чергу поділяють на загальні та спеціальні.

Загальні критерії використовують незалежно від розподілу, а спеціальні – у разі перевірки гіпотези щодо конкретного розподілу: нормального, рівномірного, експоненціального та ін.

Розрахунки емпіричного розподілу та його графічна візуалізація не



дають надійних підстав для висновку щодо закону розподілу ознаки у сукупності, з якої взято вибірку. Тим часом знання цього закону є необхідною умовою використання багатьох математичних методів.

Серед методів оцінювання законів розподілу ймовірностей випадкових величин близько 20 було спеціально розроблено для перевірки нормальності. Найпоширенішими вважаються:

- ✓ Критерії асиметрії й ексцесу;
- ✓ Хі-квадрат;
- ✓ Критерій Шапіро-Вілка.

Проте для перевірки нормальності варто використовувати критерій Шапіро-Вілка, який за рейтингом потужності посідає чільне місце.

**Крива нормального розподілу** – це симетрична крива рівномірного розподілу одиниць сукупності за значенням ознаки – балів (Рис 3.14).

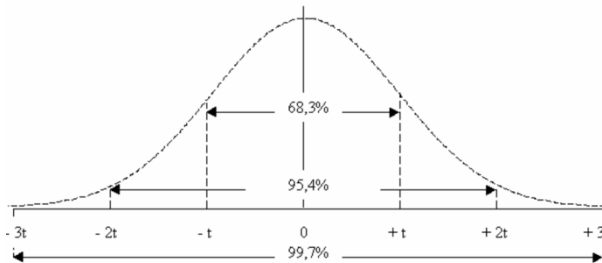


Рис. 3.14. Крива нормального розподілу

Нормальний розподіл – один з найважливіших розподілів вірогідності.

Термін "нормальний розподіл", що належить К. Пірсону (K. Pearson) (давній назви – закон Гаусса, розподіл Гаусса-Лапласа), застосовують як стосовно розподілів ймовірностей випадкових величин, так і спільних розподілів ймовірностей декількох випадкових величин (до розподілів скінченновимірних випадкових векторів), а також випадкових процесів.

У теорії ймовірностей доведено, що сума різних незалежних випадкових величин (незалежно від закону їх розподілу) виявляється випадковою величиною, розподіленою за нормальним законом (так звана центральна гранична теорема). Тому нормальний розподіл добре моделює найширше коло явищ, для яких відомо, що на них впливають декілька незалежних випадкових чинників.

Математичне очікування і дисперсія  $\epsilon$ , по суті, параметрами розподілу. Густина розподілу для чотирьох пар значень параметрів показана на Рис. 3.15. Вона задає ймовірність потрапляння випадкової величини  $x$  у малий інтервал від  $x$  до  $x + x\sigma$  Таким чином, наприклад, для кривої ( $\mu = 0, \sigma^2 = 0,2$ ) імовірність того, що випадкова величина  $x$  набуде

значення в околі нуля приблизно в чотири рази більша, ніж імовірність того, що вона набуде значення в околі  $x = 2$ . А значення випадкової величини більші ніж 5 і менші ніж  $-5$  і зовсім малоімовірні.

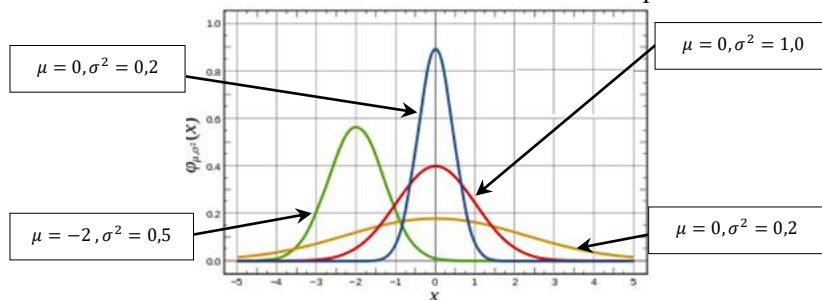


Рис. 3.15. Густина ймовірності нормальних розподілів

### 3.3.2. Критерій $\chi^2$ (хі-квадрат)

За допомогою нормального розподілу визначаються три розподіли, які на сьогодні часто використовують при статистичному опрацюванні даних. Це розподіли Пірсона ( $\chi^2$ ), Стьюдента та Фішера.

Зупинимо увагу на розподілі ( $\chi^2$ ). Вперше його дослідив у 1876 році астроном Ф. Хельмерт. З огляду на Гауссівську теорію помилок він досліджував суми квадратів  $n$  незалежних стандартно нормально розподілених випадкових величин. Пізніше К. Пірсон назвав цю функцію розподілу –  $\chi^2$ .

Завдяки тісному зв'язку з нормальним розподілом,  $\chi^2$ -розподіл відіграє важливу роль у теорії ймовірностей і математичній статистиці. Розподіл  $\chi^2$ , і багато інших розподілів, які визначаються за допомогою розподілу  $\chi^2$  (наприклад – розподіл Стьюдента), описують вибіркові розподіли різних функцій від нормально розподілених результатів спостережень. Їх використовують для побудови довірчих інтервалів і статистичних критеріїв.

Розподіл Пірсона  $\chi^2$  – розподіл випадкової величини  $X = X_1^2 + X_2^2 + \dots + X_n^2$ , де  $X_1, X_2, \dots, X_n$  – нормальні незалежні випадкові величини, причому математичне сподівання кожної з них дорівнює 0, а середнє квадратичне відхилення – 1.

При цьому кількість доданків, тобто  $n$ , називають числом ступенів свободи" розподілу  $\chi^2$ . Зі збільшенням числа ступенів свободи розподіл повільно наближається до нормального.

Щільність цього розподілу можна знайти так (Рис. 3.16):

$$f(x) = \begin{cases} 0, & \text{якщо } x \leq 0, \\ \frac{1}{2^{k/2} \cdot \Gamma(k/2)} \cdot \frac{x}{2} \cdot x^{k/2-1}, & \text{якщо } x > 0, \end{cases} \quad (3.7)$$

де  $\Gamma(x) = \int_0^\infty t^{x-1} \cdot e^{-t} dt$  – гамма-функція, а саме  $\Gamma(n+1) = n!$

Розподіл  $\chi^2$  залежить від одного параметра  $n$  – числа ступенів свободи. Функція розподілу  $\chi^2$  має вигляд:

$$f(x^2) = P(x^2 < x_0^2) = \begin{cases} 0, \\ \frac{1}{2^{n/2} \cdot \Gamma(\frac{n}{2})} \cdot \int_0^{x^2} (x^2)^{\frac{n}{2}-1} \cdot e^{-\frac{x^2}{2}} d(x^2). \end{cases} \quad (3.8)$$

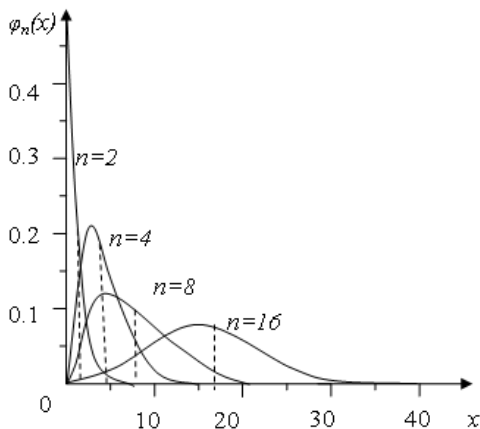


Рис. 3.16. Залежність щільності ймовірності  $\varphi(x)$  у розподілі  $\chi^2$  (хі-квадрат) за різної кількості ступенів свободи

Розподіл  $\chi^2$  використовують для оцінювання дисперсії (за допомогою довірчого інтервалу), при перевірці гіпотез згоди, однорідності, незалежності, насамперед для якісних (категоризованих) змінних, що набувають скінченного числа значень, і в багатьох інших завданнях статистичного аналізу даних.

Статистичні методи аналізу даних застосовують майже в усіх сферах діяльності людини. Їх використовують завжди, коли необхідно отримати і обґрунтувати будь-які судження про групу (об'єктів або суб'єктів) з деякою внутрішньою неоднорідністю.

Сучасний етап розвитку статистичних методів можна відраховувати з 1900 р., коли англієць К. Пірсон заснував журнал "Biometrika". Перша третина ХХ століття пройшла під знаком параметричної статистики. Вивчалися методи, що ґрунтуються на аналізі даних з параметричних

сімейств розподілів, описаних кривими сімейства Пірсона. Найпопулярнішим був нормальний розподіл. Для перевірки гіпотез використовувалися критерії Пірсона, Стьюдента, Фішера. Було запропоновано метод максимальної правдоподібності, дисперсійний аналіз, а також сформульовано основні ідеї планування експерименту.

Розподіл хі-квадрат є одним з найбільш використовуваних в статистиці для перевірки статистичних гіпотез. На основі розподілу хі-квадрат побудовано один з найпотужніших критеріїв згоди – критерій хі-квадрата Пірсона.

Критерієм згоди називають критерій перевірки гіпотези про передбачуваний закон невідомого розподілу.

Критерій  $\chi^2$  (хі-квадрат) використовується для перевірки гіпотези різних розподілів. У цьому й полягає його перевага.

Розрахункова формула критерію дорівнює:

$$\chi^2 = \sum_{i=1}^n \frac{(m - m')^2}{m}, \quad (3.9)$$

де  $m$  та  $m'$  – відповідно емпіричні та теоретичні частоти досліджуваного розподілу;  $n$  – число степенів свободи.

Для перевірки необхідно порівняти емпіричні (спостережувані) та теоретичні (обчислені за умови нормального розподілу) частоти.

Якщо емпіричні частоти збігаються з частотами, обчисленими або очікуваними  $S = E - T = 0$ , то критерій  $\chi^2$  також буде дорівнювати 0. Якщо  $S = E - T \neq 0$ , то це свідчить про невідповідність обчислених частот емпіричним частотам ряду. У таких випадках необхідно оцінити значущість критерію  $\chi^2$ , який теоретично може змінюватися від нуля до нескінченності. Це можна зробити шляхом порівняння фактично отриманої величини  $\chi_{\text{факт}}^2$  з його критичним значенням  $\chi_{\text{кр}}^2$ .

Нульова гіпотеза, тобто припущення, що розбіжність між емпіричними і теоретичними або очікуваними частотами носить випадковий характер, спростовується, якщо для прийнятого рівня значущості ( $\alpha$ ) і числа степенів свободи ( $n$ )

$$\chi_{\text{факт}}^2 \geq \chi_{\text{кр}}^2. \quad (3.10)$$

Розподіл можливих значень випадкової величини  $\chi^2$  неперервний і асиметричний. Він залежить від числа степенів свободи  $n$  і наближається до нормального розподілу в міру збільшення кількості спостережень. Тому застосування критерію  $\chi^2$  для оцінки дискретних розподілів пов'язане з деякими похибками, які позначаються на його значенні, особливо на нечисленних вибірках. Для отримання точніших оцінок вибірка, яку розкладають у варіаційний ряд, повинна мати не менше ніж 50 елементів. Правильне застосування критерію  $\chi^2$  потребує також, щоб у крайніх класах було б не менш ніж частоти варіант 5; якщо їх менше ніж 5, то їх об'єднують з частотами сусідніх класів, щоб сума була такою, що більше

ніж чи дорівнювала 5. Відповідно до об'єднання частот зменшується і число класів ( $N$ ). Число ступенів свободи встановлюється за вторинним числом класів із урахуванням числа обмежень свободи варіації.

Оскільки точність визначення критерію  $\chi^2$  у значній мірі залежить від точності розрахунку теоретичних частот  $T$ , для отримання різниці між емпіричними і обчисленими частотами слід використовувати не округлені теоретичні частоти.

### 3.3.3. Критерій Шапіро-Вілка

Критерій Шапіро-Вілка використовується для перевірки гіпотези  $H_0$ : «випадкова величина  $X$  має нормальний розподіл» і є одним з найефективніших критеріїв перевірки нормальності. Критерій, за допомогою яких перевіряють нормальність вибірки, є окремим випадком критеріїв згоди. Якщо вибірка нормальна, можна далі застосовувати потужні параметричні критерії, наприклад, критерій Фішера.

Критерій Шапіро-Вілка ґрунтується на оптимальній лінійній незміщеній оцінці дисперсії до її звичайної оцінки методом максимальної правдоподібності.

Статистика критерію  $W$  Шапіро-Вілка має вигляд:

$$W = \frac{1}{S^2} [\sum_{i=1}^n a_{n-i+1} (x_{n-i+1} - x_i)]^2, \quad (3.11)$$

$$S^2 = \sum_{j=1}^n (x_j - \bar{x})^2 = S_x^2 (n - 1), \quad \bar{x} = \sum_{j=1}^n x_j, \quad (3.12)$$

де  $n$  – об'єм вибірки.

Перевіримо за критерієм Шапіро - Вілка гіпотезу щодо відповідності нормальному законові розподілу емпіричних даних. Розглянемо покрокове виконання завдання:

#### **КРОК 1. Формулюємо гіпотези:**

$H_0$ : емпіричний розподіл не відрізняється від нормального;

$H_1$ : емпіричний розподіл відрізняється від нормального.

#### **КРОК 2. Будуємо таблицю емпіричних даних (Рис. 3.17).**

i	y <sub>i</sub>		
1	43	26	20
2	42	27	20
3	41	28	19
4	41	29	19
5	40	30	19
6	39	31	18
7	39	32	17
8	38	33	16
9	38	34	15
10	37	35	10
11	37	36	10
12	36	37	9
13	36	38	8
14	34	39	8
15	33	40	8
16	32	41	7
17	32	42	6
18	32	43	2
19	31	44	1
20	28	45	1
21	25	46	1
22	23	47	1
23	22	48	1
24	21	49	1
25	20	50	1

Рис. 3.17. Таблиця емпіричних даних результатів тестування

**КРОК 3. Визначаємо індекси  $i$  та  $(n-i+1)$** 

Розмістимо індекси  $i$  та  $(n-i+1)$ , причому  $i$  змінюється від 1 до  $m$ .

$$m = \frac{n}{2} = \frac{50}{2} = 25$$

**КРОК 4. Вводимо значення коефіцієнтів  $a_{n-i+1}$** 

У комірці стовпця  $a_{n-i+1}$  введемо значення з таблиці.

**КРОК 5. Визначимо значень у стовпцях  $x_i$  та  $x_{n-i+1}$** 

Стовпці  $x_i$  та  $x_{n-i+1}$  заповнюємо за допомогою функції ВПР, яка за індексом у лівому стовпчику таблиці повертає значення в тому самому рядку зазначеного стовпця таблиці.

**КРОК 6. Визначимо значення  $b_i$  та  $(\sum b_i)^2$** 

У комірках стовпця  $b_i$  розраховуємо значення.

$$b = a_{n-i+1} \times (x_{n-i+1} - x_i). \quad (3.13)$$

У комірці  $(\sum b_i)^2$  визначаємо квадрат суми параметрів стовпця  $b_i$ .

**КРОК 7. Визначаємо критерій  $W$  (Рис. 3.18).**

Значення критерію  $W$  обчислюємо за формулою

$$W = \frac{(\sum b_i)^2}{s^2}, \quad (3.14)$$

де  $S^2$  – квадрат значення дисперсії.

$(\sum b_i)^2$	9462,1
$S^2$	9680,3
$W$	0,9775
$W(0,1)$	0,9544

Рис. 3.18. Обчислення значення критерію Шапіро-Вілкі

**КРОК 8. Визначаємо критичні значення критерію  $W$** 

Критичне значення критерію  $W_{50}$  (0,05) отримуємо з табл. 3.1. Для педагогічний досліджень рівень значущості  $\alpha$  приймають таким, що дорівнює 0,05 (або 5%).

Таблиця 3.1

**Формули для визначення критичних значень критерію**

$\alpha$	$W_{\text{табл}}$
0,01	$\frac{-0,0148n^4 + 2,1875n^3 - 122,61n^2 + 3257,3n + 55585}{100000}$
0,05	$\frac{-0,0113n^4 + 1,656n^3 - 91,88n^2 + 2408,6n + 67608}{100000}$
0,1	$\frac{-0,0084n^4 + 1,2513n^3 - 70,724n^2 + 1890n + 73840}{100000}$

Отримані результати подано на Рис. 3.19.

$\alpha$	$W$
0,01	0,9286
0,05	0,9471
0,1	0,9544

Рис. 3.19. Показники критичних значень показника  $W$  для різних рівнів значущості

Якщо  $W \geq W_{\text{табл}}$ , то нульову гіпотезу приймають, тобто розподіл вважають нормальним. Отже, оскільки  $0,95 \geq 0,94$ , то розподіл вважається нормальним. Розходження між емпіричним і очікуваним теоретичним нормальним розподілами можна вважати статистично значущими на рівні 0,1.

**3.4. Основні характеристики тестових завдань****3.4.1. Дискримінативність тестових завдань**

**Дискримінативність завдань тесту** – це здатність окремих завдань методики розрізнити тестованих із високим або низьким тестовим результатом, а також суб'єктів з високою продуктивністю реальної (критеріальної) діяльності від тих, у котрих вона низька.

Будь-яку відповідь тестованого можна подати у дихотомічній шкалі:

- ✓ «правильно», «так» – один бал,
- ✓ «неправильно», «ні» – нуль балів.

Сума таких балів за всіма завданнями тесту являє собою первиний («сирий») бал. Міра збігу успішності розв'язування окремих завдань і всього тесту у вибірці тестованих є безпосереднім показником дискримінативності завдань тесту, який обчислюється у вигляді точково-бісеріального коефіцієнта кореляції і має назву *коефіцієнта дискримінації* (індексу дискримінації):

$$r_{pb} = \frac{\bar{x}_n - \bar{x}}{\sigma_x} \sqrt{\frac{N_n}{N - N_n}}, \quad (3.15)$$

де  $\bar{x}$  – середнє арифметичне значення усіх індивідуальних оцінок за тестом;  $\bar{x}_n$  – середнє арифметичне значення оцінок у тих тестованих, які правильно виконали дане завдання;  $\sigma_x$  – середньоквадратичне відхилення індивідуальних балів за тестом для вибірки;  $N_n$  – кількість тестованих, котрі правильно розв'язали завдання (або тих, чия відповідь на даний пункт опитувальника відповідає «ключу»);  $N$  – загальна кількість тестованих.

Наведене рівняння використовують для розрахунку коефіцієнта дискримінації лише у випадках, коли всі тестовані дали відповіді на всі завдання тесту. Коли кількість відповідей неповна, є сенс розраховувати коефіцієнт дискримінації за іншою формулою, а саме:

$$r_{pb} = \frac{\bar{x}_n - \bar{x}_d}{\sigma_{x_d}} \sqrt{\frac{N_n}{N_d - N_n}}, \quad (3.16)$$

де  $x_d$  – середнє арифметичне значення індивідуальних оцінок тих тестованих, які дали відповідь на завдання;  $\sigma_{x_d}$  – середнє квадратичне відхилення індивідуальних оцінок за тестом тих, хто дав відповідь на завдання;  $N_d$  – загальна кількість тестованих, які дали відповідь на завдання.

Коефіцієнт дискримінації може набувати значень від  $-1$  до  $+1$ . Близьке до  $1$  додатне значення  $r_{pb}$  означає, що конкретний пункт тесту надійно розподіляє тестованих, з високими та низькими балами. Близьке до  $1$  від'ємне значення  $r_{pb}$  свідчить про непридатність завдання, недостатній зв'язок часткового результату з кінцевим висновком.

Коефіцієнт дискримінації завдань тесту являє собою, по суті, показник критеріальної валідності завдання, оскільки його визначають відносно зовнішнього критерію – підсумкового результату або продуктивності критеріальної діяльності тестованого.

Індекс дискримінації можна знаходити із застосуванням *методу контрастних груп*. Необхідною умовою використання цього засобу є наявність близького до нормального розподілу оцінок за критерієм валідації. Частка тестованих у контрастних групах може змінюватися в широких межах залежно від обсягу вибірки: чим більшою за обсягом є вибірка, тим меншим відсотком тестованих можна обмежитись при



виокремленні груп із високими та низькими результатами.

*Нижня межа* виокремлення групи становить 10% для кожної з груп від загальної кількості членів вибірки, *верхня межа* – 33%.

Десяти відсоткові групи використовуються досить рідко, оскільки мала кількість випадків знижує статистичну надійність індексів дискримінації. Частіше з вибірки виокремлюють 25 - 27% тестованих.

Індекс дискримінації обчислюють як різницю між частками осіб, що успішно розв'язали завдання – окремо по високо- та низькопродуктивних групах. Такий індекс позначається ІД. Отже,

$$ІД = \frac{N_{n \max}}{N_{\max}} - \frac{N_{n \min}}{N_{\min}} \quad (3.17)$$

Оскільки  $N_{\max} = N_{\min} = 0,10: 0,33$ , то рівняння набуває вигляду:

$$ІД = \frac{N_{n \max}}{N_{\max}} - \frac{N_{n \min}}{N_{\min}} = 10,0 + 3,0 \frac{N_{n \max} - N_{n \min}}{N} \quad (3.18)$$

Врешті-решт, можна використати чотириклітковий коефіцієнт кореляції:

$$r_{phi} = \frac{f_g - f_d}{\sqrt{pq}}, \quad (3.19)$$

де  $f_g$  – частка осіб, що правильно розв'язали завдання, з загальної кількості тестованих, що належать до групи з максимальним результатом;  $f_d$  – частка осіб, що правильно розв'язали завдання, у групі мінімальних результатів;  $p$  – загальна частка ( $f_g + f_d$ ) осіб, що правильно розв'язали завдання;  $q$  – частка осіб, що неправильно розв'язали завдання ( $1 - p$ ).

Критичні значення коефіцієнта, що свідчать про діагностичну цінність пункту на рівні значущості  $p < 0,05$  залежно від кількості обстежених ( $n$ ), наведені в таблиці 3.2.

Таблиця 3.2

### ***Критичні значення коефіцієнта кореляції***

<b>Параметр</b>	<b>Значення</b>			
<b><i>n</i></b>	25	50	100	200
<b><i>r<sub>phi</sub></i></b>	0,39	0,28	0,20	0,14

Максимальна точність визначення  $r_{phi}$  досягається при обсязі контрастних груп близько 27% вибірки у кожній.

Аналізуючи дискримінативність особливу увагу слід приділяти статистичній вазі коефіцієнтів кореляції, та їх достовірності. У випадках, коли значення коефіцієнта дискримінативності наближається до нуля і рівень значущості невеликий, пункт тесту, що перевіряється, має бути переглянутий. Важливість визначення дискримінативності завдань тесту – одного з найголовніших показників, що характеризують діагностичну вагу опитувальників особистості, пояснюється орієнтованістю великої кількості

факторних опитувальників та анкет на дихотомічний розподіл досліджуваних за полярними характеристиками особистості.

Значення індексу дискримінативності для опитувальників підсилюється можливістю диференційованого підрахунку кінцевого результату. Аналіз дискримінативності завдань широко застосовують у випадках розробки та контролю об'єктивних тестів, тестів загальних та спеціальних здібностей тощо.

Дискримінативність – це здатність окремих завдань тесту і тесту в цілому диференціювати тестованих відносно максимального і мінімального результатів тесту. Її вимірюють показником дельта Фергюсона і набуває максимального значення при рівномірному (прямокутному) розподілі показників, у яких якнайповніше реалізовані всі можливі прояви вимірюваної властивості.

**Коефіцієнт Фергюсона** – це відношення між показником дискримінативності, отриманими для деякого тесту і максимальним значенням дискримінативності, яку може забезпечити такий тест.

За допомогою ретельного конструювання тесту можна забезпечити відповідний рівень дискримінативності, а це саме те, в чому тести виграють порівняно з іншими формами випробувань.

Показник роздільної здатності застосовується тільки для дихотомічного оцінювання завдань. Його обчислюють за формулою

$$D_j = (p_1)_j - (p_0)_j, \quad (3.20)$$

де  $(p_1)_j$  – частка учнів, які правильно виконали  $j$ -те завдання у підгрупі з 27% кращих учнів за результатами виконання тесту;  $(p_0)_j$  – частка учнів, які правильно виконали  $j$ -те завдання у підгрупі з 27% гірших учнів за результатами виконання тесту.

Значення індексу  $D_j$  належать інтервалу  $[-1;1]$ . Максимальне значення  $D_j = 1$  досягається у тому випадку, коли усі тестовані із підгрупи кращих, правильно виконали це завдання, а з підгрупи гірших це завдання не виконав ніхто. Тоді завдання має максимальний диференціувальний ефект. Завдання, у яких  $D_j = 0$  або  $D_j < 0$  потрібно з тесту вилучити.

В. К. Гайда та В. П. Захаров пропонують обчислювати коефіцієнт дискримінації, розраховуючи міру відповідності між успішністю розв'язання одного завдання і всього тесту. Цей показник і буде коефіцієнтом дискримінації, для його розрахунку використовують таку формулу:

$$D = \frac{x_n - x}{\sigma_x} \sqrt{\frac{n}{N_d - n}}, \quad (3.21)$$

де  $x$  – середнє арифметичне значення всіх індивідуальних оцінок по тесту;  $x_n$  – середнє арифметичне значення оцінок по тесту в тих тестованих, які правильно виконали завдання;  $\sigma_x$  – середнє квадратичне відхилення індивідуальних оцінок по тесту для вибірки;  $n$  – кількість тестованих, які

правильно виконали завдання;  $N_d$  – загальна кількість тестованих. Наведемо послідовність кроків знаходження коефіцієнта дискримінації.

**КРОК 1. Перетворення матриці результатів тестування**

У матриці результатів тестування потрібно знайти “сирі” бали, знайшовши суму балів для кожного тестованого. Відсортуємо результати тестування за зростанням.

**КРОК 2. Вибір 27% кращих та гірших результатів**

У даному випадку 27 % тасованих – 14 осіб  $\left(\frac{27 \times 51}{100}\right)$ .

У матриці результатів тестування виділяємо кольором 27% кращих і 27% гірших тестованих.

**КРОК 3. Визначення величин  $p_1$  та  $p_0$**

Засобами електронних таблиць Microsoft Excel визначаємо індекси дискримінативності кожного із завдань тесту.

Так, для визначення величини  $p_1$  необхідно ввести таку формулу:

$$= \frac{\text{кількість учнів, що виконали правильно завдання з першої групи}}{\text{загальна кількість учасників першої групи}}$$

i \ j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Сума балів	
26	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	18	
42	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	18
10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	17
17	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	1	17
23	1	1	1	1	1	1	1	1	0	1	1	0	1	1	0	1	1	1	1	1	1	17
24	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	0	17
29	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1	17
43	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	0	1	1	1	1	17
2	1	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	1	0	1	1	1	16
7	1	1	1	1	0	1	1	1	1	1	1	1	1	0	0	1	1	0	1	1	1	16
41	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	0	1	1	16
8	0	1	0	1	1	0	1	1	1	1	1	1	1	1	0	0	1	1	1	1	0	15
11	1	1	1	1	1	0	1	1	1	0	0	0	1	1	1	1	0	1	1	1	0	15
27	1	1	1	1	1	0	1	1	1	0	1	1	1	0	0	1	1	1	1	1	0	15
47	1	1	1	0	1	1	1	0	1	1	0	1	1	1	1	0	0	1	1	1	1	15
50	0	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	1	15
51	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	1	0	0	1	1	15
52	1	1	1	0	1	1	1	1	0	0	1	1	1	1	1	0	0	1	1	1	1	15
54	1	1	1	0	1	1	1	1	0	1	1	1	1	0	1	1	0	1	0	1	1	15
12	1	1	1	1	1	1	1	1	0	1	0	1	1	0	0	0	0	1	1	1	1	14
13	0	1	1	1	0	1	1	1	0	1	0	1	1	1	1	1	0	1	0	1	1	14
18	1	1	0	1	1	0	1	0	1	1	1	0	1	0	1	0	1	1	1	0	1	14
20	1	0	1	1	1	0	1	1	1	1	1	1	0	0	1	1	0	0	1	1	1	14
28	0	1	1	1	1	0	1	1	1	0	1	1	1	1	0	1	1	0	1	0	1	14
37	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	0	0	1	0	0	1	14
39	1	0	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	0	1	0	0	14
22	0	1	0	1	1	0	1	0	1	1	0	1	0	1	1	0	1	0	1	1	1	13
30	0	1	1	1	1	0	1	1	1	0	0	1	1	1	1	0	1	0	1	0	1	13
35	0	1	0	1	1	0	1	1	0	0	0	0	1	1	1	1	0	1	1	1	1	13
45	1	1	1	1	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	1	1	13
49	1	0	0	0	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	0	0	13
53	1	1	1	0	1	1	0	0	1	1	1	1	0	1	1	0	1	1	0	0	1	13
1	0	1	1	0	1	1	0	1	0	1	1	1	0	1	0	0	0	1	0	1	1	12
5	0	0	0	1	1	0	0	1	0	1	1	1	1	1	0	0	1	1	1	1	1	12
14	0	0	0	1	1	1	1	1	0	0	1	1	1	1	1	0	1	0	1	0	1	12
31	1	1	1	1	1	1	0	1	0	0	1	1	0	1	0	1	0	0	1	0	1	12
40	1	0	1	1	1	1	1	0	1	0	0	1	1	0	0	1	1	1	1	0	0	12
44	1	1	0	1	0	0	0	1	0	1	0	1	1	1	1	0	0	1	1	1	1	12
48	1	0	1	0	0	1	1	1	0	1	1	0	1	1	1	0	0	1	1	0	1	12
6	0	1	0	0	0	1	0	1	0	1	1	1	1	1	0	1	1	0	1	0	1	11
16	1	0	0	0	1	0	0	0	1	1	0	1	0	1	1	1	1	0	1	1	1	11
21	0	1	0	0	0	1	0	1	1	1	1	0	0	0	1	1	1	1	1	1	1	11
25	0	0	1	1	0	0	1	1	0	1	1	1	1	0	0	1	1	1	1	0	0	11
33	0	1	1	0	1	0	1	0	0	0	0	0	1	0	1	1	1	1	1	1	1	11
34	0	1	1	0	1	0	1	0	0	0	0	0	1	0	0	1	1	1	0	1	1	11
19	0	0	0	1	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1	1	1	10
32	0	1	0	0	1	0	1	0	0	0	0	0	1	1	1	1	1	0	1	1	1	10
3	0	0	0	0	1	0	0	1	0	1	1	1	0	1	0	0	0	1	1	1	1	9
36	0	0	1	0	0	0	1	0	1	0	1	1	1	1	1	0	0	1	0	0	0	9
38	0	1	0	1	0	1	1	0	1	0	1	0	0	0	0	1	1	0	0	1	0	9
46	0	0	1	0	0	1	1	1	0	0	1	1	1	0	0	1	0	0	0	0	1	9

Рис. 3.20. Матриця результатів тестування (перетворена)  
У середовищі Microsoft Excel ця формула матиме вигляд

$$= \frac{\text{СЧЕТЕСЛИ}(B2:B15; 1)}{\text{СЧЕТ}(B2:B15)}$$

Аналогічно визначаємо величину  $p_0$ :

=  $\frac{\text{кількість учнів, що виконали правильно завдання з третьої групи}}{\text{загальна кількість учасників третьої групи}}$

**КРОК 4. Визначаємо індекс дискримінативності** (Рис. 3.21).

Індекс дискримінативності знаходимо за формулою

$$D_j = (p_1)_j - (p_0)_j. \quad (3.21)$$

$(p_1)_j$	0,93	1	0,93	1	0,86	0,86	0,93	1	0,79	0,71	0,93	0,86	0,86	0,64	0,5	0,86	0,79	0,57	0,86	0,64
$(p_0)_j$	0,79	0,5	0,57	0,71	0,64	0,57	0,5	0,36	0,64	0,43	0,36	0,43	0,36	0,57	0,5	0,29	0,29	0,43	0,29	0,36
$D_j$	0,14	0,5	0,36	0,29	0,21	0,29	0,43	0,64	0,14	0,29	0,57	0,43	0,5	0,07	0	0,57	0,5	0,14	0,57	0,29

Рис. 3.21. Визначення індексу дискримінативності

**КРОК 5. Сортування результату дискримінативності**

Виконавши сортування даних таблиці за полем  $D_j$  отримаємо результати, отримані на рис. 3.22.

	15	14	9	18	1	5	4	6	10	20	3	12	7	13	2	17	11	16	19	8
$(p_1)_j$	0,5	0,64	0,79	0,57	0,93	0,86	1	0,86	0,71	0,64	0,93	0,86	0,93	0,86	1	0,79	0,93	0,86	0,86	1
$(p_0)_j$	0,5	0,57	0,64	0,43	0,79	0,64	0,71	0,57	0,43	0,36	0,57	0,43	0,5	0,36	0,5	0,29	0,36	0,29	0,29	0,36
$D_j$	0	0,07	0,14	0,14	0,14	0,21	0,29	0,29	0,29	0,29	0,36	0,43	0,43	0,5	0,5	0,5	0,57	0,57	0,57	0,64

Рис. 3.22. Відсортована таблиця для визначення індексу дискримінації

Проаналізувавши отримані результати, робимо такі висновки:

- завдання 15, 14, 9, 18, 1 потрібно вилучити з тесту,
- завдання 5, 4, 6, 10, 20 потрібно переглянути,
- завдання 3 потребує незначної корекції.

### 3.4.2. Складність тестових завдань

Складність завдання  $U$  є традиційним показником під час аналізу завдання.

Складність завдань тесту – це характеристика, що відображає статистичний рівень її розв'язності в даній вибірці стандартизації. Зазвичай її оцінюють при зіставленні кількості тестованих, що правильно виконали завдання, до загальної кількості тестованих. Цей показник змінюється в межах від 0 до 1. Його значення зростає зі збільшенням складності завдання. При розробленні критеріально-орієнтованих тестів показник складності використовують з метою конструювання тестів з бажаним рівнем складності. Для підвищення складності тесту до його складу включають більше завдань високого рівня складності, а для зменшення складності – легших завдань.

Прийнято виділяти суб'єктивну і статистичну (об'єктивну) складності. *Суб'єктивна складність завдання* пов'язана з індивідуальним

бар'єром, величина якого визначається як обставинами (обмеженням часу, доступністю інструкції і т.д.), так і рівнем сформованості необхідних для вирішення знань, умінь і навичок, психічним станом випробуваного і низкою інших чинників. Вплив комплексу цих чинників на результат виконання тесту знижує надійність і достовірність даних, тому важливо вживати заходи для вирівнювання їх впливу за допомогою спрямованого комплектування вибірки та стандартної процедури проведення тестування.

Показником складності завдань як *статистичної складності* є частка осіб вибірки, які розв'язали і не розв'язали завдання. Наприклад, якщо тільки 15% тестованих розв'язали завдання, його можна оцінити як важке для даної групи, якщо 85% – як легке. Складність завдань тесту є відносною характеристикою, оскільки залежить від особливостей вибірки (вікові, професійні, соціально-культурні відмінності).

Добір завдань за показником складності важливий для успішного застосування діагностичної методики. Так, при виборі занадто важких завдань валідність і надійність тесту різко знизяться (погіршення статистичних характеристик оцінювання за невеликої кількості правильних розв'язань завдань, імовірність випадкової оцінки, вплив неправильних розв'язків на подальшу роботу випробовуваних). Занадто прості завдання призведуть до одноманітності результатів (більшість опитуваних справляється з усім набором завдань тесту), їх незначної варіативності, що ускладнює оцінювання, і як наслідок, викликає негативне ставлення до тесту.

Основне завдання аналізу складності завдань тесту зводиться до вибору оптимальних за складністю завдань тесту, а значить, і впорядкування окремих завдань в ньому. Зазвичай, якщо завдання розв'язує більшість учнів, його (як легке) розміщують на початку тесту, якщо завдання розв'язує незначний відсоток випробуваних, то його (як важке) розміщують в кінці тесту. Одне - два найлегші завдання ставлять перед основними завданнями тесту і використовують як приклад. Розташування завдань за зростанням ступеня складності, вираженого часткою або відсотком осіб даної вибірки, які розв'язали відповідне завдання, дає змогу приблизно визначити поріг складності для індивідуального випробуваного, що характеризує його місце в групі тестованих.

Основним показником складності завдань тесту є індекс складності  $U_i$ , який обчислюють за формулою:

$$U_i = \left(1 - \frac{N_p}{N}\right) 100\%, \quad (3.22)$$

де  $U_i$  – індекс складності завдання у відсотках;  $N_p$  – кількість учнів, які правильно розв'язали завдання;  $N$  – загальна кількість тестованих.

Практика педагогічного тестування свідчить, що максимум інформації про індивідуальний рівень знань, навичок та умінь тестованих

забезпечують завдання середньої складності (50%).

### 3.5. Валідність результатів тестування

Вибір того чи іншого методу вимірювання рівня знань ґрунтується на критеріях, які визначають якість засобу вимірювання. Найважливішими критеріями якості є валідність, об'єктивність, надійність і точність. Розглянемо їх визначення та з'ясуємо зміст кожного з них.

Фундаментальним є критерій валідності. Валідність (від англ. valid) означає обґрунтований, дійсний, придатний, той, що має силу. Вдале визначення валідності педагогічного тесту надала американський психолог Анна Анастасі, (1908 – 2001).

Поняття валідність використовується для загальної характеристики тесту в аспекті відповідності одержаних результатів та вимогам оцінювання.

Невалідний тест – такий, у якому одержані результати не вимірюють того, що було визначено метою тестування. Це може бути наслідком дії кількох чинників. Наприклад, невідповідності інструменту вимірювання (тесту) меті оцінювання. Наведемо декілька прикладів. Тест побудовано з тестових завдань на розуміння матеріалу, а навчання було спрямовано лише на знання базового матеріалу. Метою оцінювання є рівень виконання певних навичок, а тест базується на вимірюванні знань про виконання цих навичок. Завдання, які входять до тесту, не є стандартизованими, отже, невалідність результатів може бути наслідком неякісних тестових завдань.

У стандартних вимогах до педагогічних тестів їх валідність має визначатися комплексом відомостей про характеристики тесту, процедури тестування та методики оцінювання.

Багато дослідників пропонують підхід до аналізу якості тестів і тестових завдань на підставі визначення їх головних характеристик і параметрів, коли загальне поняття валідності щодо процесу вимірювання й оцінювання рівня знань диференціюється за функціональною ознакою:

- валідність методу (валідність змісту, відповідності, прогнозу);
- валідність тесту (валідність тестових завдань, процедури тестування, процедури оцінювання).

Валідність – це комплексна характеристика, що визначається як параметрами засобу та процедури вимірювання, так і властивостями досліджуваної ознаки.

Валідність методу – це відповідність того, що вимірюється цим методом, тому, що він має вимірювати. Отже, критерій валідності встановлює сферу дійсності, для якої метод дає статистично достовірні результати. Валідність методу при вимірюванні успішності навчання можна класифікувати за такими критеріями:

- валідність змісту (змістова);
- валідність відповідності.

Валідність змісту (змістова) – характеристика тесту, що відображає ступінь упевненості, завдання тесту досить повно охоплює зміст певної сфери знань, тест точно визначає істотні для цієї діяльності навички, але водночас не перевіряє другорядні в цьому разі знання й навички, залежить від глибини попередньо проведеного авторами тесту аналізу навчального плану й програми, змісту підручників, а також від точності й повноти формулювання мети навчання.

Валідність відповідності – це відповідність результатів вимірювання та оцінювання, здобутих різними методами.

Валідність змісту і відповідності можна кількісно визначено через коефіцієнт валідності. Але необхідно зауважити, що розраховується коефіцієнт валідності опосередковано – як коефіцієнт кореляції між результатами тестування та результатами інших вимірювань, здійснених на тій самій групі з того самого предмета іншим методом. Результат тестування можна вважати валідним, якщо коефіцієнт кореляції більший ніж 0,6.

І.Є. Булах і М.Р. Мруга запропонували таке визначення коефіцієнта валідності: коефіцієнт валідності дорівнює коефіцієнту кореляції між результатами, одержаними різними методами за однакових умов, і свідчить, наскільки збігаються результати вимірів.

Валідність прогнозу (прогностична) – характеристика тесту, яка відображає ступінь упевненості, що отримані за тест оцінки добре прогнозують майбутні досягнення тестованого (наприклад, у США результати за складання тесту SAT передбачають певний рівень успішності студента у ВНЗ). Для виявлення прогностичної валідності результати одного тесту порівнюють із результатами іншого тесту, який проводять, наприклад, через півроку або рік навчання. Учитель / викладач може використати індивідуальні бали студента як такі, що передбачають його успіх у майбутньому навчанні чи професійній діяльності тощо. Прогностичну валідність оцінюють за допомогою коефіцієнта кореляції між результатами тестування й певним зовнішнім критерієм (зовнішньою змінною), що характеризуватиме вимірювану властивість у майбутньому. Основна складність визначення такої валідності – вибір вагомого зовнішнього критерію (зовнішньої змінної) в майбутньому.

Надійність методу вимірювання – це міра стійкості результатів, що впливає на точність, з якою можна виміряти ту чи іншу конкретну ознаку. Перевірка надійності методу стосується насамперед відновлення результатів у разі повторних вимірювань.

Під надійністю контрольного завдання розуміють ступінь точності, з якою можна визначити ту чи іншу ознаку. Тобто визначено, наскільки



можна довіряти результатам даного тесту. Про надійність тестів іноді роблять висновки за такою ознакою: якщо в усіх випадках перевірки тесту чи його варіантів виявиться, що учні в розподілі за показниками успішності займуть ті самі місця, то такий тест можна вважати надійним.

Надійність тесту залежить від кількості тестових завдань. Тому для достатньої надійності підсумкового контролю тест має містити не менше ніж 40 завдань.

П. Клайн вважає, що у психометрії термін “надійність” має два значення. Тест його називають надійним, якщо він є внутрішньо узгодженим. Також називається надійним, якщо він дає одні й ті самі показники для кожного тестованого (за умови, що тестований не змінився) за повторного тестування. Надійність за повторного тестування через певний час називають ретестовою надійністю.

В. Аванесов вважає, що в міру становлення нової теорії тестів починає відчуватись необхідність розширення кількості критеріїв хоча б до трьох. В якості третього педагог пропонує критерій ефективності.

Ефективність – порівняльний критерій, який дає змогу порівняти тести. Ефективним можна назвати тест, який краще, ніж інші тести, вимірює знання учнів потрібного рівня підготовки з меншою кількістю завдань, якісніше, швидше, дешевше, і все це, – якщо можливо, одночасно.

Якщо порівняти поняття "ефективність" із поняттями "надійність" і "валідність", то найістотніша відмінність нового поняття від двох традиційних полягає в переході від середнього показника до диференційованого. Достатньо пригадати, що надійність стосується тесту, який складається із зафіксованої кількості завдань, що пред'являються всім випробовуваним; тільки тоді можна знайти коефіцієнт надійності тесту як середню міру точності вимірювання. Аналогічно це стосується й валідності тесту. Ефективний же тест, навпаки, припускає відхід від фіксованої для всіх випробовуваних кількості завдань.

Психологи Л. Бурлачук і С. Морозов до ефективних тестів висувують ще одну істотну вимогу – складність завдання тесту.

Складність завдання тесту – характеристика завдання тесту, що відображає статистичний рівень розв'язування в даній вибірці проблеми стандартизації. Показником складності тестового завдання є частка вибірки випробовуваних, які виконали чи не виконали дане завдання. Наприклад, якщо лише 20 % випробовуваних виконали завдання, його можна вважати складним для даної вибірки, якщо 80 % – легким.

Добір завдань за показниками складності важливий для успішного використання тестів. Так, при доборі надто складних тестових завдань валідність і надійність тесту різко зменшуються. Надто прості тестові завдання призведуть до одноманітності тесту та його неефективності.

Ще однією особливістю ефективних тестів, на думку П. Клайна, є

дискримінативність. Досягнення задовільного розподілу показників є однією з цілей автора тестів. Нема необхідності підкреслювати те, що є очевидним: яка цінність тесту, за яким усі випробувані показали однаковий результат? За допомогою ретельного конструювання тесту можна забезпечити відповідний рівень дискримінативності, а це саме те, у чому тести виграють порівняно з іншими формами випробувань.

Як бачимо, вимоги, які ставлять тестологи до ефективних тестів, між собою дещо розрізняються. Мета вимірювань завжди полягає в отриманні відомостей про ознаки об'єктів чи подій. Вимірюється не сам об'єкт, а тільки властивості або відмінні ознаки об'єкта.

У педагогічній діагностиці намагаються визначити якість результатів вимірювання (тут – вимірювання успішності). Варто зазначити, що всі вище наведені критерії спрямовані на єдину загальну вимогу: тест повинен давати як найточнішу відомості про досліджуване явище. Вони лише конкретизують цю вимогу й характеризують її з різних боків.

Наведемо чинники, що знижують валідність результатів оцінювання:

- ✓ Завдання не відповідають меті тестування.
- ✓ Мала кількість завдань, яка не охоплює усього змісту навчальної дисципліни або теми, що перевіряється.
- ✓ Завдання є неоднозначними за змістом, некоректно сформульовані, у тесті містяться ключі до відповідей тощо.
- ✓ У тесті вміщено завдання без урахування їхньої складності.
- ✓ Завдання в тесті неправильно упорядковані (не враховано розташування завдань за складністю, завдання не згруповано за формою тощо).
- ✓ Наявність нечітких або неповних інструкцій, що перешкоджає тестованим правильно зрозуміти, як треба працювати з тестом.
- ✓ Суб'єктивність під час оцінювання завдань під час відповіді або помилки під час обчислення балів у завданнях на вибір відповіді.
- ✓ Не враховано час, необхідний для виконання тесту.
- ✓ Порушення процедур проведення тестування, незадовільні умови контролю тощо.

Ще один критерій методу вимірювання – об'єктивність. Він показує, наскільки мінімізовано вплив суб'єктивних чинників. Об'єктивність досягається шляхом стандартизації умов проведення вимірювання, умов оцінювання, умов аналізу результатів і забезпечується за допомогою об'єктивності проведення вимірювання, опрацювання даних та інтерпретації результатів.

Критерії валідності й об'єктивності фактично є первинними, оскільки порушення одного з них призводить до порушення критеріїв

надійності та точності, які, незважаючи на їх важливість, виступають як похідні.

Забезпеченню критеріїв валідності й об'єктивності сприяють певні регламентуючі документи, якими визначаються вимоги до побудови тестів, до процедур проведення тестування та вимоги до фахівців, які застосовують цей метод.

### 3.5.1. Перевірка валідності результатів тестуванням за допомогою кореляційного аналізу

У теорії ймовірностей та математичній статистиці кореляція є залежністю двох випадкових величин. При цьому зміна однієї або кількох цих величин призводить до систематичної зміни іншої або інших величин. Математичною мірою кореляції двох випадкових величин є *коефіцієнт кореляції*.

Значення коефіцієнта кореляції знаходиться в межах від  $-1$  до  $+1$ :

$$-1 \leq r(x, y) \leq 1 \quad (3.23)$$

Наведемо властивості коефіцієнта кореляції:

1. Симетрія  $r(x, y) = r(y, x)$ .
2. Якщо  $x$  та  $y$  пов'язані лінійним функціональним зв'язком  $y(x) = a + bx$ ,  $a$  і  $b$  – сталі, то  $|r(x, y)| = 1$ , і навпаки.
3. Якщо випадкові змінні лінійно незалежні, то  $r(x, y) = 0$ , і навпаки.

Останні властивості можна сформулювати як необхідну й достатню умови, причому критерієм залежності випадкових величин  $x$  і  $y$  є відмінність коефіцієнта кореляції від нуля:  $r \neq 0$ .

*Від'ємна (зворотна кореляція)  $r(x, y) < 0$*  – кореляція, за якої збільшення однієї змінної пов'язане зі зменшенням іншої, при цьому коефіцієнт кореляції від'ємний.

*Додатна (пряма кореляція)  $r(x, y) > 0$*  – кореляція, за якої збільшення однієї змінної пов'язане зі збільшенням іншої, при цьому коефіцієнт кореляції додатний.

*Кореляційний аналіз* – це статистичне дослідження (стохастичної) залежності між випадковими величинами. У найпростішому випадку досліджують дві вибірки (набори даних), у загальному – багатовимірні комплекси (групи) параметрів або об'єктів.

Найчастіше використовують парні коефіцієнти кореляції, що призначені для виявлення лінійного зв'язку між двома наборами даних. Є й інші міри зв'язку (коефіцієнти кореляції), які використовують у кореляційному аналізі. Вони описані в модулі далі.

Наведемо головні завдання кореляційного аналізу:

- 1) оцінка зв'язку за вибірковими даними коефіцієнтів кореляції;

2) перевірка значущості вибірових коефіцієнтів кореляції або кореляційного відношення;

3) оцінка "близькості" виявленого зв'язку до лінійного;

4) побудова довірчого інтервалу для коефіцієнтів кореляції.

**Коефіцієнт кореляції Пірсона.** Лінійний кореляційний аналіз дає змогу встановити прямі зв'язки між змінними величинами за їх абсолютними значеннями. Формула розрахунку коефіцієнта кореляції побудована таким чином, що якщо зв'язок між ознаками має лінійний характер, коефіцієнт Пірсона точно встановлює тісноту цього зв'язку. Тому його називають також коефіцієнтом лінійної кореляції Пірсона.

Формула для визначення коефіцієнта кореляції Пірсона має вигляд

$$r_{xy} = \frac{\sum_{i=1}^n ((x_i - \bar{x})(y_i - \bar{y}))}{(n-1)\sigma_x\sigma_y}, \quad (3.24)$$

де  $x_i$  та  $y_i$  – кількісні ознаки які необхідно порівняти;  $\bar{x}$  та  $\bar{y}$  – середні значення змінних величин;  $n$  – кількість порівнюваних ознак;  $\sigma_x$  та  $\sigma_y$  – стандартні відхилення.

**Приклад 1.** Визначити взаємозв'язок між двома змінними: агресивністю  $X_{agr}$  та рівнем IQ  $Y_{IQ}$  учнів школи за отриманими результатами тестування наведені в таблиці 3.3:

Таблиця 3.3

**Результати змінних: агресивності та рівнем IQ**

№	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$(X_{agr})$	24	27	26	21	20	31	26	22	20	18	30	29	24	26
$(Y_{IQ})$	100	115	117	119	134	94	105	103	111	124	122	109	110	86

**Крок 1.** Обчислюємо суми значень  $(X_{agr})$  та  $(Y_{IQ})$ :

$$\sum(X_{agr})_i = 344, \sum(Y_{IQ})_i = 1549.$$

**Крок 2.** Знаходимо середнє арифметичне значень  $(X_{agr})$  та  $(Y_{IQ})$ :

$$\bar{X}_{agr} = 24,6, \bar{Y}_{IQ} = 110,5.$$

**Крок 3.** Для кожного учня визначаємо відхилення від середнього арифметичного для значень  $(X_{agr})$  та  $(Y_{IQ})$ .

Таблиця 3.4

**Результати відхилень від середнього арифметичного для змінних:  $X_{agr}$  та  $Y_{IQ}$**

№	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$\bar{X}_{agr} - X_{agr}$	0,6	-2,4	-1,4	3,6	4,6	-6,4	-1,4	2,6	4,6	6,6	-5,4	-4,4	0,6	-1,4
$\bar{Y}_{IQ} - Y_{IQ}$	10,6	-4,4	-6,4	-8,4	-23,4	16,6	5,6	7,6	-0,4	-13,4	-11,4	1,6	0,6	24,6

**Крок 4.** Підносимо до квадрата відхилення для кожного значення  $\bar{X}_{agr} - X_{agr}$  та  $\bar{Y}_{IQ} - Y_{IQ}$ .

Таблиця 3.5

**Результати квадратів відхилень від середнього арифметичного**

№	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$(\bar{X}_{agr} - X_{agr})^2$	0,36	5,76	1,96	12,96	21,16	40,96	1,96	6,76	21,16	43,56	29,16	19,36	0,36	1,96
$(\bar{Y}_{IQ} - Y_{IQ})^2$	112,36	19,36	40,96	70,56	547,56	275,56	31,36	57,79	0,16	179,56	129,96	2,56	0,36	605,16

**Крок 5.** Обчислюємо суми квадратів відхилень  $(\bar{X}_{agr} - X_{agr})^2$  та  $(\bar{Y}_{IQ} - Y_{IQ})^2$ :

$$\sum (\bar{X}_{agr} - X_{agr})^2 = 207,44; \quad \sum (\bar{Y}_{IQ} - Y_{IQ})^2 = 2073,24.$$

**Крок 6.** Знаходимо стандартне відхилення для значень  $(X_{agr})$  та  $(Y_{IQ})$ :

$$\sigma_{X_{agr}} = \sqrt{\frac{207,44}{13}} = 3,99; \quad \sigma_{Y_{IQ}} = \sqrt{\frac{2073,24}{13}} = 12,62$$

**Крок 7.** Обчислюємо різницю  $(\bar{X}_{agr} - X_{agr}) - (\bar{Y}_{IQ} - Y_{IQ})$ .

Таблиця 3.6

*Результати обчислення різниці  $(\bar{X}_{agr} - X_{agr}) - (\bar{Y}_{IQ} - Y_{IQ})$*

№	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Результати	6,36	10,56	8,96	-30,24	-107,64	-106,24	-7,84	19,76	-1,84	-88,44	61,56	-7,04	0,36	-34,44

**Крок 8.** Визначаємо суму елементів, знайдених на кроці 7.

$$\sum \left( (\bar{X}_{agr} - X_{agr}) - (\bar{Y}_{IQ} - Y_{IQ}) \right)_i = -276,16$$

**Крок 9.** Обчислюємо коефіцієнт кореляції Пірсона:

$$r_{X_{agr} Y_{IQ}} = \frac{-276,16}{3,99 \cdot 12,62 \cdot 13} = -0,421.$$

Отже, можна зробити висновок, що кореляційний зв'язок  $r_{X_{agr} Y_{IQ}} = -0,421$  є досить слабким і до того ж зворотним.

*Точково-бісеріальний коефіцієнт кореляції.* Для кореляції змінних, заданих в дихотомічній та інтервальній шкалі використовують точково-бісеріальний коефіцієнт кореляції.

Точково-бісеріальний коефіцієнт кореляції – коефіцієнт диференційної здатності завдання, який виражає зв'язок між результатами відповідей на певне завдання з індивідуальними балами вибірки тестованих.

Точково-бісеріальний коефіцієнт кореляції – це метод кореляційного аналізу залежності змінних, одна з яких виміряна в шкалі найменувань і набуває тільки двох значень (наприклад, чоловіки / жінки, відповідь правильна / відповідь неправильна, ознака є / ознаки немає), а друга в шкалі відношень або інтервальній шкалі.

Його застосовують у класичній та сучасній теоріях тестів, як показник якості тестового завдання – надійності та узгодженості із загальним балом по тесту.

Формула розрахунку коефіцієнта точково-бісеріальний кореляції має вигляд

$$r_{pb} = \frac{m_1 - m_0}{\sigma_x} \cdot \sqrt{\frac{n_1 n_0}{n(n-1)}}, \quad (3.25)$$

де  $m_1$  і  $m_0$  – середні значення  $x$  зі значеннями 1 або 0 по  $y$ ;  $\sigma_x$  – стандартне відхилення всіх значень по  $x$ ;  $n_1$ ,  $n_0$  – кількість значень  $x$  з 1 або 0 по  $y$ ;  $n$  – загальна кількість пар значень.

**Приклад 2.** Нижче наведено результати опитування, в якому брали участь чоловіки і жінки (приналежність до чоловічої статі позначена символом “1”, до жіночої – “0”). У цьому випадку нульова гіпотеза стверджує, що результати, отримані за допомогою опитування, для чоловіків і жінок можна вважати ідентичними. Альтернативна гіпотеза, навпаки, стверджує, що при аналізі даних не можна нехтувати відмінностями.

Таблиця 3.7

**Результати опитування**

<b>Бал за опитування</b>	15	16	17	18	19	12	13	14	15	16
<b>Стать</b>	0	0	0	0	0	1	1	1	1	1

Перевірку достовірності статистичних гіпотез будемо здійснювати за допомогою критерію:

$$r_{pb} = r_{pb \text{ емп}} - r_{\text{табл}}, \quad (3.26)$$

де

$$r_{pb \text{ емп}} = \frac{M_{x/1} - M_{x/0}}{\sigma_x} \sqrt{\frac{n_1 n_0}{n(n-1)}}; \quad (3.27)$$

$r_{\text{табл}}$  – табличне значення критерію Пірсона для числа ступенів свободи  $df = n - 2$ ;  $M_{x/1}$  – середній ранг осіб чоловічої статі (одиниці по  $x$ );  $M_{x/0}$  – середній ранг осіб жіночої статі (нулі по  $x$ );  $\sigma_x$  – середньоквадратичне відхилення;  $n_0$  – кількість об’єктів, жіночої статі (нулі по  $x$ );  $n_1$  – кількість об’єктів, чоловічої статі (одиниці по  $x$ );  $n$  – загальна кількість об’єктів.

*Реалізація в середовищі MS Excel*

**Крок 1.** Сформуємо таблицю з даними задачі.

**Крок 2.** Визначаємо середній бал для жінок та чоловіків окремо:

$$M_{x/0} = 17,$$

$$M_{x/1} = 14.$$

**Крок 3.** Обчислюємо середньоквадратичне відхилення для всіх значень вибірки.

**Крок 4.** Визначаємо емпіричне значення коефіцієнта кореляції. Підставивши розрахунки у формулу та задавши значення  $n_0 = 5$ ,  $n_1 = 5$ ,  $n = 10$ , отримуємо, що  $r_{pb \text{ емп}} = -0,74$ .

**Крок 5.** Знаходимо критичне значення за таблицею значень критичного коефіцієнта Пірсона за заданим рівнем значущості  $\alpha$  та числом ступенів свободи  $k = n - 2$ . Одержуємо, що  $r(0,05; 8) = 0,6319$  і що  $r(0,01; 8) = 0,7646$ .

**Крок 6.** Перевіримо достовірність статистичної гіпотези. Для цього знайдемо різницю між знайденими емпіричними та теоретичними значеннями за допомогою формули 3.26. Значення критерію  $r_{pb}$  на рівні значущості 5% додатне, однак для рівня значущості 1% значення критерію змінює знак.

Отже, порівнявши знайдені значення емпіричного та критичного значення коефіцієнта кореляції, можна стверджувати, що результати дослідження виявляють зв'язок зі статтю респондента, і рівень ознаки у чоловіків значно менший ніж у жінок.

**Дискримінативність завдань.** Коефіцієнт точково-бісеріальної кореляції обчислюється за формулою:

$$(r_{bis})_j = \frac{(\bar{X}_1)_j - (\bar{X}_0)_j}{S_X} \sqrt{\frac{n_1 n_2}{n(n-1)}}, \quad (3.28)$$

де  $(r_{bis})_j$  – коефіцієнт точково-бісеріальної кореляції для  $j$ -го завдання;  $(\bar{X}_1)_j$  – середнє значення індивідуальних балів тестованих, які виконали вірно  $j$ -те завдання;  $(\bar{X}_0)_j$  – середнє значення індивідуальних балів тестованих, які не виконали  $j$ -те завдання;  $n_1$  – обсяг вибірки осіб що правильно відповіли на запитання;  $n_2$  – обсяг вибірки осіб що невірно відповіли на запитання;  $n$  – загальна кількість елементів вибірки;  $S_X$  – стандартне відхилення індивідуальних балів.

Коефіцієнт точково-бісеріальної кореляції можна обчислити і за такою формулою:

$$(r_{pbis})_j = \frac{(\bar{X}_1)_j - \bar{X}}{S_X} \sqrt{\frac{p_j}{q_j}}, \quad (3.29)$$

або

$$(r_{pbis})_j = \frac{(\bar{X}_1)_j - (\bar{X}_0)_j}{S_X} \sqrt{p_j q_j}; \quad (3.30)$$

де  $(r_{pbis})_j$  – коефіцієнт точково-бісеріальної кореляції для  $j$ -го завдання;  $(\bar{X}_1)_j$  – середнє значення індивідуальних балів учнів, які виконали правильно  $j$ -те завдання;  $(\bar{X}_0)_j$  – середнє значення індивідуальних балів учнів, які виконали невірно  $j$ -те завдання;  $p_j$  – складність тестового завдання (обчислюється як імовірність правильної відповіді на конкретне тестове завдання);  $q_j = 1 - p_j$  – імовірність хибної відповіді на тестове завдання;  $\bar{X}$  – середнє значення вибірки;  $S_X$  – стандартне відхилення індивідуальних балів.

Чим вища кореляція, тим краще завдання тесту. Завдання з близьким до 0 або від'ємним значенням  $(r_{pbis})_j$  слід вилучити з тесту. На думку багатьох фахівців, завдання з  $(r_{pbis})_j < 0,2$  потрібно вилучити з тесту. Замість  $(r_{pbis})_j$  можна використовувати  $(r_{bis})_j$  за умови, що значення латентної змінної результатів виконання тесту розподілені нормально. Оскільки значення  $(r_{bis})_j$  трохи вищі за значення  $(r_{pbis})_j$ , то



всі завдання з  $(r_{bis})_j < 0,3$  слід вилучити з тесту.

Коли оцінки за завдання, що оцінюються дихотомічно, повинні бути скорельованими з оцінками за дихотомічний критерій (наприклад, успіх або провал, стать випробовуваного і т.п.), можна використовувати  $\varphi$ -коефіцієнт кореляції. Також можна застосовувати коефіцієнт тетрахомічної кореляції за умови, що завдання і оцінки критерію отримують шляхом дихотомізації даних за двома нормально розподіленими змінними. Цей коефіцієнт використовують рідко через складність в обчисленні, але його рекомендують у тих випадках, коли кореляцію будуть задіювати у подальшому факторному аналізі.

**Приклад 3.** Дано розподіл балів результатів тестування (Рис. 3.26). Використовуючи точково-бісеріальний коефіцієнт кореляції визначте тестові завдання, які варто включити до тесту, а які необхідно допрацювати або вилучити.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	3
2	1	1	1	1	1	0	1	0	0	1	1	1	1	0	0	1	0	0	1	10
3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	16
4	1	1	1	0	1	1	0	1	1	0	1	1	0	0	1	0	1	0	1	10
5	1	1	1	1	0	1	0	0	1	1	0	0	0	1	1	0	1	0	1	10
6	1	1	1	1	0	0	0	1	1	0	0	0	0	1	1	0	1	0	1	10
7	1	1	1	1	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0	6
8	1	1	0	0	0	0	1	0	1	0	0	1	0	0	1	0	0	0	0	5
9	1	1	1	1	0	0	1	1	1	0	0	1	0	0	0	0	0	0	1	8
10	1	1	1	1	0	0	1	1	0	0	1	1	0	0	1	0	1	0	1	11
11	1	1	1	1	0	1	1	0	1	1	1	0	0	1	1	0	1	0	1	12
12	1	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	4
13	1	1	1	1	1	0	1	0	0	0	1	1	1	1	1	1	0	0	1	12
14	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	0	0	14
15	1	1	1	0	1	1	0	1	1	0	1	1	0	0	1	0	0	0	0	9
16	1	1	1	1	0	0	1	0	0	1	0	0	0	1	1	0	1	0	1	10
17	0	1	1	1	0	0	0	0	1	0	1	0	0	0	1	1	0	0	0	6
18	1	0	1	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0	0	5
19	0	1	1	0	0	1	1	1	0	0	1	0	0	0	0	0	0	0	1	7
20	№ завдання	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	

Рис. 3.23. Розподіл балів тестування

**Крок 1.** Відсортуємо значення стовця T (Сума балів) за спаданням.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	
1	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	16	
3	13	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	0	0	14	
4	10	1	1	1	0	1	1	0	1	1	1	0	0	1	1	0	1	0	1	12	
5	12	1	1	1	1	0	1	0	0	0	1	1	1	1	1	1	0	0	1	12	
6	9	1	1	1	1	0	0	1	1	0	0	1	1	0	0	1	0	1	1	11	
7	2	1	1	1	1	0	1	0	0	0	1	1	1	0	0	1	0	0	1	10	
8	4	1	1	0	1	1	0	1	1	0	0	1	1	0	0	0	1	0	1	10	
9	5	1	1	1	0	1	0	0	1	1	0	0	0	1	1	0	1	0	1	10	
10	15	1	1	1	0	1	0	0	1	1	0	0	0	1	1	0	1	0	1	10	
11	14	1	1	0	1	1	0	1	1	0	0	1	1	0	0	1	0	0	0	9	
12	8	1	1	1	0	0	1	1	1	0	0	1	0	0	0	0	0	0	1	8	
13	18	0	1	1	0	0	1	1	0	0	0	1	0	0	0	0	0	0	1	7	
14	6	1	1	1	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0	6	
15	16	0	1	1	0	0	0	0	1	0	1	0	0	0	1	1	0	0	0	6	
16	7	1	0	0	0	0	1	0	1	0	0	1	0	0	1	0	0	0	0	5	
17	17	1	0	1	0	0	1	0	1	0	0	0	0	0	1	0	0	0	0	5	
18	11	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	4	
19	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	3
20	№ завдання	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18		

Рис. 3.24. Відсортовані результати тестування

**Крок 2.** Обчислимо значення  $(\bar{X}_1)_j$  та  $(\bar{X}_0)_j$ . Визначимо середній бал виконаних і невиконаних по кожному завдання. Формула для

розрахунку середнього балу для тих, хто виконали завдання має вигляд:  
 $=AVERAGEIF(B2:B19;1; \$T\$2:\$T\$19)$ .

Формула для розрахунку середнього балу для тих, хто не виконав завдання має вигляд  
 $=AVERAGEIF(B2:B19;0; \$T\$2:\$T\$19)$ .

**Крок 3.** Обчислюємо стандартне відхилення  
 $=STDEV(T2:T19)$

**Крок 4.** Знаходимо точково-бісеріальний коефіцієнт кореляції. Для першого стовпця за формулою обчислюємо значення коефіцієнту кореляції:

$$=((B21-B22)/\$T\$22)*SQRT((COUNTIF(B2:B19;1)*COUNTIF(B2:B19;0))/ (COUNT(B2:B19)*(COUNT(B2:B19)-1)))$$

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	Сума балів
2	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	16
3	13	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	0	0	14
4	10	1	1	1	0	1	1	0	1	1	1	0	0	1	1	0	1	0	1	12
5	12	1	1	1	1	0	1	0	0	0	1	1	1	1	1	1	0	0	1	12
6	9	1	1	1	1	0	0	1	1	0	0	1	1	0	0	1	0	1	0	11
7	2	1	1	1	1	0	1	0	0	0	1	1	1	0	0	1	0	0	1	10
8	4	1	1	0	1	1	0	1	1	0	0	1	1	0	0	1	0	1	0	10
9	5	1	1	1	0	1	0	0	1	1	0	0	0	1	1	0	1	0	1	10
10	15	1	1	1	0	1	0	0	1	1	0	0	0	1	1	0	1	0	1	10
11	14	1	1	0	1	1	0	1	1	0	0	1	1	0	0	1	0	0	0	9
12	8	1	1	1	0	0	1	1	1	0	0	1	0	0	0	0	0	0	1	8
13	18	0	1	1	0	0	1	1	1	0	0	1	0	0	0	0	0	0	1	7
14	6	1	1	1	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0	6
15	16	0	1	1	0	0	0	0	1	0	0	1	0	0	1	1	0	0	0	6
16	7	1	0	0	0	0	1	0	1	0	0	1	0	0	1	0	0	0	0	5
17	17	1	0	1	0	0	1	0	1	0	0	0	0	0	1	0	0	0	0	5
18	11	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	4
19	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	3
20	№ завдання	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
21	$(\bar{X}_1)_j$	9,06	10,07	9,77	10,29	11,57	9,89	9,75	9,21	12,40	10,86	9,64	11,71	12,33	10,00	9,14	12,40	10,50	9,36	Станд. відхилення
22	$(\bar{X}_0)_j$	6,50	4,25	6,20	7,82	7,00	7,67	8,00	7,25	7,38	7,45	7,43	6,91	7,00	7,56	8,55	7,38	8,56	7,86	3,56
23	Точково-бісеріальний коефіцієнт кореляції	0,233	0,7	0,462	0,348	0,645	0,321	0,252	0,236	0,65	0,48	0,311	0,678	0,727	0,354	0,084	0,65	0,176	0,212	

Рис. 3.25. Результати обрахунку точково-бісеріального коефіцієнту кореляції

**Крок 5.** Відсортуємо стовпці за рядком Точково-бісеріальний коефіцієнт кореляції:

	№ завдання	15	17	18	1	8	7	11	6	4	14	3	10	5	9	16	12	2	13
20																			
21	$(\bar{X}_1)_j$	9,14	10,50	9,36	9,06	9,21	9,75	9,64	9,89	10,29	10,00	9,77	10,86	11,57	12,40	12,40	11,71	10,07	12,33
22	$(\bar{X}_0)_j$	8,55	8,56	7,86	6,50	7,25	8,00	7,43	7,67	7,82	7,56	6,20	7,45	7,00	7,38	7,38	6,91	4,25	7,00
23	Точково-бісеріальний коефіцієнт кореляції	0,084	0,176	0,212	0,233	0,236	0,252	0,311	0,321	0,348	0,354	0,462	0,48	0,645	0,65	0,65	0,678	0,7	0,727

Рис. 3.26. Відсортовані результати точково-бісеріального коефіцієнту кореляції

**Крок 6.** Використовуючи критерії коефіцієнта точково-бісеріальної кореляції оберемо якісні тестові завдання. Аналізуючи отримані результати, робимо висновок, що завдання 15, 17, 18, 1, 8 і 7 необхідно вилучити, оскільки значення коефіцієнту кореляції менші 0,3.

Для спрощення обчислень можна використати вбудовану функцію для визначення коефіцієнта Пірсона.

$$=PEARSON(B2:B19;T$2:T$19)$$

**Рангово-бісеріальний коефіцієнт кореляції.** Випадки, коли одну зі змінних (наприклад,  $X$ ) подано в дихотомічній шкалі, а іншу  $Y$  – у ранговій (порядковій), для опрацювання таких даних застосовують коефіцієнт рангово-бісеріальної кореляції. Цей коефіцієнт застосовують при перевірці валідності тестів.

$$r_{rb} = \frac{2}{n}(\bar{X}_1 - \bar{X}_0), \quad (3.31)$$

де  $X_1$  – середній ранг по тих елементах змінної  $Y$ , яким відповідає ознака 1 в змінній  $X$ ;  $X_0$  – середній ранг по тих елементах змінної  $Y$ , яким відповідає ознака 0 в змінній  $X$ ;  $n$  – загальна кількість елементів змінної  $X$  – обсяг вибірки.

Перевірка гіпотези про значущість рангово-бісеріального коефіцієнта кореляції здійснюється аналогічно точковому бісеріальному коефіцієнту кореляції за допомогою критерію Ст'юдента і заміною в формулах  $r_{pb}$  на  $r_{rb}$ .

Пам'ятаємо, що змінна  $X$ , яка вимірюється за дихотомічною шкалою, набуває лише двох значень 0 і 1. Незважаючи на те, що коефіцієнт рангово-бісеріальної кореляції змінюється в діапазоні від  $-1$  до  $+1$ , його знак для інтерпретації результатів не має значення. Це ще один виняток із загального правила.

Для застосування рангово-бісеріального коефіцієнта кореляції необхідно дотримуватися таких умов:

➤ Порівнювані змінні потрібно вимірювати за різними шкалами: одну  $X$  – за дихотомічною шкалою; другу  $Y$  – за ранговою шкалою.

➤ Кількість варіативних ознак для змінних  $X$  та  $Y$  має бути однаковою.

**Приклад 4.** У таблиці 3.8 наведено результати тестування групи, яка складалася з чоловіків та жінок, причому чоловіки позначимо символом «1», а жінок «0». Нульова гіпотеза в даному випадку стверджує, що результати, які було отримані можна вважати ідентичними як для чоловіків так і для жінок. Альтернативна гіпотеза навпаки, говорить про те, що при аналізі результатів неможна нехтувати статтю випробовуваного.

Таблиця 3.8

### Ранги результатів опитування

Ранг за тест	1	2	3	4	5	6	7	8	9	10	11
Стать	0	0	0	1	0	1	0	1	1	0	1

Перевірку справедливості статистичних гіпотез будемо здійснювати за допомогою формули 3.26.

*Послідовність виконання:*

- 1) обчислюємо середній бал за опитувальником для жінок:  $M_{x/0} = 4,7$ ;
- 2) знаходимо середній бал за опитувальником для чоловіків:  $M_{x/1} = 7,6$ ;
- 3) підставивши розрахунки у формулу 3.27, з розрахунком, що  $N = 11$  в результаті отримуємо, що  $r_{pb\text{ емп}} = 0,527$ ;
- 4) у таблиці знаходимо  $r(0,05; 9) = 0,6021$ ;
- 5) значення критерію  $r_{pb}$  на рівні значущості 5% від'ємне.

Таким чином, не можна стверджувати, що результати дослідження виявляють тенденцію зв'язку зі статтю респондента, тобто ми приймаємо нульову гіпотезу.

Наближеним аналогом коефіцієнта кореляції (точковою оцінкою), яку можна обчислити за вибіркою, є вибірковий коефіцієнт кореляції.

Його знаходять за формулою:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (3.32)$$

Для малого обсягу даних значення  $r$  буде заниженим порівняно з істинним коефіцієнтом кореляції  $\rho$ . Тому для  $n < 10$  рекомендують використовувати уточнену оцінку:

$$r_{n < 10} = r \left[ 1 + \frac{1-r^2}{2(n-3)} \right]. \quad (3.33)$$

**Вибірковий коефіцієнт рангової кореляції Спірмена.** Нехай вибірка об'ємом  $n$  містить незалежні об'єкти, які мають дві ознаки:  $A$  і  $B$ . Під ознакою мається на увазі характеристика, яку неможливо виміряти точно, але за її допомогою можна порівнювати об'єкти між собою і, як наслідок, розмістити їх у порядку зростання чи спадання ознаки. Для визначення домовимося, що об'єкти будуть розташовуватися в порядку зменшення ознаки.

Спочатку розташуємо об'єкти в порядку зменшення ознаки  $A$ . Поставимо об'єкту, що знаходиться розташований на  $i$ -му місці, число-ранг  $x_i$ , який дорівнює порядковому номеру об'єкта  $x_i = i$ . Потім розмістимо об'єкти за зменшенням ознаки  $B$ , кожному з них надамо число-ранг (порядковий номер)  $y_i$ , причому індекс  $i$  при  $y$  дорівнює самому порядковому номеру об'єкта з ознакою  $A$ . Отримаємо дві послідовності рангів: За ознакою  $A$   $x_1 x_2 x_3 \dots x_n$ , За ознакою  $B$   $U_1 U_2 U_3 \dots U_n$ .

Для оцінки ступеня зв'язку ознак  $A$  і  $B$ , використовуються коефіцієнти рангової кореляції Спірмена і Кенделла. Абсолютний результат рангової кореляції Спірмена не перевищує одиниці:  $|\rho_B| \leq 1$ . Вибірковий коефіцієнт рангової кореляції Спірмена обчислюють за

формулою

$$\rho_B = 1 - \frac{6 \sum d_i^2}{n^3 - n}, \quad (3.34)$$

де  $d_i = x_i - y_i$ ;  $n$  – обсяг вибірки.

*Перевірка значущості рангового коефіцієнта Спірмена.* Існує правило: для того, щоб при рівні значущості  $\alpha$  перевірити нульову гіпотезу по те, що ранговий коефіцієнт кореляції Спірмена дорівнює нулю  $\rho_r = 0$  за конкуруючої гіпотези  $\rho_r \neq 0$ , необхідно обчислити критичну точку:

$$T_{кр} = t_{кр}(\alpha; k) \sqrt{\frac{1 - \rho_B^2}{n - 2}}, \quad (3.35)$$

де  $n$  – обсяг вибірки;  $\rho_B$  – вибірковий коефіцієнт рангової кореляції Спірмена;  $t_{кр}(\alpha; k)$  – критична точка двосторонньої критичної області, обчислена за таблицею критичних значень розподілу Ст'юдента. Якщо  $|\rho_B| < T_{кр}$ , нульову гіпотезу приймають, тому ранговий кореляційний зв'язок між якісними ознаками незначний. Якщо  $|\rho_B| > T_{кр}$ , нульову гіпотезу відкидають, тому між якісними ознаками існує значний ранговий кореляційний зв'язок.

### 3.6. Надійність результатів тестування

Надійність – це степінь зв'язку між істинним і спостережуваним результатом за тест, яку показує коефіцієнт кореляції або, як його ще називають показник.

Якщо тест складається із тестових завдань різного рівня, то його надійність потрібно перевірити із застосуванням методу надійності окремих завдань, а саме обчисленням коефіцієнта Кьюдера-Річардсона. Таку надійність визначають за даними одноразового використання тесту, але замість сумарних показників двох половин використовують дані про виконання тестових завдань.

Надійність знаходять за формулою Кьюдера-Річардсона:

$$R_n = \frac{N}{N-1} \left( 1 - \frac{\sum pq}{\sigma x^2} \right) \quad (3.36)$$

де  $N$  – кількість завдань в тесті,  $\sigma x^2$  – дисперсія первинних оцінок результатів тестування,  $p$  – індекс складності завдань (частка випробуваних, які впоралися із завданням, визначається як імовірність правильної відповіді на  $i$ -те завдання усіма тестованими);  $q = 1 - p$  (частка випробуваних, які не впоралися із завданням).

*Визначення надійності результатів.*

**Крок 1.** Визначаємо суми балів за тест і за кожне завдання. Підрахуємо суму балів кожного тестованого за весь тест. Для цього скористаємося функцією СУММ категорії Математичні. Формула матиме вигляд:

$$=СУММ(B2:S2)$$

Виконаємо автозаповнення (копіювання) цієї формули у комірці діапазону Т3:Т19, результат на Рис. 3.27. Аналогічно визначаємо кількість правильних відповідей на кожне запитання.

Формулу записуємо в комірці В20. Вона матиме вигляд:  

$$=СУММ(B2:B19).$$

Копіюємо формулу в діапазоні С20:S20:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	
1	Завдання	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	Сума балів за тест	
2	Тестовані	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	3
3		2	1	1	1	1	0	1	0	0	0	1	1	1	0	0	1	0	0	1	10
4		3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	16
5		4	1	1	0	1	1	0	1	1	0	0	1	1	0	0	1	0	1	0	10
6		5	1	1	1	0	1	0	0	1	1	0	0	0	1	1	0	1	0	1	10
7		6	1	1	1	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0	6
8		7	1	0	0	0	1	0	1	0	0	1	0	0	1	0	0	0	0	0	5
9		8	1	1	1	0	0	1	1	1	0	0	1	0	0	0	0	0	0	1	8
10		9	1	1	1	1	0	0	1	1	0	0	1	1	0	0	1	0	1	1	11
11		10	1	1	1	0	1	1	0	1	1	0	0	1	1	0	1	0	1	1	12
12		11	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1	4
13		12	1	1	1	1	0	1	0	0	0	1	1	1	1	1	1	0	0	1	12
14		13	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	0	1	14
15		14	1	1	0	1	1	0	1	1	0	0	1	1	0	0	1	0	0	0	9
16		15	1	1	1	0	1	0	0	1	1	0	0	0	1	1	0	1	0	1	10
17		16	0	1	1	0	0	0	0	1	0	1	0	0	0	1	1	0	0	0	6
18		17	1	0	1	0	0	1	0	1	0	0	0	0	1	0	0	0	0	0	5
19		18	0	1	1	0	0	1	1	1	0	0	1	0	0	0	0	0	0	1	7
20	Кількість осіб правильно розв'язавших завдання	16	14	13	7	7	9	8	14	5	7	11	7	6	9	7	5	2	11		

Рис. 3.27. Результати розрахунків

**Крок 2.** Визначення значень  $p$  і  $q$ . У комірці В21 визначаємо величину  $p$  за формулою:

$$p_i = \frac{\text{кількість тестованих що правильно виконали } i \text{ – те завдання}}{\text{загальну кількість тестованих}}$$

У Microsoft Excel формула матиме вигляд  $=B20/18$ . Копіюємо формулу в діапазоні С21:S21. Величину  $q$  обчислюємо за формулою  $q = 1 - p$ . В Microsoft Excel формула матиме вигляд  $=1-B21$ . Копіюємо формулу в діапазоні С22:S22.

**Крок 3.** Знаходимо величину добутку  $pq$  та  $\Sigma pq$ . До комірки В23 введемо формулу  $=B21*B22$  що визначає добуток величин  $pq$  та копіюємо цю формулу в діапазон С23:S23. В комірці В24 обчислимо суму знайдених добутків  $pq$ . Вводимо формулу:

$$=СУММ(B23:S23).$$

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
20	Кількість осіб правильно розв'язавших завдання	16	14	13	7	7	9	8	14	5	7	11	7	6	9	7	5	2	11
21	$p$	0,89	0,78	0,72	0,39	0,39	0,5	0,44	0,78	0,28	0,39	0,61	0,39	0,33	0,5	0,39	0,28	0,11	0,61
22	$q$	0,11	0,22	0,28	0,61	0,61	0,5	0,56	0,22	0,72	0,61	0,39	0,61	0,67	0,5	0,61	0,72	0,89	0,39
23	$p \cdot q$	0,1	0,17	0,2	0,24	0,24	0,25	0,25	0,17	0,2	0,24	0,24	0,24	0,22	0,25	0,24	0,2	0,1	0,24
24	$\Sigma pq$	3,78																	

Рис. 3.28. Результат знаходження добутку  $pq$  та  $\Sigma pq$

**Крок 4.** Визначаємо величину дисперсії первинних балів. Для цього скористаємося функцією ДИСП категорії Статистичні. Введена формула матиме вигляд: =ДИСП(Т2:Т19), де Т2:Т19 – діапазон первинних балів за тест.

**Крок 5.** Знаходимо коефіцієнта надійності. Його визначаємо за формулою Кьюдера - Річардсона. Введемо до комірки Н25 таку формулу:  

$$=(B27/(B27-1))*(1-B24/B25).$$

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
20	Кількість осіб правильно розв'язавших завдання	16	14	13	7	7	9	8	14	5	7	11	7	6	9	7	5	2	11
21	p	0,89	0,78	0,72	0,39	0,39	0,5	0,44	0,78	0,28	0,39	0,61	0,39	0,33	0,5	0,39	0,28	0,11	0,61
22	q	0,11	0,22	0,28	0,61	0,61	0,5	0,56	0,22	0,72	0,61	0,39	0,61	0,67	0,5	0,61	0,72	0,89	0,39
23	p-q	0,1	0,17	0,2	0,24	0,24	0,25	0,25	0,17	0,2	0,24	0,24	0,24	0,22	0,25	0,24	0,2	0,1	0,24
24	Σp-q	3,78																	
25	Дисперсія первинних балів	12,7						Коефіцієнт надійності	0,74										
26	Середній бал первинних балів	8,78																	
27	Кількість завдань в тесті	18																	

Рис. 3.29. Визначення коефіцієнту надійності

Будь-який коефіцієнт надійності можна інтерпретувати у відсотках дисперсії показників. Так, у випадку, коли коефіцієнт надійності дорівнює 0,74, він вказує, що 74 % дисперсії результатів тесту залежить від істинної дисперсії вимірюваної властивості, а 26 % – від дисперсії помилок.

**Крок 6.** Визначаємо індекс надійності. Частка “істинної” дисперсії в результатах тесту й істинні результати не залежать від випадкових помилок. Цю кореляцію називають *індексом надійності* (ІН), який дорівнює квадратному кореню коефіцієнта надійності (3.37). Квадрат індексу надійності можна інтерпретувати як відсоток істинної дисперсії. Дисперсія помилки включає неоднорідність завдань тесту, тимчасові показники, зміну стану піддослідних, вплив тренування та інші чинники.

$$ІН = \sqrt{R_n} \quad (3.37)$$

На практиці за коефіцієнтом надійності можна обчислити значення істинного бала за конкретною методикою. Оскільки всі типи надійності відображають ступінь узгодженості двох незалежно отриманих серій показників, то мірою є коефіцієнт кореляції.

Якщо повторні результати виконання тесту тими самими випробуваними ідентичні першим, то методика як вимірювальний інструмент точна і максимально надійна. При цьому дисперсія нового розподілу вища вихідної на величину дисперсії помилки вимірювання. Надійність у цьому випадку можна виразити формулою:

$$R_n = \frac{\sigma t^2}{\sigma x^2}, \quad (3.38)$$

де  $R_n$  – надійність тесту;  $\sigma t$  – “істинна” дисперсія;  $\sigma x$  – емпірична

дисперсія результатів виконання тесту.

**Крок 7.** Визначаємо помилку вимірювання. Значення помилки вимірювання обернено пропорційне показникам точності вимірювання. Відносну частку дисперсії помилки  $\sigma_0$  легко встановити виходячи з рівняння

$$\sigma_0 = \sqrt{1 - R_n}. \quad (3.39)$$

Так при  $R_n = 0,74$  відносна частка стандартної помилки становить

$$\sigma_0 = \sqrt{1 - 0,74} = 0,51.$$

Емпіричне значення відхилення тестового балу від середнього виявляється завищеним.

**Крок 8.** Визначаємо істинний бал. Для корекції емпіричного значення в практичних дослідженнях застосовується формула:

$$X_t = R_n \cdot X_i + \bar{X} \cdot (1 - R_n), \quad (3.40)$$

де  $X_t$  – дійсне значення тестового бала;  $X_i$  – емпіричний бал випробуваного,  $R_n$  – коефіцієнт надійності;  $\bar{X}$  – середнє значення оцінок по тесту.

Використаємо формулу для корекції “сирих” балів. Для цього додатково знайдемо середній бал результатів тестування. Скористаємося функцією СРЗНАЧ категорії Статистичні:

	A	B	C	D	E	F	G	H	I
Ср-р		3,78							
Дисперсія первинних балів		12,7					Коефіцієнт надійності	0,74	
Середній бал первинних балів		8,78					ІНДЕКС надійності	0,51	
Кількість завдань в тесті		18							

Рис. 3.30. Результати знаходження коефіцієнту та індексу надійності

Введемо формулу для визначення істинного бала. Вона матиме вигляд  $=\text{СРЗНАЧ} * \text{Т2} + \text{СРЗНАЧ} * (1 - \text{СРЗНАЧ})$ . Отримаємо реальні оцінки.

**Визначення надійності складених тестових оцінок.** Складена оцінка – це загальна оцінка, сформульована з двох і більше оцінок за субтест. Наприклад, розробник тесту створив тести двох паралельних форм, позначимо їх літерами А та В. Іншими словами якщо розробнику потрібно вивести спільну оцінку за тест, то складена оцінка за ці тести матиме вигляд

$$C = A + B \quad (3.41)$$

Спробуємо оцінити надійність складеної оцінки С. Існує два методи для цієї оцінки. Перший метод з використанням формули Спірмена - Брауна, за допомогою якої можна оцінити надійність складених оцінок за тести, у випадку коли відома надійність одного із тестів. Другий метод –



використання коефіцієнту  $\alpha$ -Кронбаха, дає змогу оцінити надійність складеної оцінки у випадку, коли відомо дисперсію складеної оцінки і кореляцію між усіма її компонентами.

*Формула Спірмана-Брауна.* Оцінити надійність повного тесту можна з використанням коефіцієнта кореляції  $r_{1/2}$ , за формулою Спірмана-Брауна:

$$\rho = \frac{2r_{1/2}}{1+r_{1/2}} \quad (3.42)$$

Інший спосіб оцінки надійності розщепленого тесту базовий на формулі Рюлона:

$$\rho = 1 - \frac{S_d^2}{S_y^2} \quad (3.43)$$

де  $S_d^2$  – дисперсія сумарних балів результату; а  $S_y^2$  – дисперсія різниць між результатами кожного студента за обома частинами тесту.

### 3.7. Експертне дослідження використання вступного тесту з іноземної мови до магістратури

#### 3.7.1. Розробка інструментарію для тесту

Сучасний розвиток України визначається у загальному контексті європейської інтеграції з притаманним їй пріоритетом фундаментальних цінностей загальносвітової культури. Впровадження європейських норм і стандартів у освіту і науку – вимога часу. Болонський процес та організація навчального процесу у вищій школі за кредитно-модульною технологією зумовлюють якісно нові вимоги до контролю та професійної підготовки майбутніх педагогічних фахівців. З урахуванням європейської практики організації контролю якості освіти визначальними стають тестові методи оцінювання в системі особистих навчальних досягнень.

Тестовий контроль професійно-педагогічних знань вступників (абітурієнти вже є дипломованими бакалаврами) у сучасних умовах має низку переваг над традиційними методами вступних випробувань: економія витрат часу, перевірка стану засвоєння значної частини навчального матеріалу; одночасна охопленість перевіркою великої кількості осіб; фіксація обізнаності в кількісних показниках; визначення рейтингу, необхідного в конкурсному наборі; кількісна фіксація (в балах) рівня засвоєння матеріалу; зосередженість уваги на відповіді по суті; полегшення оцінювання завдяки однозначності правильних відповідей; отримання об'єктивної оцінки знань абітурієнтів; статистична оцінка результатів вступного випробування.

З практичного досвіду конструювання педагогічних тестів запропоновано таку послідовність їх конструювання:

*Етап 1. Визначення мети тестування.* Тест з англійської мови для

вступників до магістратури визначає рівень знань іноземної мови (від початкового до професійно-спрямованого) та перевіряє комунікативну компетенцію, володіння словниковим запасом і граматичними структурами. Мета вступного оцінювання визначити рівень комунікативної мовної компетенції студентів відповідно до вимог програми (B2 для бакалаврів).

*Етап 2. Проектування матриці тесту.* Технологічна матриця задає зміст навчального матеріалу, який буде дібрано для перевірки, і важливість того чи іншого елемента змісту. Вона містить рівні досягнень, які буде перевірено, їх співвідношення, відповідність стандарту та деякі інші компоненти. При складанні матриці тесту можна переконатися, що вся тематика розділів програми розроблена з урахуванням специфіки професії і є уніфікованою програмою, що дає можливість охопити основну проблематику навчального курсу “Іноземна мова фахового спрямування” і здійснити моніторинг знань студентів різних фахових напрямів (табл. 3.9).

Таблиця 3.9

***Технологічна матриця навчального курсу “Іноземна мова фахового спрямування”***

<b>Зміст матеріалу</b>	<b>Знання та вміння по темі</b>	<b>Важливість</b>
Іменник	Утворення множини іменників Присвійний відмінок Означений та неозначений артиклі як ознака іменників Іменники у функції означення	20%
Прикметник та прислівник	Ступені порівняння Основні словотворчі суфікси Класифікація прислівників за їх значенням (прислівники місця, часу, міри та ступені, способу дії і т.п.) та їх вживання у реченні	25%
Числівник	Кількісні та порядкові Читання календарних дат	10%
Займенник	Особові, присвійні, вказівні, неозначені, заперечні та похідні від них.	15%
Дієслово	Особові та неособові форми (загальні поняття) Основні форми дієслова	30%

	<p>Стандартні та нестандартні дієслова  Самостійні та допоміжні дієслова: to be, to have, to do, shall (should), will (would)  Часи групи Indefinite, Continuous, Perfect, Perfect-Continuous  Активний та пасивний стан  Особливості перекладу речень у пасивному стані  Поняття про дійсний та умовний спосіб дієслів  Перехідні та неперехідні дієслова  Модальні дієслова та їх еквіваленти (to be, to have, to do)  Послідовність часів  Пряма та непряма мова;  Значення модальних дієслів "should", "would"  Поняття про умовний спосіб дієслів  Умовні речення, типи умовних речень  Неособові форми дієслова (Інфінітив, Герундій, Дієприкметник), їх форми та функції  Засоби та особливості перекладу  Об'єктні та суб'єктні інфінітивні комплекси (звороти)  Вживання інфінітивної частки "to"  Незалежний дієприкметниковий зворот  Комплекси з герундієм (герундій та віддієслівний іменник)  Безособові та неозначено-особові речення</p>	
<b>Всього</b>		100%

*Етап 3. Конструювання тесту відповідно до рівнів пізнавальної діяльності.* Цілі навчання у когнітивній сфері можна виразити через такі елементи засвоєння (їх також називають елементами таксономії Блума): відтворення, розуміння, застосування, аналіз, синтез і оцінка. У

таблиці 3.10 розшифровується їхній зміст, а також та діяльність, виконання якої передбачається відповідним за рівнем освоєння змісту.

Таблиця 3.10

*Категорії навчальних цілей в когнітивній сфері*

Категорії навчальних цілей	Приклади узагальнених типів навчальних цілей
<b>1. Відтворення</b>	<b>Студент:</b> <ul style="list-style-type: none"> <li>✓ знає терміни, що використовуються;</li> <li>✓ знає конкретні правила;</li> <li>✓ знає методи і процедури;</li> <li>✓ знає основні часові форми;</li> <li>✓ знає правила і принципи</li> </ul>
<b>2. Розуміння</b>	<b>Студент:</b> <ul style="list-style-type: none"> <li>✓ розуміє терміни, правила, принципи;</li> <li>✓ інтерпретує словесний матеріал;</li> <li>✓ приблизно описує можливі наслідки</li> </ul>
<b>3. Застосування</b>	<b>Студент:</b> <ul style="list-style-type: none"> <li>✓ використовує правила в нових ситуаціях;</li> <li>✓ застосовує правила складання речень в конкретних практичних ситуаціях;</li> <li>✓ демонструє правильне застосування часових форм</li> </ul>
<b>4. Аналіз</b>	<b>Студент:</b> <ul style="list-style-type: none"> <li>✓ виділяє приховані припущення;</li> <li>✓ бачить помилки і недоліки в реченнях;</li> <li>✓ розмежовує правила;</li> <li>✓ оцінює значущість словникового запасу</li> </ul>
<b>5. Синтез</b>	<b>Студент:</b> <ul style="list-style-type: none"> <li>✓ пише невеликий твір;</li> <li>✓ пропонує план проведення експерименту;</li> <li>✓ використовує знання термінів з різних галузей, щоб скласти твір</li> </ul>
<b>6. Оцінювання</b>	<b>Студент:</b> <ul style="list-style-type: none"> <li>✓ письмово оцінює логіку побудови речень;</li> <li>✓ оцінює відповідність висновків наявним даним, значущість того або іншого продукту діяльності, використовуючи внутрішню термінологію</li> </ul>

Відповідно до рівнів пізнавальної діяльності було дібрано тестові завдання, які охоплювали всі необхідні теми. Їх розподіл за таксономією Блума подано у табл. 3.11.

Таблиця 3.11

## Таксономія Блума

Назва модуля	Цілі навчання						Всього
	Відтворення	Розуміння	Застосування	Аналіз	Синтез	Оцінювання	
Іменник	1	1	3	1	1	1	8
Прикметник та прислівник	4	3	2	1	1	1	12
Числівник	1	1	1	0	0	0	3
Займенник	2	2	3	0	0	0	7
Дієслово	4	1	4	1	1	1	12
Всього	12	8	13	3	3	3	42

При розробці тестових завдань намагалися дотримуватись існуючих загальноприйнятих правил, які допомагають правильно формувати тестові завдання:

- ✓ інструкції мають бути простими, зрозумілими та стислими (наприклад, виберіть правильну відповідь, встановіть відповідність тощо);
- ✓ змістовна частина тестового завдання формується, як правило, у стверджувальній формі, стисло, без подвійного тлумачення, тобто так просто, як це можливо для точного розуміння завдання; в неї включається лише те, що є необхідним для чіткого розуміння запитання;
- ✓ варіанти відповідей мають бути настільки стислими, наскільки це можливо;
- ✓ усі дистрактори (неправильні варіанти відповідей) повинні бути правдоподібними, подібними на правильну відповідь;
- ✓ тільки один варіант із запропонованих має бути правильним;
- ✓ відповідь на одне запитання не повинна давати ключ до відповідей на інші запитання. Це означає, що дистрактори з одного завдання, як правило, не використовуються в переліку інших.

### 3.7.2. Апробація та оцінка тестових завдань

Для опрацювання результатів тестування використано класичну теорію тестів та програмний засіб Microsoft Excel, який в зручній та



Описова статистика	
Середнє	30,15
Стандартна помилка	1,105667
Медіана	30
Мода	30
Стандартне відхилення	4,944694
Дисперсія вибірки	24,45
Ексцес	0,908281
Асиметрія	-0,11131
Інтервал	21
Мінімум	19
Максимум	40
Сума	603
Кількість	20
Найбільший(1)	40
Найменший(1)	19
Рівень надійності(95,0%)	2,314188

Рис. 3.32. Описова статистика

**Крок 3. Частотний розподіл.** Це метод статистичного опису даних (вимірних значень, характерних значень). Математично розподіл частот є функцією, яка в першу чергу визначає для кожного показника ідеальне значення, оскільки цю величину зазвичай вже виміряно. Такий розподіл можна подати у вигляді таблиці або графіка, моделюючи функціональні рівняння (Рис. 3.33).



Рис. 3.33. Частотний розподіл

Застосовуючи методи математичної статистики, дуже важливо знати закон розподілу досліджуваної характеристики чи ознаки. Будь-яка статистична обробка даних починається зі спроби визначити закон розподілу. Залежно від того, яким є закон розподілу для аналізу даних використовують різні статистичні методи, а у випадку, коли методи застосовують для розподілів, що відрізняються від гіпотетичного, отримують хибні результати та помилкові висновки.

Критерії перевірки гіпотез щодо законів розподілу називають критеріями згоди, які в свою чергу поділяють на загальні та спеціальні. Загальні використовують незалежно від розподілу, а спеціальні – у разі перевірки гіпотези щодо конкретного розподілу: нормального,

рівномірного, експоненціального та інших.

Розрахунки емпіричного розподілу та його графічна візуалізація не дають надійних підстав для висновку щодо закону розподілу ознаки у сукупності, якої взято вибірку. Тим часом знання цього закону є необхідною умовою використання багатьох математичних методів. Серед методів оцінювання законів розподілу ймовірностей випадкових величин було спеціально розроблено близько двох десятків методів для перевірки нормальності розподілу. Найбільш розповсюдженими вважаються:

2. Критерії асиметрії й ексцесу.
3. Хі-квадрат ( $\chi^2$ ).
4. Критерій Шапіро-Вілка та інші.

Критерій асиметрії та ексцесу застосовують для приблизної перевірки гіпотези про нормальність емпіричного розподілу.

Асиметрія розподілу балів є додатною, якщо основна частина значень індивідуальних балів знаходиться праворуч від середнього значення, що зазвичай характерно для надто легких завдань.

Асиметрія розподілу балів є від'ємною, якщо більшість тестованих одержали оцінки нижчі від середнього бала.

Ефект від'ємної асиметрії діагностується у випадку занадто складних тестів, які неправильно збалансовані за складністю при доборі завдань. У добре збалансованому за складністю тесті розподіл балів має вигляд нормального розподілу. Для нього характерною є нульова асиметрія, що зрозуміло, оскільки за повної симетрії кожне значення бала, меншого ніж середнє значення, урівноважується іншим симетричним балами, що перевіряють середній бал.

За допомогою ексцесу можна отримати уявлення про те, чи є функція розподілу частот гостроверхою, середньоверхою чи плоскою. (Ексцес – міра кривизни розподілу одиниць сукупності тестованих за значенням ознаки (бали тестів). Характеризує величину відхилення верхівки фактичного розподілу від верхівки нормального розподілу вгору (гостроверхий) чи донизу (плосковерхий).

Критерій Шапіро-Вілка використовується для перевірки гіпотези  $H_0$ : “випадкова величина  $X$  має нормальний розподіл” і є одним з найефективніших критеріїв перевірки нормальності. Критерії, за допомогою яких перевіряють нормальність вибірки, є окремим випадком критеріїв згоди. Якщо вибірка нормальна, можна далі застосовувати потужні параметричні критерії, наприклад критерій Фішера.

Критерій Шапіро - Вілка ґрунтується на оптимальній лінійній незміщеній оцінці дисперсії до її звичайної оцінки методом максимальної правдоподібності.

За допомогою MS Excel визначено асиметрію, ексцес вибірки та критерій Шапіро - Вілка. Який дорівнює 0,91. Отже, з похибкою 10%



можна вважати розподіл нормальним. Отримані результати подано на рис. 3.34.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Студент	Результат		Асиметрия	-0,11131								
2	1	40		Експес	0,908281								
3	2	39											
4	3	35		i	n-1+i	a <sub>n-1+i</sub>	X <sub>i</sub>	X <sub>n-1+i</sub>	bi		(Ebi)2	446	
5	4	34		1	20	0,4734	40	19	9,9414		S2	489	
6	5	33		2	19	0,3211	39	22	5,4587		W	0,91	
7	6	32		3	18	0,2565	35	26	2,3085		W(0,01)	0,86815	
8	7	31		4	17	0,2085	34	27	1,4595		W(0,05)	0,90468	
9	8	31		5	16	0,1686	33	27	1,0116		W(0,1)	0,920168	
10	9	31		6	15	0,1334	32	28	0,5336				
11	10	30		7	14	0,1013	31	28	0,3039		W<W(0,1)	Розподіл нормальний	
12	11	30		8	13	0,0711	31	30	0,0711				
13	12	30		9	12	0,0422	31	30	0,0422				
14	13	30		10	11	0,014	30	30	0				
15	14	28											
16	15	28											
17	16	27											
18	17	27											
19	18	26											
20	19	22											
21	20	19											
22													

Рис. 3.34. Результати обробки показників

Дискримінативність завдань тесту – це здатність окремих завдань методики розрізняти тестованих із високим або низьким тестовим результатом, а також суб'єктів з високою продуктивністю реальної (критеріальної) діяльності від тих, у котрих вона низька.

Коефіцієнт дискримінації може набувати значень від  $-1$  до  $+1$ . Близьке до  $1$  додатне значення означає, що конкретний пункт тесту надійно розподіляє тестованих, із високими та низькими балами. Близьке до  $1$  від'ємне значення свідчить про непридатність завдання, недостатній зв'язок часткового результату з кінцевим висновком.

Будь-яку відповідь тестованого можна подати у дихотомічній шкалі: 1) “правильно”, “так” –  $1$  бал; 2) “неправильно”, “ні” –  $0$  балів.

Сума таких балів за всіма завданнями тесту являє собою первинний “сирий” бал. Міра збігів успішності розв'язання окремих завдань і всього тесту у вибірці тестованих є безпосереднім показником дискримінативності завдань тесту, який обчислюється у вигляді точково-бісеріального коефіцієнта кореляції і називається коефіцієнтом дискримінативності (індекс дискримінативності). Отримані результати подані на рис. 3.35.

Складність завдань тесту – це характеристика, що відображає статистичний рівень її розв'язності в даній вибірці стандартизації. Зазвичай її оцінюють при порівнянні кількості тестованих, що правильно виконали завдання до загальної кількості тестованих. Цей показник змінюється в межах від  $0$  до  $1$ . Його значення збільшується зі збільшенням складності завдання. При розробленні критеріально-орієнтованих тестів показник складності використовується з метою конструювання тестів з бажаним рівнем складності. Для підвищення складності тесту до його складу включають більше завдань високого рівня складності, а для зменшення складності – легших завдань.



+1 або менше ніж -1, то кажуть, що сталася помилка в обчисленнях. Отримані результати подано на Рис. 3.37.

	A	B	C	D	E	F	G	H
1		тест 1	іспит	Xi-Xсер.	Yi-Yсер.	(Xi-Xсер.) <sup>2</sup>	(Yi-Yсер.) <sup>2</sup>	(Xi-Xсер.)*(Yi-Yсер.)
2	1	40	95	9,9	17,5	97,02	304,5	171,88
3	2	30	70	-0,1	-7,6	0,02	57,0	1,13
4	3	27	60	-3,2	-17,6	9,92	308,0	55,28
5	4	28	70	-2,2	-7,6	4,82	57,0	16,23
6	5	35	86	4,9	8,5	23,52	71,4	40,98
7	6	31	77	0,9	-0,5	0,72	0,3	-0,47
8	7	22	73	-8,2	-4,6	66,42	20,7	37,08
9	8	33	90	2,9	12,5	8,12	155,0	35,48
10	9	31	63	0,9	-14,6	0,72	211,7	-12,37
11	10	26	80	-4,2	2,5	17,22	6,0	-10,17
12	11	27	75	-3,2	-2,6	9,92	6,5	8,03
13	12	34	83	3,9	5,5	14,82	29,7	20,98
14	13	30	81	-0,1	3,5	0,02	11,9	-0,52
15	14	19	60	-11,2	-17,6	124,32	308,0	195,68
16	15	32	73	1,9	-4,6	3,42	20,7	-8,42
17	16	30	90	-0,1	12,5	0,02	155,0	-1,87
18	17	30	63	-0,1	-14,6	0,02	211,7	2,18
19	18	28	80	-2,2	2,5	4,82	6,0	-5,27
20	19	39	92	8,9	14,5	78,32	208,8	127,88
21	20	31	90	0,9	12,5	0,72	155,0	10,58
22	Сума	603	1551	0,0	0,0	464,55	2305,0	684,35
23	Середнє	30,2	77,6					
24	Кореляція	0,661349889						
25								
26		0,661349889						
27								

Рис. 3.37. Коефіцієнт кореляції

Тест можна вважати ефективним за його відповідності таким основним умовам: застосування шкали інтервалів, надійність, валідність, дискримінативність, наявність нормативних даних та їх стандартизованість.

Надійність як один із основних критеріїв забезпечує незалежність результатів тестів від дії випадкових чинників. Надійність характеризує точність тесту як вимірювального інструменту, стійкість його до дії перешкод (стану випробовуваних, їх ставлення до процедури тестування, випадковим помилок і т.п.). Якісний тест не можна створити без ретельного вивчення цього важливого аспекту вимірювання.

У випадку політомічних відповідей (частково правильні відповіді) для визначення надійності тесту використовується формула  $\alpha$ -Кронбаха. Коефіцієнт  $\alpha$ -Кронбаха – це узагальнена міра внутрішньої гомогенності завдань тесту. У цьому методі порівнюється розкид кожного елемента (тестового завдання) із загальним розкидом всієї шкали. Якщо розкид результатів за весь тест менший, ніж розкид результатів для кожного окремого тестового завдання, то це означає, що кожне окреме тестове завдання спрямоване на дослідження однієї і тієї самої ознаки, властивості або явища. Констатують, що коефіцієнт  $\alpha$  задає нижню межу коефіцієнта надійності. Наприклад, якщо отриманий коефіцієнт  $\alpha$  дорівнює 0,8, то можна стверджувати, що 80% дисперсії отриманих результатів є наслідком дисперсії істинних результатів тестованих. Значення критерію  $\alpha$ -Кронбаха наведено в табл. 3.12.



### СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. About the system distantsiynogo navchannya "Virtualny Universitet" [Електронний ресурс]. Режим доступу: <http://vu.net.ua>. Назва з екрана.
2. Cattell R.B., Warbuiton F.W. Objective Personality and Motivation Tests. Urbana: University of Illinois Press, 1967.
3. Easy Test Maker [Електронний ресурс]. Режим доступу: <https://www.easytestmaker.com>. Назва з екрана.
4. IBM Lotus Workplace Collaborative Learning [Електронний ресурс]. Режим доступу: <http://icc.mpei.ru/documents/00000817.pdf>. Назва з екрана.
5. ISO/IEC 9126-1:2001 9126-1:2001 [Електронний ресурс]. Режим доступу: [http://www.iso.org/iso/catalogue\\_detail.htm?csnumber=22749](http://www.iso.org/iso/catalogue_detail.htm?csnumber=22749). Назва з екрана.
6. Miraxtest [Електронний ресурс]. Режим доступу: <http://www.mirax-software.ru>. Назва з екрана.
7. MyTest [Електронний ресурс]. Режим доступу: <http://mytest.klyaksa.net>. Назва з екрана.
8. Open Source e-Testing [Електронний ресурс]. Режим доступу: <http://www.tao.lu>. Назва з екрана.
9. PDF-tools [Електронний ресурс]. Режим доступу: <http://www.pdf-tools.com>. Назва з екрана.
10. TCExam is a FLOSS Computer-Based Assessment system [Електронний ресурс]. Режим доступу: <http://www.tcexam.org>. Назва з екрана.
11. Web-class KPI [Електронний ресурс]. Режим доступу: <http://dl.kpi.kharkov.ua>. Назва з екрана.
12. Weiss D.J. Computerized Adaptive Testing for Effective and Efficient Measurement in Counseling and Education. Measurement and Evaluation in Counseling and Development. 2004. Vol. 37. P. 70–84.
13. ХАМРР [Електронний ресурс]. Режим доступу: <http://uk.wikipedia.org/wiki/ХАМРР>. Назва з екрана.
14. ХАМРР for windows [Електронний ресурс]. Режим доступу: <https://www.apachefriends.org/download.html>. Назва з екрана.
15. Аванесов В.С. Композиция тестовых заданий. Москва: Центр тестирования, 2002. 240 с.
16. Аванесов В.С. Математические модели педагогического измерения. Москва: Исслед. Центр, 1994. 26 с.
17. Аналіз критеріїв педагогічного тесту [Електронний ресурс]. Режим доступу: <http://zmyo.npu.edu.ua/quiz/index.shtml>. Назва з екрана.
18. Беспалько В.П. Слагаемые педагогической технологии. Москва: Просвещение, 1989. 192 с.
19. Биков В.Ю., Руденко В.Д. Системи управління інформаційними базами даних в освіті. Київ: ІЗМН, 1996. 288 с.

20. Блэк Дж. Экономика. Толковый словарь. Общая редакция: д.э.н. Осадчая И.М. Москва: Весь Мир, 2000. 848 с.
21. Бойко С.М. Розробка та впровадження комп'ютерного тестування у навчальний процес в аграрних ЗВО 1–2 рівнів акредитації [Електронний ресурс]. Режим доступу: <http://int-konf.org/konf022014/726-boyko-s-m-rozrobka-ta-vprovadzheniya-kompyuternogo-testuvannya-u-navchalniy-proces-v-agrarnih-vnz-rvnyv-akreditacyi.html>. Назва з екрана.
22. Бронетко В.О. Системи комп'ютерного тестування: огляд, аналіз, порівняння. Зб. наук. праць Кам'янець-Подільського нац. Ун-ту ім. Івана Огієнка. Серія педагогічна. Кам'янець-Подільський, 2009. 352 с.
23. Булах І.Є., Мруга М.Р. Створюємо якісний тест. Навчальний посібник. Київ: Майстер-клас, 2006. 160 с.
24. Булах І.Є. Теорія і методика комп'ютерного тестування успішності навчання (на матеріалах медичних навчальних закладів): дис... доктора пед. наук: 13.00.01, Київський національний університет імені Т.Г. Шевченка. К., 1995. 430 с.
25. Владимиров В.Н., Урусов Н.А. О возможностях компьютеризованного тестового контроля. Компьютер и историческое знание. Барнаул, 1994. С.177–183.
26. Выготский Л.С. Диагностика развития и педологическая клиника трудного детства: Собр. соч. Москва, 1983. С. 257–321.
27. Гастев Ю.А. Модель. Философская энциклопедия. Москва, 1964. 634 с.
28. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике. Москва: Высшая школа, 1998. 400 с.
29. Гончаренко С.У. Педагогічний словник. Довідкове видання. Київ: Либідь, 1997. 366 с.
30. Державна національна програма "Освіта" (Україна ХХІ століття)
31. Дуглас У. Тестирование умений и навыков: основне. Директор школы. 1994. № 6. С. 14–22.
32. Закон України "Про Національну програму інформатизації." Київ: Офіц. вид., Парламентське видавництво, 2002. 20 с.
33. Закон України "Про вищу освіту", Законодавство України про освіту. Збірник законів. Київ: Парламентське видавництво. 2002.
34. Ковальчук Ю.О. Теорія освітніх вимірювань. Ніжин: Видавець ПП Лисенко М.М., 2012. 200 с.
35. Великая дидактика. Приводится по изданию: Я.А. Коменский, Д. Локк, Ж. Ж. Руссо, И.Г. Песталоцци. Педагогическое наследие. Москва: Педагогика, 1989. 416 с.
36. Комп'ютерне тестування – інноваційний метод контролю знань навчальних досягнень школярів [Електронний ресурс]. Режим доступу: <http://www.monitoring.in.ua/publications/view/7>. Назва з екрана.

37. Конструктор тестов easyQuizzy [Електронний ресурс]. Режим доступу: <http://easyquizzy.ru>. Назва з екрана.
38. Корнетов Г.Б. Становление и развитие воспитания в доклассовом обществе: автореф. дис... канд. пед. наук. Москва, 1986. 19 с.
39. Котяк В.В. Оцінка якості веб-орієнтованих систем тестування навчальних досягнень. Наук. часопис, Інформаційні технології в освіті. Кіровоград: 2013. № 14. С. 120–126.
40. Кривошеев А.О. Разработка и использование компьютерных обучающих программ. Информационные технологии. 1996. №4. С. 45.
41. Крившенко Л.П. Педагогика: Учеб., под ред. Л.П. Крившенко. Москва: ТК Велби, 2008. 432 с.
42. Кудін А.П. Інформаційно-комунікаційні технології в навчанні: Навч. посіб. Київ: Волиньполіграф, 2012. 414 с.
43. Курлянд З.Н., Хмельюк Р.І., Семенова А.В. Педагогіка вищої школи: Навч. посіб., 2-ге вид., перероб. і доп. Київ: Знання, 2005. 399 с.
44. Кухар Л.О., Сергієнко В.П. Конструювання тестів. Курс лекцій: Навч. Посіб. Луцьк, 2010. 182 с.
45. Лендюк Т.В. Моделювання комп'ютерного адаптивного навчання і тестування. Праці Одеського політехнічного університету. 2013. Вип. 1(40). С. 112.
46. Лисенко В.Ф. Вимірювання в освіті: Підруч; за ред. О.В. Авраменко. Кіровоград, 2011. 360 с.
47. Лісова Т.В. Моделі та методи сучасної теорії тестів. Навчально-методичний посібник. Ніжин: Видавець ПП Лисенко М.М., 2012. 112 с.
48. Лодатко Є.О. Педагогічні моделі, педагогічне моделювання і педагогічні вимірювання: that is that?. Педагогіка вищої школи: методологія, теорія, технології. Вища освіта України: Теоретичний та науково-методичний часопис. 2011. Вип 3. С. 339–344.
49. Мархель И.И. Компьютерная технология обучения. Среднее специальное образование. 1989. С. 27.
50. Машбиц Е.И., Андриевская В.В., Комиссаров Е.Ю. Диалог в обучающей системе. Київ: Вища школа, 1989. 184 с.
51. Микитенко П.В., Сергієнко В.П., Кухар Л.О. Використання вбудованої системи аналізу тестових завдань в LCMS MOODLE. Інформаційні технології і засоби навчання. 2014. № 3. С. 196–208.
52. Микитенко П.В., Сергієнко В.П. Комп'ютерно орієнтовані технології освітніх вимірювань як педагогічна проблема. Збірник наукових праць Кам'янець-Подільського національного університету ім. Івана Огієнка. Сер. Педагогічна. 2014. № 20. С. 297–301.
53. Микитенко П.В., Сергієнко В.П., Франчук В.М. Методичні рекомендації зі створення тестових завдань та тестів у системі управління

навчальними матеріалами MOODLE. Київ: Вид-во НПУ імені М.П. Драгоманова, 2014. 100 с.

54. Микитенко П.В. Педагогічна модель системи комп'ютерно орієнтованих технологій освітніх вимірювань. Комп'ютерно-орієнтовані системи навчання природничо-математичних дисциплін: матеріали Міжнародного науково-практичного семінару. Київ, 2014. С. 101–102.

55. Микитенко П.В. Реалізація адаптивного тестування засобами комп'ютерно орієнтованих технологій освітніх вимірювань. Науковий часопис Національного педагогічного університету імені М.П. Драгоманова. Педагогічні науки: реалії та перспективи. Київ: Вид-во НПУ імені М.П. Драгоманова, 2013. С. 141–148.

56. Микитенко П.В., Сергієнко В.П. Сутність і специфіка методів та методик комп'ютерної педагогічної діагностики знань. Вища освіта України. Тематичний випуск "Педагогіка вищої школи: методологія, теорія, технології". Київ; Кіровоград, 2014. № 3. С. 20–24.

57. Морзе Н.В. Методика навчання технологій. Навч. Посіб., у 3-х ч., за ред. М. І. Жалдака. Загальна методика навчання технологій. Київ, 2004. 256 с.

58. Подготовка и проведение учебных курсов в заочно-дистанционной форме обучения. Методические рекомендации преподавателям, под ред. И.А. Цикина. С-Пб.: СПбГТУ, 2000. 126 с.

59. Подласый И.П. Педагогика: 100 вопросов – 100 ответов. Учеб. пособ. для вузов. Москва: ВЛАДОС-пресс, 2004. 365 с.

60. Подласый И.П. Педагогика: Новый курс. В 2 кн. Общие основы. Москва: Гуманит, 2002. 576 с.

61. Програма тестирования OpenTEST [Електронний ресурс]. Режим доступу: <http://opentest.com.ua>. Назва з екрана.

62. Програмна оболонка для створення тестових завдань Hot Potatoes [Електронний ресурс]. Режим доступу: <http://hotpot.uvic.ca>.

63. Раков С.А. Формування математичних компетентностей учителя математики на основі дослідницького підходу у навчанні з використанням інформаційних технологій: дис... докт. пед наук: 13.00.02, ХНПУ імені Г.С. Сковороди. Харків, 2005. 538 с.

64. Романов А.В. Методика подготовки и проведения тестового контроля в учебном процессе. Чебоксары: КЛИО, 1988. С.11–13.

65. Сергієнко В.П., Матвійчук О.В., Пустовий О.М. Тестові завдання з курсу загальної фізики. Навч. посіб. для студентів вищих навчальних закладів. Луцьк: 2010. 70 с.

66. Сергієнко В.П., Кухар Л.О. Методичні рекомендації зі складання тестових завдань. Київ: НПУ ім М.П. Драгоманова, 2011. 41 с.

67. Сергієнко В.П., Франчук В.М., Микитенко П.В. Використання інформаційно-комунікаційних технологій управління якістю освіти в



- педагогічному університеті. Інформатика та інформаційні технології. 2012. № 4 . С. 56.
68. Сліпчук І.Ю. Методика навчання біології учнів 8-9 класів з використанням комп'ютерних технологій: дис... канд.пед.наук: 13.00.02. Київ, 2008. 239 с.
69. Третьяков П.И. Школа: управление по результатам: практика педагогического менеджмента. Москва: Нова школа, 2001. 320 с.
70. Федорук П.І. Адаптивні тести: статистичні методи аналізу результатів тестового контролю знань. Математичні машини і системи. 2007. № 3. 380 с.
71. Янченко О.І. Форми і методи контролю знань в умовах сучасних навчальних технологій: Методична розробка. Кривий Ріг: 2008. 29 с.
72. Ярощук Л.Г. Основи педагогічних вимірювань та моніторингу якості освіти: навч. посіб. Київ: Видавничий Дім "Слово", 2010. 304 с.
73. Яшанов С.М., Смекалін І.М. Комп'ютерне тестування: Навч.-метод. посіб. для вищ. пед. закл. освіти. Київ: Вид-во НПУ імені М.П. Драгоманова, 2008. 84 с.

## ДОДАТКИ

**Додаток 1**  
**Порівняння комп'ютерно-орієнтованих технологій освітніх**  
**вимірювань**

Характеристика	Назва								
	TCEam	TAO	Open Test2	MOODLE	MyTest	Hot Potatoes	Mirax Test	EasyTest	Easy Quizzy
Стандартні типи тестових завдань	1	1	1	1	1	1	1	1	1
Додаткові типи тестових завдань	0	2	0	2	2	2	1	1	2
Банк тестових завдань	2	1	2	2	2	1	1	2	1
Обмеження в часі на проходження тесту	1	1	1	1	1	1	1	1	1
Вагові коефіцієнти для тестових завдань	2	1	1	2	1	1	1	1	1
Підтримка різних шкал оцінювання	2	2	2	2	2	0	1	1	2
Облікові записи користувачів та груп	2	1	1	2	1	1	1	1	2
Експорт та імпорт даних про користувачів	2	2	2	2	2	0	0	0	0
Експорт та імпорт тестових завдань	2	2	2	2	2	2	1	0	1
Експорт тесту для виконання на папері	1	1	1	1	1	0	1	1	1

Імпорт сканованих бланків тесту	1	0	0	0	0	0	0	0	0
Захист персональних даних	2	2	2	2	1	0	1	0	0
Підтримка мультимедійних даних	1	1	2	2	1	0	1	0	2
Засоби аналізу тесту та тестових завдань	1	1	1	2	1	1	0	0	1
Створення резервних копій	1	1	1	1	1	1	1	0	1
Простота використання	1	1	2	2	1	0	1	1	1
Простота модифікації чи зміни функціоналу	1	0	0	1	0	0	1	0	1
Простота встановлення	1	1	1	0	1	1	1	0	1
Зрозумілість	2	2	2	2	2	1	2	1	2
<b>Всього балів</b>	<b>26</b>	<b>23</b>	<b>24</b>	<b>31</b>	<b>22</b>	<b>13</b>	<b>17</b>	<b>11</b>	<b>21</b>
<b>Рейтинг</b>	<b>II</b>	<b>IV</b>	<b>III</b>	<b>I</b>	<b>IV</b>	<b>VIII</b>	<b>VII</b>	<b>IX</b>	<b>VI</b>

## Додаток 2

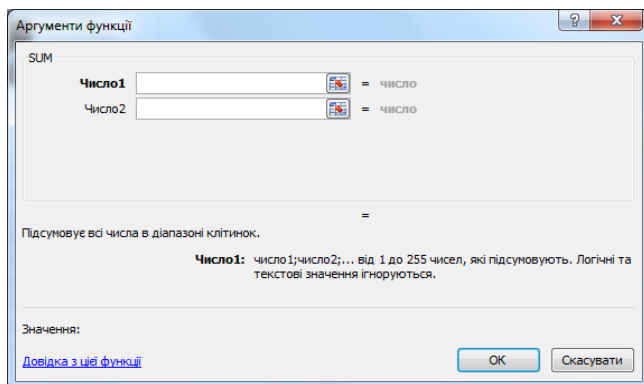
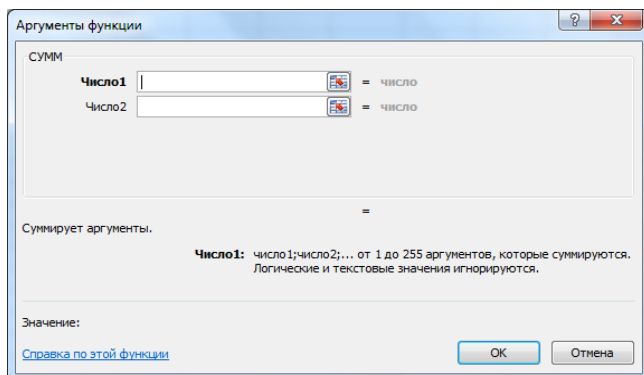
### Синтаксис функцій в середовищі MS Excel

#### СУМА

*Синтаксис функції:* **СУММ(число1,число2,...); SUM(число1,число2,...)**

*Результат:* Обчислює суму заданого масиву(діапазону).

*Аргументи:* Число1,число2, ... – від 2 до 255 аргументів для яких обчислюється сума.



## СЕРЕДНЄ ЗНАЧЕННЯ

*Синтаксис функції:* **СРЗНАЧ**(число1; число2;...); **AVERAGE**(число1; число2;...)

*Результат:* Обчислює середнє значення ряду (масиву).

*Аргументи:* Число1, число2, ... від 1 до 255 чисел для яких обчислюється середнє значення.

Аргументы функции

---

СРЗНАЧ

**Число1**  = число

**Число2**  = число

=

Возвращает среднее арифметическое своих аргументов, которые могут быть числами, именами, массивами или ссылками на ячейки с числами.

**Число1:** число1;число2;... от 1 до 255 числовых аргументов, для которых вычисляется среднее.

---

Значение:

[Справка по этой функции](#)

Аргументи функції

---

AVERAGE

**Значення1**  = число

**Значення2**  = число

=

Повертає середнє (арифметичнє) аргументів, які можуть бути числами або іменами, масивами або посиланнями на клітинки з числами.

**Значення1:** значення1;значення2;... від 1 до 255 аргументів, для яких необхідно обчислити середнє.

---

Значення:

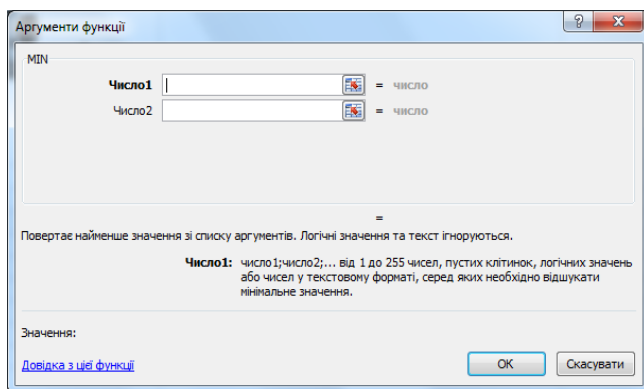
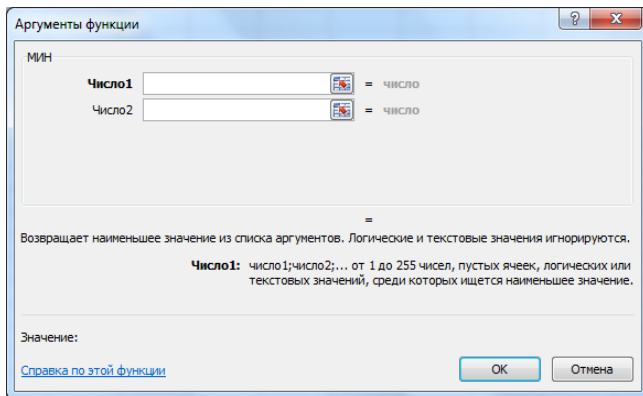
[Довідка з цієї функції](#)

## МІНІМУМ

*Синтаксис функції:* **МИН(число1, число2,...); MIN(число1, число2,...)**

*Результат:* Обчислює найменше значення серед ряду розподілу.

*Аргументи:* Число1, число2,... – від 1 до 255 чисел, серед яких порівнювати знайти найменше.



## МАКСИМУМ

*Синтаксис функції:* **МАКС(число1, число2,...); МАХ(число1, число2,...)**

*Результат:* Обчислює найбільше значення серед ряду розподілу.

*Аргументи:* Число1, число2,... – від 1 до 255 чисел, серед яких порівнювати найбільше.

Аргументы функции

МАКС

Число1  = число

Число2  = число

=

Возвращает наибольшее значение из списка аргументов. Логические и текстовые значения игнорируются.

**Число1:** число1;число2;... от 1 до 255 чисел, пустых ячеек, логических или текстовых значений, среди которых ищется наибольшее значение.

Значение:

[Справка по этой функции](#) OK Отмена

Аргументы функції

МАХ

Число1  = число

Число2  = число

=

Повертає найбільше значення зі списку аргументів. Логічні значення та текст ігноруються.

**Число1:** число1;число2;... від 1 до 255 чисел, пустих клітинок, логічних значень або чисел у текстовому форматі, серед яких необхідно відшукати максимальне значення.

Значення:

[Довідка з цієї функції](#) OK Скасувати

## КВАРТИЛЬ

*Синтаксис*

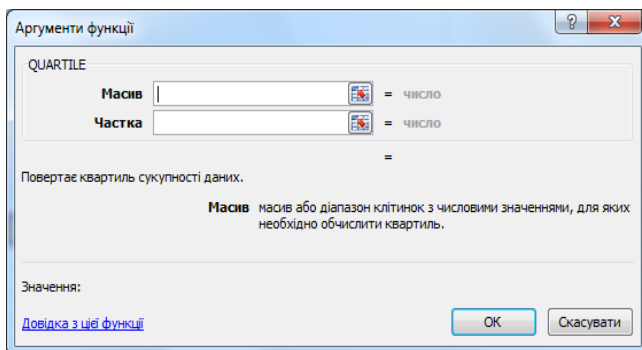
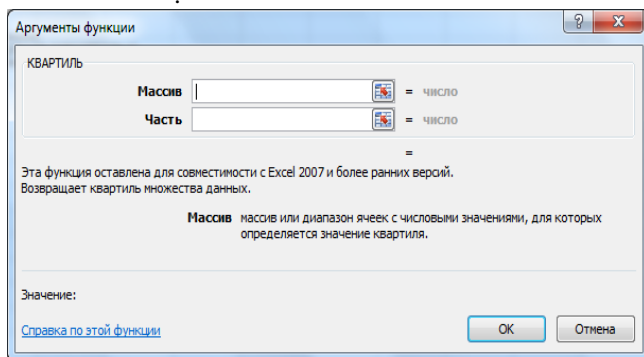
*функції:* **КВАРТИЛЬ**(масив,часть);

**QUARTILE**(масив,часть)

*Результат:* Квартиль заданого масиву (діапазону чисел). (Квартиль –одна з числових характеристик випадкових величин, за допомогою якої ми відокремлюємо частину ряду значень. Наприклад, квантиль 0.25 (також називається 25-м процентилям або нижнім квартилем) змінної — це таке значення, що 25% значень змінної попадають нижче даного значення).

*Аргументи:* **Масив** – діапазон комірок з числовими значеннями, для яких обчислюється значення квартиля.

**Часть** – значення, яке потрібно обчислити



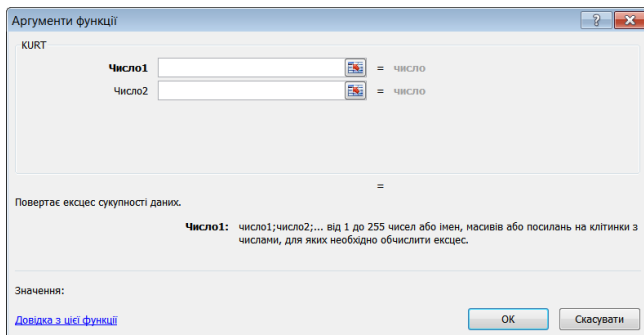
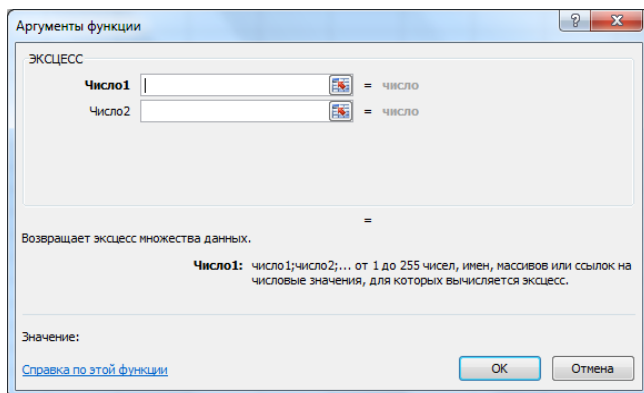


## ЕКСЦЕС

*Синтаксис функції:* **ЭКЦЕСС(число1,число2,...)**

*Результат:* Ексцес заданої множини (масиву) чисел. (Ексцес – числова характеристика розподілу ймовірностей дійсної випадкової величини. Коефіцієнт ексцесу характеризує «крутість», тобто, стрімкість підвищення кривої розподілу у порівнянні з нормальною кривою)

*Аргументи:* Число1, число2, ... – від 1 до 255 аргументів, для яких обчислюється ексцес. У функції ексцес можна використовувати крім чисел, ще й масиви або посилання на них.

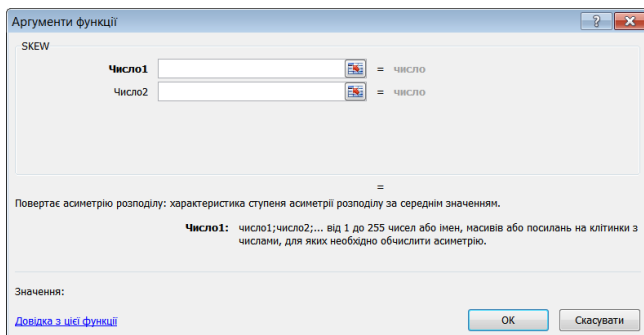
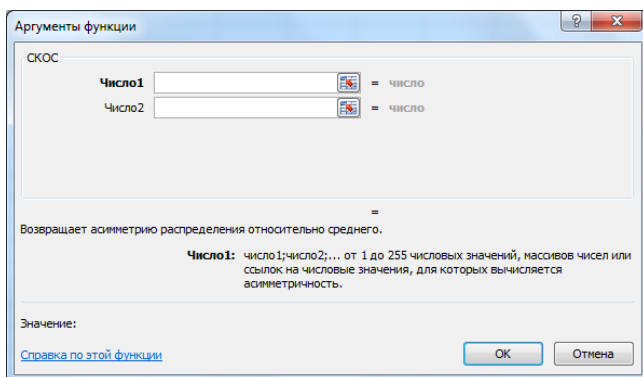


## АСИМЕТРИЯ

**Синтаксис функції:** *СКОС*(число1, число2,...)

**Результат:** Асиметрія заданої множини (масиву) чисел. (Асиметрія - це зміщення від кривої нормального розподілу. При зміщенні вправо від центра асиметрія буде характеризуватися додатнім числом, при зміщенні вліво - від'ємним).

**Аргументи:** Число1, число2,... – від 1 до 255 аргументів, для яких розраховується асиметрія. У функції асиметрія можна використовувати крім чисел, ще й масиви або посилання на них.



## РАНГ

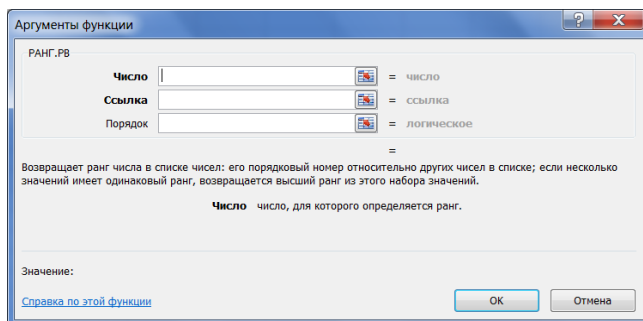
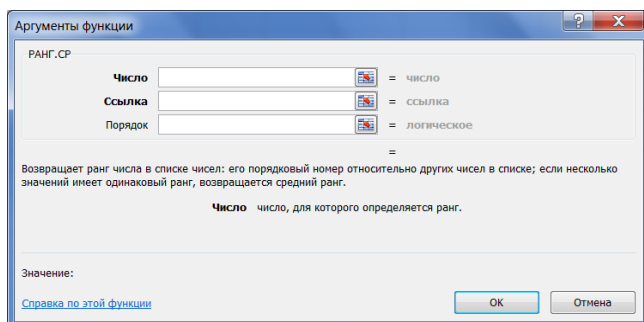
**Синтаксис функції: РАНГ (число; ссылка; порядок)**

**Результат:** Ранг числа в списке чисел. Ранг числа - це показник його величини відносно інших значень в списку. (Якщо список відсортувати, то ранг числа буде його позицією)

**Аргументи:** **число** – посилання на комірку, для якої визначається ранг;

**ссылка** – посилання на діапазон індивідуальних балів (вибірка);

**порядок** – число, яке вказує на спосіб упорядкування. Якщо порядок дорівнює 0 (нулю), або не вказаний, то програма визначає ранг числа так, якщо б посилання було списком, відсортованим за спаданням. Якщо порядок – будь яке ненульове значення, то ранг числа визначається так, якщо б посилання було списком, відсортованим за зростанням.

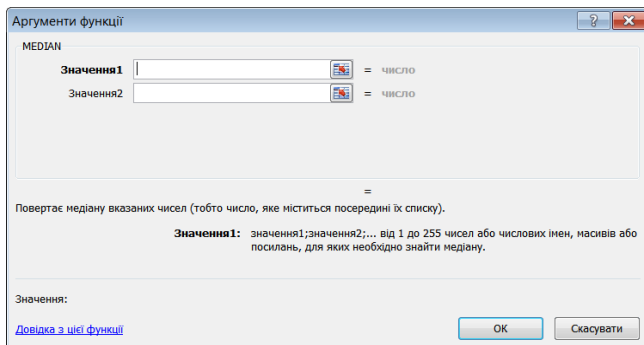
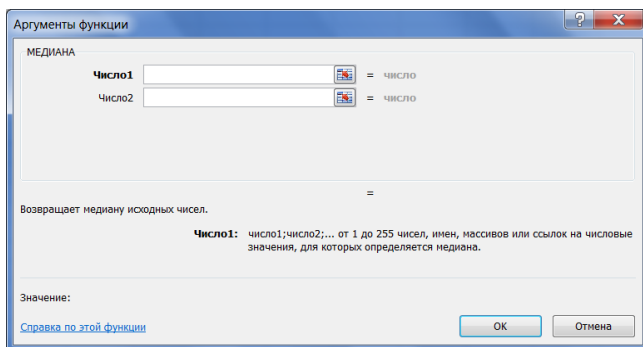


## МЕДІАНА

*Синтаксис функції:* **МЕДІАНА** (число1, число2, ...)

*Результат:* Медіана заданої множини чисел (число, яке є серединою множини чисел: половина чисел більша, ніж медіана, а половина чисел менша, ніж медіана).

*Аргументи:* число1, число2, ... – числа чи імена, масиви або адресні посилання на діапазон комірок, що містить посилання.

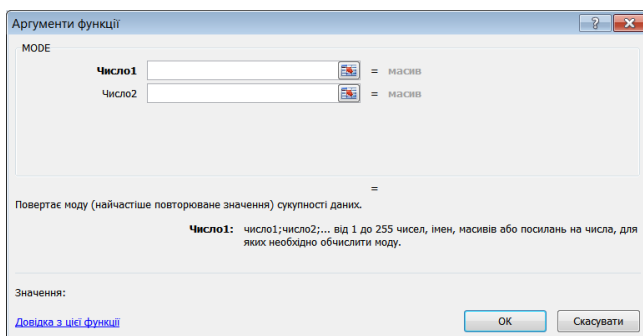
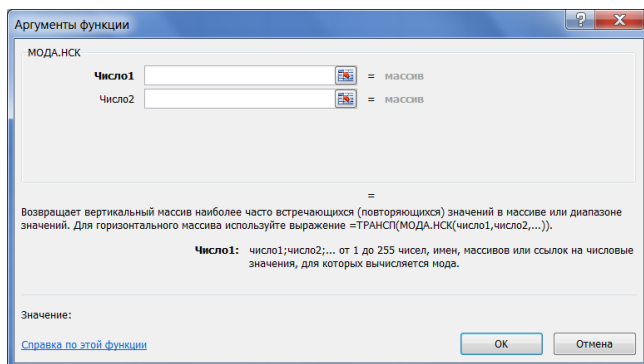


## МОДА

**Синтаксис функції:** *МОДА* (*число1*, *число2*, ...)

**Результат:** Значення масиву, які найчастіше зустрічаються в масиві чи інтервалі даних.

**Аргументи:** *число1*, *число2*, ... – від 1 до 30 аргументів, для яких обчислюється функція *МОДА*; в функції *МОДА* можна використовувати замість аргументів масив або посилання на масив.

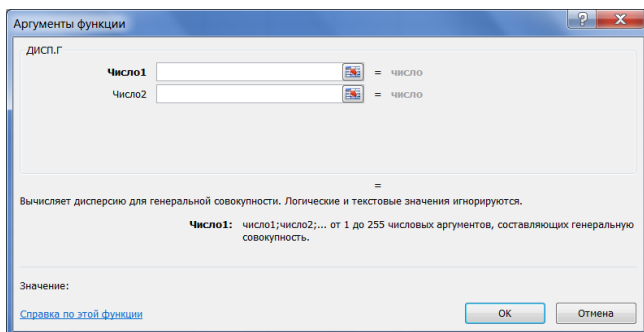
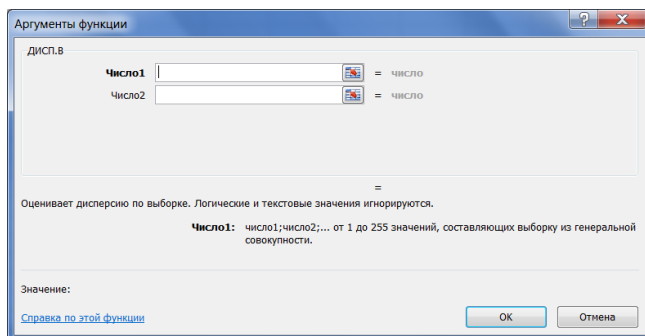


## ДИСПЕРСІЯ

*Синтаксис функції:* **ДИСП (число1,число2, ...)**

*Результат:* Дисперсія генеральної сукупності. Аргументи представляють усю генеральну сукупність.

*Аргументи:* число1, число2, ... – не більше 30 аргументів; текстові, логічні та порожні поля приводять до помилки.

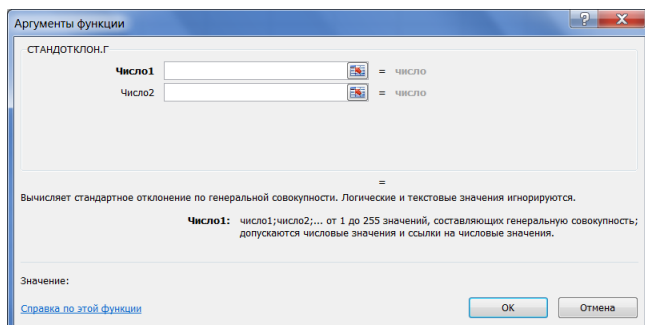
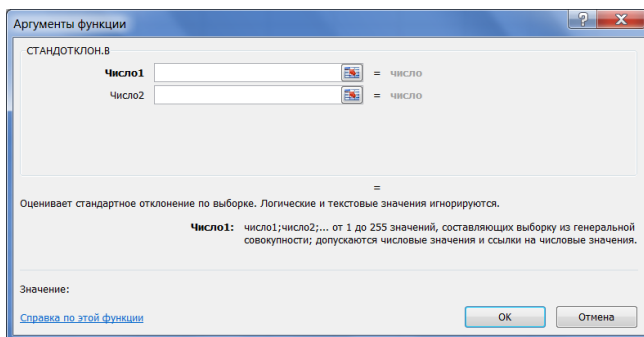


## СТАНДАРТНЕ ВІДХИЛЕННЯ

*Синтаксис функції:* **СТАНДОТКЛОН**(число1, число2, ...)

*Результат:* Оцінка стандартного відхилення вибірки. Стандартне відхилення – це міра того, наскільки широкий розкид точки даних відносно їх середнього.

*Аргументи:* число1, число2, ... – від 1 до 30 числових аргументів, які відповідають вибірці з генеральної сукупності.



Наукове видання

Микитенко Павло Васильович  
Галицький Олександр Вадимович  
Кучеренко Інна Іванівна

КОМП'ЮТЕРНО-ОРІЄНТОВАНІ ТЕХНОЛОГІЇ ОСВІТНІХ  
ВИМІРЮВАНЬ

Київ, Науково-виробниче підприємство  
«Видавництво “Наукова думка” НАН України», 2019

Підп. До друку 02.04.2019. Формат 60×90/16.  
Папір. офс. № 1. Гарн. Таймс. Друк. офс.

Ум. друк. арк. 00,00. Ум. фарбо-відб. 00,0. Обл.-вид. арк. 00,0

Тираж 300 прим., Зам. № 00-000