

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ, МОЛОДІ ТА СПОРТУ УКРАЇНИ
НАЦІОНАЛЬНИЙ ПЕДАГОГІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ М.П. ДРАГОМАНОВА**

Сергієнко В.П., Малежик М.П., Сіткар Т.В.

Комп'ютерні технології в тестуванні

Навчальний посібник

для студентів вищих педагогічних навчальних закладів

Київ – 2012

УДК 37.091.26:004(075.8)

ББК 74.002.8:32.97я73

С32

Рецензенти:

Рамський Ю.С. – кандидат фізико-математических наук, професор, завідувач кафедри програмування та інформаційних технологій НПУ імені М.П. Драгоманова;

Маслянюк П.П. – кандидат технічних наук, доцент кафедри прикладної математики НТУУ „КПІ”

Сергієнко В.П., Малежик М.П., Сіткар Т.В. **Комп’ютерні технології в тестуванні: навч. посіб.** – Луцьк: СПД Гадяк Жанна Володимирівна, друкарня «Волиньполіграф»™, 2012. – 290 с.

У навчальному посібнику розкрито основи педагогічних вимірювань та моніторингу якості освіти, а також історія становлення та розвитку тестування, роль міжнародного тестування. Систематизовані основні принципи, правила складання тестових завдань. Особливе місце відведено опрацюванню результатів тестування, особливостям застосування комп’ютерних технологій у тестуванні.

Розроблені тестові завдання створюють умови для здійснення об’єктивного контролю та самоконтролю знань.

Методичний комплект (посібник, презентації, праці відомих учених, які допоможуть підготуватися до занять) подано і в електронному варіанті.

Посібник призначений для студентів вищих навчальних закладів, викладачів, учителів-практиків, тих, хто цікавиться педагогічними вимірюваннями.

Друкується за рішенням вченої ради інституту інформатики
Національного педагогічного університету імені М.П. Драгоманова
(протокол № 9 від 20 червня 2012 р.)

ЗМІСТ

ПЕРЕДМОВА.....	8
РОЗДІЛ I. ІНФОРМАТИЗАЦІЯ ОСВІТИ. НЕОБХІДНІ УМОВИ ДЛЯ СТВОРЕННЯ ІНФОРМАТИЗАЦІЇ СУСПІЛЬСТВА.....	9
РОЗДІЛ II. ТИПОЛОГІЯ ПРОГРАМНИХ ЗАСОБІВ НАВЧАЛЬНОГО ПРИЗНАЧЕННЯ.....	19
2.1 Типи програмних засобів навчального призначення.....	19
2.2 Програмні засоби, що реалізують ідеї програмованого навчання.....	20
2.3 Програмні засоби, що моделюють досліджуваний об'єкт, процес, експеримент або явище.....	21
2.4 Ігрові компоненти в програмних засобах.....	22
2.5 Дидактична доцільність використання ПЗ навчального призначення різних типів.....	22
2.6 Типологія ПЗ за функціональним призначенням.....	23
2.7 Методичні цілі використання ПЗ навчального призначення.....	26
2.8 Типологія ПЗ за методичним призначенням.....	27
РОЗДІЛ III. ПЕДАГОГІЧНЕ ТА ПСИХОЛОГІЧНЕ ТЕСТУВАННЯ.....	37
3.1 Переваги та недоліки тестування.....	37
3.2 Педагогічне та психологічне тестування.....	41
3.3 Бланкове тестування.....	42
3.4 Переваги комп'ютерного тестування.....	46
3.5 Загальні переваги і недоліки тестових технологій.....	48
3.6 Культура та філософія адаптивного тестування.....	51
РОЗДІЛ IV. ТЕХНОЛОГІЇ МАСОВОЇ КОМП'ЮТЕРНОЇ АТЕСТАЦІЇ ЗУН. ЦІЛІ, ЗАВДАННЯ, ХАРАКТЕРИСТИКИ, ЕТАПИ, ШЛЯХИ, РЕЗУЛЬТАТИ..	56
4.1 Мета і завдання педагогічного тестового контролю.....	56
4.2 Функцій ОІТ контролю навчальної діяльності.....	59
4.3 Види, типи, етапи, вимоги, характеристики комп'ютерного тестування.....	60
4.4 Характеристики ОІТ атестації ЗУН.....	61

4.5	Загальні етапи комп'ютерної атестації	62
4.6	Шляхи досягнення мети комп'ютерної атестації.....	62
4.7	Типи сценаріїв комп'ютерного тестування	63
4.8	Запитання теоретику і практику	68
4.9	Загальні та часткові вимоги до комп'ютерного тестового комплексу	70
РОЗДІЛ V. НЕДОЛІКИ ІНТЕРФЕЙСУ КОМПЛЕКСІВ ТЕСТОВОГО КОНТРОЛЮ		73
5.1	Групи основних недоліків інтерфейсу.....	73
5.2	Зауваження користувачу.	74
5.3	Результати комп'ютерної атестації ЗУН і фактори впливу	75
5.4	Інтерпретації результатів тестувань	77
5.5	Тривалість сеансу тестування.....	79
5.6	Кількість завдань у тесті	82
5.7	Специфікація тесту	83
РОЗДІЛ VI. РЕЗУЛЬТАТИ ПЕДАГОГІЧНИХ ДОСЛІДЖЕНЬ.....		84
6.1	Артефакти	84
6.2	Динаміка результатів педагогічних вимірювань	90
6.3	ОІТ та менеджмент якості. Принцип безперервного покращення	10
4		
6.4	Основні вимоги до комп'ютерного комплекс для навчання і контролю ЗУН?	10
6		
РОЗДІЛ VII. ТЕСТОВІ ВИПРОБУВАННЯ ТА УПРАВЛІННЯ ОСВІТНІМИ РЕЗУЛЬТАТАМИ		10
8		

7.1	Види	мотивації	10
8			
7.2	Мотивація	учнів до підвищення якості ЗУН	10
9			
7.3	Технологія	м'якого тестування	11
0			
7.4	Мотивація	студента до підвищення якості ЗУН	11
6			
7.5	Привабливі	риси процедури комп'ютерного тестування	11
7			
	РОЗДІЛ VIII. ІНФОРМАЦІЙНА БЕЗПЕКА ПРОЦЕДУР МАСОВОГО КОМП'ЮТЕРНОГО ТЕСТУВАННЯ		12
1			
8.1	Основні вимоги інформаційної безпеки тестувань та забезпечення їх виконання		12
1			
8.2	Цікаві явища з досвіду масових комп'ютерних тестувань		12
2			
8.3	Шляхи забезпечення інформаційної безпеки масових комп'ютерних тестувань		12
3			

8.4	Канали витоку конфіденційної інформації. Забезпечення секретності БТЗ	12
6		
8.5	Стимули збереження конфіденційності тестових матеріалів у США	12
8		
8.6	Стимули збереження конфіденційності тестових матеріалів у Росії	13
0		
8.7	Забезпечення конфіденційності тесту в масштабі регіону і країни	13
1		
8.8	Принципи безпечної технології створення БТЗ	13
1		
8.9	Десять кроків безпечної технології тестування	13
3		
8.10	Безпечна технологія обробки результатів і публікацій	13
5		
8.11	Напрями зниження загальної собівартості та собівартості заходів безпеки	13
6		
8.12	Способи фальсифікації, які відомі в практиці тестування	13
7		

РОЗДІЛ	ІХ.	КІЛЬКІСТЬ	ТА	ЯКІСТЬ	ЗУН	
.....						13
9						
9.1	Іспит	як	форма	виміру	ЗУН	
.....						13
9						
9.2		Сепарабельність			оцінки.	
.....						14
0						
9.3	Що	вимірюють	у	процесі	тестування?	
.....						14
2						
9.4	Як	виміряти		кількість	ЗУН?	
.....						14
7						
9.5	Як	виміряти		якість	ЗУН?	
.....						14
8						
9.6	Як	оцінити	якість	оцінювання	якості	ЗУН?
.....						14
9						
9.7	Тестування,	як	засіб	об'єктивізації	оцінки	якості та кількості.
.....						15
1						
9.8	Роздільна	здатність	і	об'єктивність	вимірювальних	процедур
.....						15
3						
9.9	Систематичні	помилки		в	тестуванні	
.....						15
6						

9.10 Рейтинг і оцінка. Їх визначення та призначення
.....15

8

9.11 Системи кумулятивного індексування
.....16

2

9.12 Визначення скалярного рейтингу
.....16

5

РОЗДІЛ X. ОІТ ДЛЯ ЗАХИСТУ ЯКОСТІ ТА УПРАВЛІННЯ ЯКІСТЮ ОСВІТИ

.....16
9

10.1 ОІТ і керування якістю.
.....16

9

10.2 ОІТ і організація управління якістю.
.....17

1

10.3 ОІТ і кадрове забезпечення
.....17

2

10.4 ОІТ і захист якості освіти
.....17

3

10.5 ОІТ і проблеми фінансування
.....17

6

РОЗДІЛ XI. ЗАСТОСУВАННЯ КОМП'ЮТЕРНИХ ТЕХНОЛОГІЙ МАСОВОГО ТЕСТУВАННЯ

.....	17
8	
11.1 Історичні передумови масового тестування	17
.....	17
8	
11.2 Масові тестування знань, умінь та навичок в Росії	18
.....	18
0	
11.3 Засоби профілактики негативних ситуацій при комп'ютерному тестуванні	18
.....	18
1	
11.4 Дидактичний принцип індуктивної мотивації в організації масових комп'ютерних тестувань	18
.....	18
3	
11.5 Моніторинг якості освіти в Росії. Міжнародне співробітництво	18
.....	18
5	
11.6 Організація та проведення дослідження TIMSS-R (1999) у Росії	18
.....	18
9	
11.7 Єдиний державний іспит у Росії.	19
.....	19
3	
11.8 Моніторинг, атестація, сертифікація. Досвід США	19
.....	19
6	

11.9	Розвиток	технологій	тестування	19			
8							
11.10	Адаптивне		тестування.	20			
1							
11.11	Незвичайні	форми	тестування	20			
3							
РОЗДІЛ	ХІІ.	ПРОГРАМИ	ДЛЯ	СТВОРЕННЯ	ТЕСТІВ	20	
6							
12.1		Quiz		Press	20		
6							
12.2		Wondershare		QuizCreator	20		
8							
12.3		Schoolhouse		Test	21		
4							
12.4		Adit		Testdesk	21		
7							
РОЗДІЛ	ХІІ.	ПРОБЛЕМИ	АДАПТИВНОГО	ТЕСТУВАННЯ.	22		
3							
РОЗДІЛ	ХІV.	ОСНОВНІ	ВІДОМОСТІ	З	ОБЛАСТІ	ТЕСТОЛОГІЇ	23
4							

14.1	Загальні	відомості	з	тестології	23
.....					
4					
14.2	Принципи конструювання тесту, створення тестових завдань і вибору технології			тестування	23
.....					
5					
14.3	Педагогічний			тест	24
.....					
4					
14.4	Класифікації та види комп'ютерних педагогічних тестів				24
.....					
7					
14.5	Недоліки		педагогічних		тестів
.....					
1					
14.6	Валідність		і		надійність
.....					
3					
14.7	Надійність тесту. Контрольовані та неконтрольовані фактори впливу на надійність				25
.....					
4					
14.8	Типи надійності тесту та способи її визначення				25
.....					
5					
14.9	Розвиток поняття		валідності		та
.....					
8					

14.10	Інформаційний	граф		
.....				26
1				
14.11	Типи помилок вимірювань в тестології та джерела їх виникнення			
.....				26
3				
14.12	Педагогічне	тестове	завдання	
.....				26
4				
14.13	Чотири	форми	тестових	завдань
.....				26
7				
14.14	Закрита	форма	тестового	завдання
.....				26
9				
14.15	Відкрита	форма	тестового	завдання
.....				27
1				
14.16	Завдання	на	встановлення	відповідності
.....				27
3				
14.17	Завдання	на	встановлення	послідовності
.....				27
4				
14.18	Вага тестового завдання. Чинники, які знижують цінність поняття ваги			
TЗ				
.....				27
4				

14.19	Трудомісткість і складність тестового завдання	27
6		
14.20	Підстави для типології тестових завдань	27
8		
14.21	Характеристика ЗУН, що виявляється шляхом тестування	28
2		
14.22	Вимоги до тестових завдань	28
3		
14.23	Недоліки тестових завдань та причини їх виникнення	28
5		
Список	рекомендованої	літератури
291		

ПЕРЕДМОВА

У запропонованому навчальному посібнику подано орієнтовний зміст лекцій навчального курсу «Комп'ютерні технології в тестуванні», який за навчальним планом підготовки магістрів (напряму 0403 Системні науки та кібернетика, напрям підготовки *6.040302 Інформатика**) структурно входить до циклу «Освітні вимірювання» за вибором навчального закладу.

Дисципліна «Комп'ютерні технології в тестуванні» реалізує додаткову підготовку вчителя інформатики в галузі теорії та практики педагогічних вимірювань з використанням комп'ютерних технологій в тестуванні.

Одним з найважливіших завдань дисципліни є підготовка майбутніх учителів, а також підвищення кваліфікації викладачів у галузі теорії і практики педагогічного тестування.

Сучасний стан системи освіти зобов'язує педагогів шукати нові підходи до оцінювання навчальних досягнень учнів. Це означає, що необхідно ознайомитися з підходами до оцінювання результатів навчання в інших країнах, результатами міжнародних досліджень, в яких бере участь Україна, визначати сучасні пріоритети в освіті.

Дисципліна передбачає вивчення питань прикладного спрямування: конструювання тестів різних рівнів застосування, контрольного матеріалу за цілями використання, інтерпретація результатів тестування, виконання статистичного опрацювання даних, тлумачення його результатів та формування висновків щодо його відповідності меті, завданням, розуміння основних характеристик якості контрольного матеріалу і можливості використання комп'ютерних технологій при створенні, проведенні і опрацюванні результатів тестування.

Цей навчальний посібник стане в нагоді майбутнім педагогам, учителям, викладачам, та усім, хто цікавиться тестуванням, технологіями конструювання тестів та їх якістю.

РОЗДІЛ І. ІНФОРМАТИЗАЦІЯ ОСВІТИ. НЕОБХІДНІ УМОВИ ІНФОРМАТИЗАЦІЇ СУСПІЛЬСТВА

Рівень розвитку країни значною мірою визначається рівнем розвитку освіти, яка повинна на сучасному етапі розвитку людства швидко й адекватно реагувати на потреби суспільства, позбавляючись шляхом проведення кардинальних реформ, притаманного теперішній освіті консерватизму. Одним із важливих чинників реформування освіти є її інформатизація. Побудова ефективних систем інформатизації освіти з урахуванням світового досвіду, особливостей і реалій стану вітчизняної освіти – одна із актуальних і важливих наукових і практичних проблем.

Різним аспектам інформатизації освіти присвячено численні дослідження. Проте, загальні методи і закономірності створення і використання засобів і систем інформатизації освіти з урахуванням необхідних напрямків реформування, видів діяльності, що здійснюються в системі освіти, сучасного стану інформатизації освіти і розвитку галузі інформаційних технологій в Україні, досліджені недостатньо. До основних публікацій, у яких висвітлені деякі аспекти зазначених загальних питань, можна віднести Закон України «Про основні засади розвитку інформаційного суспільства в Україні на 2007–2015 роки» [1], у якому, зокрема, дано оцінку нинішнього стану інформатизації освіти України і визначено основні напрямки її розвитку, статтю В.Г. Кременя [2], в якій наведені основні напрямки реформування освіти України, і статтю Б.М.Богатиря [3], у якій пропонуються деякі загальні підходи до вирішення проблем інформатизації системи освіти Російської Федерації. Проте публікацій, у яких висвітлювалося би питання відповідності існуючих засобів інформатизації вимогам реформування освіти і визначалися б напрямки досліджень, які необхідно здійснити, щоб забезпечити побудову ефективних систем інформатизації освіти, нами не знайдено. Ціль цієї статті – якоюсь мірою заповнити цю прогалину.

Сьогодні, на жаль, в Україні рівень інформатизації суспільства в цілому і освіти зокрема суттєво нижчий за рівень інформатизації суспільства й освіти розвинутих країн.

Важливою віхою у створенні інформаційного суспільства в Україні в цілому та інформатизації освіти зокрема повинен стати прийнятий 9 січня 2007 року за № 537-V Верховною Радою України Закон України «Про основні засади розвитку інформаційного суспільства в Україні на 2007–2015 роки» [1].

У Законі констатується, що ступінь розбудови інформаційного суспільства в Україні порівняно із світовими тенденціями, є недостатнім і не відповідає потенціалу та можливостям України. Причинами відставання є:

- низька ефективність використання фінансових, матеріальних, кадрових ресурсів, спрямованих на інформатизацію, слабе впровадження ІКТ у соціально-економічну сферу;

- недостатній розвиток нормативно-правової бази інформаційної сфери;

- недостатній рівень комп'ютерної та інформаційної грамотності населення та повільне впровадження нових методів навчання із застосуванням сучасних ІКТ;

- рівень державної підтримки виробництва засобів інформатизації, виробництво засобів інформатизації, програмних засобів та впровадження ІКТ є замалим, що не забезпечує всіх потреб економіки і суспільного життя;

- спостерігається нерівномірність забезпечення можливості доступу населення до комп'ютерних і телекомунікаційних засобів, поглиблення «інформаційної нерівності» між окремими регіонами, галузями економіки та різними верствами населення.

У Законі вказано також, що розвиток інформаційного суспільства в Україні та впровадження новітніх ІКТ в усі сфери суспільного життя і в діяльність органів державної влади та органів місцевого самоврядування визначається одним із пріоритетних напрямів державної політики.

Серед основних стратегічних цілей розвитку інформаційного суспільства в Україні названі такі:

- прискорення розробки та впровадження новітніх конкурентоспроможних ІКТ в усі сфери суспільного життя;

- забезпечення комп'ютерної та інформаційної грамотності населення, насамперед шляхом створення системи освіти, орієнтованої на використання новітніх ІКТ у формуванні всебічно розвиненої особистості;

- створення загальнодержавних інформаційних систем, насамперед у сферах охорони здоров'я, освіти, науки, культури, охорони довкілля.

Основними напрямками розвитку інформаційного суспільства в Україні, зокрема, визначені:

- надання кожній людині можливості для здобуття знань, умінь і навичок із використанням ІКТ під час навчання, виховання та професійної підготовки;

- створення умов для забезпечення комп'ютерної та інформаційної грамотності усіх верств населення, створення системи мотивацій щодо впровадження і використання ІКТ для формування широкого попиту на такі технології в усіх сферах життя суспільства.

Інформатизація освіти визнана одним із пріоритетних державних завдань. Таким чином, інформатизація системи освіти повинна бути невід'ємною складовою інформатизації України в цілому та здійснюватися згідно із єдиними державними нормативами, враховуючи при цьому особливості системи освіти.

Інформатизація освіти важлива не сама по собі. Вона повинна сприяти виконанню тієї місії, яка покладається на освіту суспільством. Можна погодитися з думкою Б.М.Богатиря [3], що «найголовнішою місією освіти в сучасних умовах є забезпечення стійкого соціально-економічного і науково-технічного розвитку країни з урахуванням її національних і регіональних культурних і соціальних особливостей, а також глобальних тенденцій у світі», додавши, що цей розвиток повинен бути спрямований на благо як суспільства в цілому, так і кожної людини зокрема.

Для виконання своєї місії освіта повинна бути реформована таким чином, щоб своєчасно й адекватно реагувати на виклики суспільства, продиктовані розвитком людської цивілізації.

У статті В.Г. Кременя «Суспільство знань і якісна освіта» [2] розкрито проблеми, які необхідно вирішувати в процесі реформування освіти, щоб вона відповідала сучасним цивілізаційним змінам суспільства.

По-перше, у зв'язку з тим, що зміна ідей, знань і технологій відбувається швидше, ніж зміна людського покоління, навчити людину на все життя під час звичної, традиційної освіти неможливо. Тому слід змінювати функції навчального процесу в освітніх закладах. Поряд із засвоєнням базових знань необхідно навчати учнів самостійно оволодівати новими знаннями та інформацією, навчити навчатися, самостійно вміти відчувати потребу в навчанні впродовж життя. Суттєвою є також функція навчального процесу – навчити людину використовувати отримані знання у своїй практичній діяльності.

Особливо це важливо в умовах, коли людство рухається до нової якості суспільного розвитку – суспільства знань, вирішальним чинником якого буде Людина, здатна діяти на основі отриманих знань і їх практичного використання.

По-друге, враховуючи, що процес глобалізації, який супроводжується розвитком сучасних інформаційних технологій, суттєвим чином збільшує сферу комунікації, у якій живе і функціонує людина. Вона отримує нескінченну множину інформаційних впливів з усього світу, вступає у відносини і контакти з громадянами своєї та інших країн. Ці впливи не тільки різноманітні, а й часто суперечливі, протилежні, що суттєвим чином ускладнює визначення самостійної позиції людини. Тому, навчально-виховний процес в освітніх закладах, а також і соціальне середовище суспільства в цілому, мають бути максимально зорієнтовані на формування розвиненої, самодостатньої особистості.

По-третє, зважаючи, що сенсом і основним показником прогресу людства є розвиток кожної окремої людини на основі її здібностей, актуальним завдання є максимально наблизити навчання і виховання кожної дитини до її сутності, здібностей та особливостей розвитку. На перший план тут виступає принцип

дитиноцентризму в значенні уваги до кожної дитини з її сутнісними характеристиками. Організація навчально-виховного процесу за принципом дитиноцентризму є єдиним шляхом формування людиноцентристського, гуманного, демократичного й ефективного сучасного суспільства і єдиним шляхом до щастя кожної людини. Упровадження принципу дитиноцентризму потребує як зміни навчальних планів, так і зміни взаємовідносин учителя й учня (професора і студента). З огляду, насамперед, на застосування нових інформаційних технологій, учитель має бути партнером учня в навчанні і розвитку.

По-четверте, для виконання освітою її місії необхідно, щоб держава забезпечила максимально наближені стартові можливості дітей із різних сімей у здобутті освіти і здійсненні перших самостійних кроків у житті.

Можна стверджувати, що інформатизація освіти – один з основних чинників вирішення вище зазначених проблем в освіті.

У вирішенні першої із зазначених проблем сучасні інформаційні технології, враховуючи можливості Інтернету, мають вирішальне значення. Такі їх властивості як:

- можливість надавати практично необмежені обсяги інформації з будь-яких галузей знань;
- можливість доступу до інформаційних ресурсів у будь-який час і в будь-якому місці;
- наявність в Інтернеті значної кількості спеціальних навчальних курсів із різних дисциплін, кількість яких постійно збільшується;
- можливість дистанційного навчання, що дозволяє кожній людині, яка має можливість користуватися сучасними інформаційними технологіями, навчатися в будь-який час впродовж усього свого життя, самостійно вибираючи бажану галузь і траєкторію навчання.

Завдання інформатизації освіти у вирішенні даної проблеми – постійне вдосконалення програмних і технічних засобів з урахуванням досягнень

педагогічних, психологічних і технічних наук, спрямованих на спрощення пошуку необхідних знань, їх засвоєння і практичне застосування.

Друга проблема – необхідність формування розвиненої, самодостатньої особистості, здатної приймати правильні рішення в умовах дії на неї все зростаючої кількості різноманітних впливів, часто суперечливих і протилежних – значною мірою є породженням інформатизації суспільства в цілому і освіти зокрема. Її вирішення, окрім відповідної орієнтації навчально-виховного процесу в освітніх закладах на формування особистості, може здійснюватися і засобами інформаційних технологій. Завдання інформатизації освіти у вирішенні цієї проблеми – розробка з урахуванням досягнень психологічної науки спеціального програмного забезпечення, присвяченого формуванню особистості, а також врахування під час розробки програм з інших навчальних предметів фактору впливу їх майбутнього використання на формування особистості.

У вирішенні третьої проблеми – організації навчально-виховного процесу за принципом дитиноцентризму – важливу роль повинні відігравати інформаційні технології. Уже сьогодні створені навчальні комп'ютерні програми і системи, які певною мірою можна адаптувати до здібностей учня. Завданням інформатизації освіти у вирішенні цієї проблеми є створення спільно з педагогами і психологами навчальних комп'ютерних систем, які дозволять максимально наблизити навчання і виховання кожної дитини до її сутності, здібностей та особливостей. Особлива увага повинна бути приділена створенню навчальних комп'ютерних систем для дітей з особливими потребами. Природно, що в таких навчальних системах учитель (викладач) повинен бути не ментором, а партнером учня (студента) у навчанні і розвитку.

Вирішення четвертої проблеми лежить як у науково-технічній, так і в фінансово-економічній і адміністративній площинах.

Одним із важливих факторів, що суттєво впливають на можливість одержання якісної освіти, є можливість користуватися сучасними інформаційними технологіями. Тому «забезпечити максимально наближені

стартові можливості дітей із різних сімей у здобутті освіти» в сучасних умовах означає, зокрема, забезпечити для всіх дітей рівні можливості користуватися інформаційними технологіями. Особливо це стосується загальної середньої освіти, яку безкоштовно (за рахунок держави) повинні одержати всі громадяни України. У процесі інформатизації освіти Академії педагогічних наук спільно з МОН України необхідно визначити науково обґрунтовані вимоги до засобів інформатизації, які необхідні та достатні для вивчення навчальних предметів в обсягах, передбачених державними стандартами загальної освіти, а державі забезпечити такими засобами всі навчальні заклади, що надають загальну середню освіту, згідно з єдиними нормами.

Вище наведені приклади свідчать, що інформатизація освіти є важливим засобом і складовою реформування освіти. Перехід системи освіти на якісно новий рівень без її інформатизації просто неможливий. Водночас необхідно зауважити, що використання інформаційних технологій необхідне і корисне не лише у реформованій, але й у нинішній освіті.

Здійснення інформатизації освіти вимагає з'ясування сутності цього процесу і визначення основних напрямків його реалізації.

Є різні визначення поняття інформатизація освіти [4, 5]. Вони відображають різні аспекти і складові процесу впровадження в систему освіти інформаційних технологій. Узагальнено ж можна сказати, що інформатизація освіти – це створення і використання інформаційних технологій для підвищення ефективності видів діяльності, що здійснюються в системі освіти.

Система освіти є складною ієрархічною системою, цілі і правила функціонування якої встановлюються Законами та іншими нормативно-правовими актами України у сфері освіти. Досягнення поставлених цілей і реалізація встановлених правил забезпечується шляхом побудови відповідної організаційної структури системи освіти і здійснення в ній певних видів діяльності. Реформування освіти – це передусім внесення змін у нормативно-правові акти у сфері освіти, які викликають відповідні зміни в структурі та діяльності в сфері освіти.

Основне завдання інформатизації освіти – впливати на виконання видів діяльності таким чином, щоб досягати поставлених цілей із меншими затратами ресурсів (часових, матеріальних, фінансових тощо).

Створення ефективної системи інформатизації освіти потребує дослідження системи освіти як об'єкта інформатизації. Необхідно розробити методiku досліджень, застосування якої забезпечувало б одержання параметрів об'єкта інформатизації, які необхідні для визначення оптимальних параметрів системи інформатизації. Слід зауважити, що під системою інформатизації системи освіти розуміється сукупність систем інформатизації всіх об'єктів освіти всіх рівнів ієрархії.

Основною сутністю інформатизації освіти є використання інформаційних технологій у різних видах діяльності, які здійснюються в системі освіти. Тому, досліджуючи систему освіти як об'єкт інформатизації, основну увагу необхідно приділити дослідженню цих видів діяльності, визначити критерії їх класифікації, виходячи з психолого-педагогічних та інформаційних характеристик і класифікувати їх за визначеними критеріями.

Для всіх кваліфікаційних груп видів діяльності в освіті необхідно визначити основні вимоги до засобів ІКТ для їх інформатизації.

Здійснення інформатизації освіти України на належному сучасному рівні з максимальною вигодою, зокрема, економічною, потребує також:

- дослідити сучасний стан застосування засобів ІКТ в освіті;
- дослідити фактори, що впливають на ефективність використання засобів ІКТ;
- дослідити відповідність існуючих засобів ІКТ визначеним основним вимогам та можливість їх ефективного використання для інформатизації усіх видів діяльності в освіті;
- сформулювати (можливо, у вигляді проекту нормативного документу) основні вимоги до перспективних засобів ІКТ;

- проаналізувати існуючі способи організації розробки і виробництва засобів ІКТ та забезпечення ними закладів і установ освіти в Україні і за кордоном;

- розглянути науково-технічний і промисловий потенціал України щодо розробки і тиражування засобів ІКТ для інформатизації освіти;

- розробити рекомендації щодо найбільш доцільних способів організації розробки і тиражування засобів ІКТ та забезпечення ними закладів і установ освіти;

- розробити рекомендації щодо підвищення ефективності використання засобів ІКТ в закладах і установах освіти України;

- дослідити існуючі способи моніторингу стану інформатизації установ і закладів освіти і ефективності використання засобів ІКТ.

Отже, аналіз напрямків реформування освіти і можливостей застосування інформаційних технологій під час їх здійснення, дає підстави стверджувати, що вирішальним чинником досягнення основних цілей реформування освіти є її інформатизація. Ефективні системи інформатизації освіти можуть бути створені, якщо вони базуються на дослідженні видів діяльності, що здійснюються в системі освіти, і враховують як сучасний стан, так і можливі трансформації освіти у процесі її реформування, сучасний стан і перспективи розвитку галузі інформаційних технологій, останні досягнення педагогіки, психології та інформаційних технологій навчання.

Для успішного вирішення проблем інформатизації освіти України доцільно комплексно проводити дослідження за вище зазначеними напрямками з широким залученням до виконання цієї роботи наукових установ АПН України, вищих навчальних закладів та виробників засобів ІКТ.

РОЗДІЛ II. ТИПОЛОГІЯ ПРОГРАМНИХ ЗАСОБІВ НАВЧАЛЬНОГО ПРИЗНАЧЕННЯ

2.1 Типи програмних засобів навчального призначення

Аналіз дидактичної доцільності використання програмних засобів (реалізація на ПЕОМ) із метою навчання проводиться багатьма фахівцями в галузі інформатизації освіти. Такого роду аналіз дозволяє з позицій положень, що описують дидактичну доцільність їх використання, створити обґрунтовані підходи до типології програмних засобів (ПЗ), використовуваних у навчальних цілях, вимоги до їх якості. В даний час у процесі навчання курсу інформатики і загальноосвітніх предметів використовується як матеріальне забезпечення певного типу ПЕОМ, так і програмне забезпечення, що відповідає цілям і завданням навчального процесу, і спеціально розроблені ПЗ і системи для цілей навчання і виховання. Процес розробки ПЗ навчального призначення, в основному, проводиться у двох напрямках:

а) ініціативна розробка авторськими колективами або окремими авторами, що створюють ПЗ на свій розсуд як відповідно до тематики досліджуваних навчальних курсів, так і незалежно від них;

б) розробка згідно з різними координаційними планами окремих відомств, наукових організацій, фірм, що має на меті насичення сфери освіти ПЗ, які можна використовувати в процесі навчання.

Програмні засоби навчального призначення використовуються як:

1) окремі програми, що забезпечують частину заняття або весь урок, декілька уроків;

2) пакети програм, що забезпечують тему навчального предмета, розділ або курс комп'ютерною підтримкою;

3) практикуми, що забезпечують певну програму практичних занять навчального курсу;

4) компонент комп'ютерного курсу або мультимедійного курсу.

Використання ПЗ навчального призначення в процесі вивчення загальноосвітніх предметів робиться в переважній більшості випадків разом з

навчально-методичною літературою, що забезпечує самий процес застосування того або іншого ПЗ і його включення в канву заняття.

Дидактична доцільність застосування ПЗ навчального призначення в більшості випадків обґрунтовується авторами, як необхідність автоматизації процесів контролю і тестування, рідше як необхідність демонстрації досліджуваних об'єктів, процесів або явищ з елементами моделювання. У переважній більшості ПЗ навчального призначення контроль обмежується пропозицією суб'єкту тестування вибрати вірну відповідь із декількох поданих на екрані (зазвичай 3-5), констатацією правильно отриманої відповіді (“вірно/невірно”), або пропозицією прочитати вірну відповідь у готовому виді (при неправильній відповіді). Такий підхід до контролю знань ніяк не можна назвати методично виправданим в умовах використання сучасних ПЕОМ. Це пояснюється тим, що їх можливості дозволяють забезпечити діагностику помилок за результатами навчання й у залежності від цього направити подальшу пошуково-творчу діяльність учня, організувати діалог у процесі виправлення помилок і забезпечити поетапний контроль (із зворотним зв'язком) на кожному логічно закінченому етапі діяльності учня.

2.2 Програмні засоби, що реалізують ідеї програмованого навчання

Програмні засоби, що реалізують ідеї програмованого навчання з використанням можливостей комп'ютера, досить широко подані як у вітчизняних, так і в зарубіжних розробках. Вони являють собою лінійні або розгалужені програми (за Скіннером або Краудером), реалізовані за допомогою комп'ютера, і нічого принципово нового з погляду дидактичних можливостей не містять.

Використання деяких програмних засобів підтримки навчання, наприклад, курсів інформатики, математики, фізики, мають мету – створення предметно-орієнтованого програмного середовища, робота з яким забезпечує різні види взаємодії з користувачем. Такі ПЗ є перспективними стосовно контролюючих, тренінгових програм або програм, що використовують ідеї програмованого навчання, тому що набагато повніше використовують можливості сучасних ПЕОМ (графічні, обчислювальні), надаючи учню інструмент для дослідження певної предметної галузі та реалізуючи ідеї діяльнісного підходу до навчання.

Інструментальні програмні засоби, призначені для конструювання програмних засобів (систем) навчального призначення або створення графічних включень, сервісних надбудов програми, підготовки навчальних і організаційних матеріалів, мають такі можливості:

1. Функціонування засобів діалогової підтримки процесу спілкування користувача з програмою.
2. Забезпечення контролю (самоконтролю) засвоєння навчального матеріалу і реакції програми на результати контролю.
3. Включення в програму засобів для здійснення обчислювальних операцій.
4. Включення в програму засобів, що забезпечують якісне оформлення, дизайн програми.
5. Генерування і розсилання роздаткового матеріалу на робочі місця учнів.
6. Забезпечення взаємодії інструментальних засобів з програмними засобами загального призначення (наприклад, із редактором тексту).
7. Забезпечення декількох робочих полів на екрані для створення активних зон, що забезпечують реакцію комп'ютерної програми на дії користувача.

Певний практичний інтерес представляє розробка сервісних ПЗ для викладача, що забезпечують автоматизацію процесу контролю результатів засвоєння або просування в навчанні, генерування завдань, завантаження і передачу ПЗ по мережі, керування процесом навчання (наприклад, із викладацької ПЕОМ). Сервісні ПЗ вкрай необхідні викладачу, тому що слугують цілям керування навчальним процесом і забезпечують комфортність роботи користувача.

2.3 Програмні засоби, що моделюють досліджуваний об'єкт, процес, експеримент або явище

У різних програмних системах досить часто зустрічаються ПЗ, що моделюють досліджуваний об'єкт, процес, експеримент або явище. Вони звичайно об'єднують демонстрацію досліджуваного процесу або експерименту, поданого моделлю, і контроль, будучи в деякому роді програмами-тренажерами. Вивчення властивостей моделі, що відбиває певний об'єкт або процес, наочне представлення досліджуваного

матеріалу через модель, можливість одержання негайного зворотного зв'язку при роботі з моделлю – все це підвищує дидактичний рівень такого типу ПЗ.

2.4 Ігрові компоненти в програмних засобах

Використання ігрової компоненти в ПЗ навчального призначення дозволяє програвати навчальну ситуацію та на фоні ігрового сюжету пропонувати вирішення нескладних навчальних завдань. Ці програми, швидше за все, мають дозвільне призначення, тому що ігрова компонента найчастіше в них превалює над навчальною. Набагато рідше за їхньою допомогою організується діяльність, що забезпечує розвиток логічного способу мислення. У переважній більшості такі ПЗ є програмами-тренажерами, з якими в кращому випадку хочеться поспілкуватися, але їх не можна назвати навчальними хоча б тому, що вони не сприяють процесу засвоєння знань або вмінь.

Окремо слід зазначити як позитивний напрямок використання ПЗ з ігровою компонентою, що реалізує в тій або іншій мірі ідею формування алгоритмічного або логічного стилю мислення. Використання таких ПЗ доцільно, у тому числі і в позаурочній роботі.

2.5 Дидактична доцільність використання ПЗ навчального призначення різних типів

Таким чином, дидактична доцільність використання ПЗ навчального призначення різних типів обпирається на наступні висновки:

- ✓ використання в практиці навчання ПЗ навчального призначення доцільно орієнтувати на тренування засвоєваних умінь, навичок, на контроль результатів навчання, на використання ігрової компоненти;
- ✓ застосування програмних засобів і систем у навчальному процесі значно підвищує мотивацію навчання за рахунок надання можливості самостійного вибору режиму роботи з ПЗ, забезпечення різноманітних видів самостійної роботи, комп'ютерної візуалізації, використання ігрових ситуацій;
- ✓ реалізація в ПЗ навчального призначення можливостей сучасної комп'ютерної графіки, різноманітних засобів наочності, формує і розвиває наочно-образний, наочно-діючий вид мислення;

✓ виконання учнем експериментально-дослідницької діяльності, організованої за допомогою ПЗ, формує дослідницькі вміння, ініціює самостійне придбання знань;

✓ реалізація ідей алгоритмізації навчання в ПЗ навчального призначення ініціює навчання оптимальному пошуку стратегії вирішення завдань певного класу, розвиває алгоритмічний, логічний стиль мислення;

✓ використання навчальних баз даних, електронних таблиць, застосування системи підготовки текстів, включення в навчальні програми діяльності з класифікації навчальної інформації формує інформаційну культуру учня.

Таким чином, варто констатувати необхідність висування на передній план (при розробці перспективних ПЗ навчального призначення) концептуальних психолого-дидактичних аспектів проблеми створення і використання ПЗ із метою навчання і розвитку інтелектуального потенціалу індивіда, у тому числі визначення умов реалізації унікальних (із погляду методичних застосувань) можливостей сучасних інформаційних технологій.

Перейдемо до розгляду типології програмних засобів, використовуваних у навчальних цілях за функціональним і методичним призначенням.

2.6 Типологія ПЗ за функціональним призначенням

Їх доцільно розділити на наступні типи.

Прикладні програми, призначені для організації і підтримки навчального діалогу користувача з комп'ютером. Частіше всього їх називають педагогічні програмні засоби (ППЗ). Функціональне призначення ППЗ – давати навчальну інформацію і спрямовувати навчання, з огляду на індивідуальні можливості і переваги учня. Як правило, ППЗ припускають засвоєння нової інформації при наявності зворотного зв'язку користувача з програмою.

Діагностичні, тестові програми, мета яких констатація причин помилкових дій учня, оцінка його знань, умінь, навичок, установлення рівня його навченості або рівня інтелектуального розвитку.

Інструментальні програмні засоби (ІПЗ) прикладного призначення можна розділити на наступні основні види:

1. Інструментальні системи, призначені для розробки автоматизованих засобів або систем контролюючого, консультуючого, тренінгового призначення, що дозволяють звести до мінімуму «паперове» подання навчального матеріалу, замінюючи його «екранним».

2. Авторські програмні системи, призначені для конструювання програмних засобів (систем) навчального призначення.

3. Системи комп'ютерного моделювання (демонстраційного, імітуючого).

4. Програмні засоби з убудованими елементами технології навчання, що включають як предметне середовище, так і елементи педагогічної технології для її вивчення.

5. Інструментальні програмні засоби, що забезпечують здійснення операцій з систематизації навчальної інформації на основі використання системи обробки даних (наприклад, інформаційно-пошукові системи, навчальні бази даних з різних галузей знань).

6. Експертні системи навчального призначення як засіб подання знань, призначені для організації діалогу між користувачем і системою, здатною за вимогою користувача представити хід міркувань при вирішенні того або іншого навчального завдання у виді, прийнятному для учня.

Предметно-орієнтовані програмні середовища дозволяють здійснювати моделювання досліджуваних об'єктів або їхніх відношень у певному предметному середовищі (при необхідності, не відображаючи об'єктивну реальність); за їх допомогою звичайно організується навчальна діяльність із моделями, що відображають об'єкти, закономірності деякої предметної області.

Програмні засоби, призначені для формування культури навчальної діяльності, інформаційної культури на основі застосування системи підготовки текстів, електронних таблиць, графічних і музичних редакторів або інтегрованих систем їхнього комплексного використання.

Програмні засоби, призначені для автоматизації процесу обробки результатів навчального експерименту, у тому числі що вимірюють і контролюють програми для датчиків, які дозволяють одержувати, записувати і візуалізувати інформацію про процеси, що реально протікають.

Керуючі ПЗ, мета яких – керування діями реальних об'єктів.

Навчальні середовища програмування, призначені для початкового навчання навичкам програмування і формування основних компонентів алгоритмічного і програмістського стилю мислення.

ПЗ, що забезпечують виконання деяких функцій викладача.

Ці ПЗ звичайно видають команди, що стосуються роботи на комп'ютері, вказівки про припинення роботи, про виконання перевірки, про необхідність модифікації інформації, про одержання додаткових даних, про обговорення ходу роботи.

Сервісні програмні засоби, що забезпечують комфортність роботи користувача (автоматизація процесу контролю результатів навчання, генерування і розсилання організаційно-методичних матеріалів, завантаження і передача ПЗ по мережі, керування процесом навчання).

Ігрові ПЗ, що є засобом, який забезпечує різні види ігрової й навчально-ігрової діяльності.

Крім поданої вище типології ПЗ за функціональним призначенням, можлива і доцільна типологія ПЗ також за методичним призначенням.

Методичне призначення кожного типу ПЗ відбиває методичну ціль (або цілі) його використання в процесі навчання і ті можливості ПЗ, реалізація яких інтенсифікує навчальний процес, переводить його на якісно більш високий рівень. Доцільність проведення такої типології ПЗ викликана низкою обставин, із яких основними є необхідність: а) вибору викладачем, методистом потрібного ПЗ із наявних або запропонованих; б) порівняння ПЗ у рамках одного типу для добору найкращого; в) створення ієрархії ПЗ за складністю і трудністю; г) орієнтування користувача в множині наявних ПЗ різного методичного призначення.

2.7 Методичні цілі використання ПЗ навчального призначення

Назвемо найбільш значущі, з позиції дидактичних принципів методичні цілі, що найбільш ефективно реалізуються з використанням ПЗ:

1) індивідуалізація і диференціація процесу навчання (наприклад, за рахунок можливості поетапного просування до цілі по лініях різного ступеня складності і трудності;

2) здійснення контролю зі зворотним зв'язком, із діагностикою помилок (констатація причин помилкових дій того, кого навчають, і подання на екрані комп'ютера відповідних коментарів) результатів навчання і оцінкою результатів навчальної діяльності;

3) здійснення самоконтролю і самокорекції;

4) здійснення тренування в процесі засвоєння навчального матеріалу і самопідготовки учнів;

5) вивільнення навчального часу за рахунок виконання на ЕОМ трудомістких обчислювальних робіт і діяльності, зв'язаної з числовим аналізом;

6) комп'ютерна візуалізація навчальної інформації: по-перше, досліджуваного об'єкта (наочне подання на екрані ЕОМ об'єкта, його складових частин або їхніх моделей, а при необхідності – у всіляких ракурсах, у деталях, із можливістю демонстрації внутрішніх взаємозв'язків складових частин); по-друге, досліджуваного процесу (наочне подання на екрані ЕОМ даного процесу або його моделі, у тому числі схованого в реальному світі, а при необхідності – у розвитку, у часі і просторовому русі, подання графічної інтерпретації досліджуваної закономірності процесу, що вивчається);

7) моделювання або імітація всіх досліджуваних об'єктів, процесів або явищ;

8) проведення лабораторних робіт (наприклад, з фізики, хімії) в умовах імітації в комп'ютерній програмі реального дослідження або експерименту;

9) створення і використання інформаційних баз даних, необхідних у навчальній діяльності, і забезпечення доступу до мережі інформації;

10) посилення мотивації навчання (наприклад, за рахунок образотворчих засобів програми або вкраплення ігрових ситуацій);

11) озброєння того, кого навчають, стратегією оптимального засвоєння навчального матеріалу;

12) розвиток певного виду мислення (наприклад, наочно-образного, теоретичного, фізичного);

13) формування вміння приймати оптимальне рішення або варіативні рішення в складній ситуації;

14) формування культури навчальної діяльності, інформаційної культури учня та вчителя.

2.8 Типологія ПЗ за методичним призначенням

На основі перерахованих методичних цілей використання ПЗ приведемо типологію за методичним призначенням:

1 Навчальні програмні засоби, методичне призначення яких – повідомлення суми знань, формування вмінь і (або) навичок навчальної і (або) практичної діяльності і забезпечення необхідного рівня засвоєння, який встановлюється зворотним зв'язком, реалізується засобами програми.

2 Програмні засоби (системи) – тренажери, призначені для відпрацювання вмінь, навичок навчальної діяльності, здійснення самопідготовки. Вони звичайно використовуються при закріпленні раніше вивченого матеріалу.

3 Програми, призначені для контролю (самоконтролю) рівня оволодіння навчальним матеріалом, – контролюючі програмні засоби.

4 Інформаційно-пошукові програмні системи, інформаційно-довідкові програмні засоби, що дають можливість вибору і виводу необхідної користувачу інформації. Частіше всього їх методичне призначення – формування вмінь і навичок щодо систематизації інформації.

5 Імітаційні програмні засоби (системи), що представляють певний аспект реальності для вивчення його основних структурних або функціональних характеристик за допомогою деякого обмеженого числа параметрів.

6 Моделюючі програмні засоби довільної композиції, що дають у розпорядження учня основні елементи і типи функцій для моделювання певної

реальності. Вони призначені для створення моделі об'єкта, явища, процесу або ситуації (реальних або віртуальних) з метою їхнього вивчення, дослідження.

7 Демонстраційні програмні засоби, що забезпечують наочне подання навчального матеріалу, візуалізацію досліджуваних явищ, процесів і взаємозв'язків між об'єктами.

8 Навчально-ігрові програмні засоби, призначені для програвання навчальних ситуацій.

9 Дозвільні програмні засоби, які використовуються для організації навчально-ігрової і просто ігрової діяльності учнів, що мають, як правило, ціллю розвиток уваги, реакції, пам'яті і т.д.

Аналіз практики розробки і використання ПЗ із метою навчання переконує в тому, що частіше всього вони мають змішане методичне призначення. Тому доцільна розробка ПЗ навчального призначення, що забезпечує реалізацію комплексу методичних цілей.

1. Психологічні проблеми комп'ютеризації освіти.

Впровадження комп'ютерів у сферу освіти змінює форми, зміст, засоби навчання, наближає його до реалій сучасного життя. Основним завданням викладача в такому разі є: «розвиток особистості того, хто вчиться, творчий пошук в організації навчального процесу, добір, розробка та вибір найкращих і найдоцільніших для навчання програмних продуктів»(В.Васильєв). Під час роботи на ЕОМ наявні комфортні умови для активного пізнання, прийняття самостійних рішень, моделювання реальних процесів. Усе це є значним фундаментом для випускників вузів. Виділяють кілька етапів комп'ютеризації освіти. Зрозуміло, що застосування ЕОМ у навчальному процесі призвело до зміни засобів та організаційних форм навчання. Є можливість використання, залежно від бажання учнів, різних електронних довідників, ілюстративного матеріалу, методичних порад. Ці можливості пропонуються в будь-якому індивідуальному обсязі. Можна всіляко варіювати самостійну діяльність учнів; включення його в процес мислення та засвоєння нових знань відбувається в умовах спілкування з комп'ютером; існує можливість візуалізації не тільки

безпосередньо отриманого результату, а й усього ходу мислення; комп'ютерні технології можна використати в різних формах управління навчальною діяльністю. Саме комп'ютер може виступати як пасивним елементом, так і активним, все залежатиме від того, яка роль для нього відведена користувачем; виникають різні нові форми організації навчання: дистанційне навчання, електронні конференції, електронна пошта.

Діяльність учнів теж помітно зазнає змін. Вони одержують можливість самостійно приймати рішення. Не звертаючись до педагога, можуть отримати необхідну інформацію з комп'ютерної системи і навіть проєкспериментувати з нею, розглядаючи різні шляхи розв'язання задачі, з частковою, повною допомогою комп'ютера або взагалі без неї. Завдяки творчому пошуку — можуть не боятися зробити помилку: комп'ютер не дозволить. Учні, таким чином, застраховані від зайвих емоційних стресів. Ще одним важливим аспектом є можливість займатися дослідницькою роботою, простежуючи в динаміці розвиток різних явищ та процесів.

З активним упровадженням комп'ютерної техніки виникають питання, пов'язані з комп'ютерною грамотністю та інформаційною культурою.

Проблему комп'ютерної грамотності деякі автори розглядають дуже вузько, як «оволодіння вміннями розв'язувати задачі на комп'ютері». З цим не можна погодитися тому, що розв'язуванням задач функції комп'ютера не обмежуються, бо, застосовуючи комп'ютери, можна продивлятися текстову інформацію, наприклад, з різних галузей знань. ЕОМ можна використовувати при прогнозуванні, моделюванні різних явищ, процесів і т. ін. Існує точка зору, що комп'ютерна грамотність включає в себе тільки знання про можливості комп'ютера, інші дослідники називають тільки практичне застосування ЕОМ при розв'язуванні складних задач. Ми вважаємо, що комп'ютерна грамотність включає в себе знання новітніх інформаційних технологій та їх функціонування; умови більш ефективного застосування тих або інших технологій; обов'язкове знання мережі Internet та вміння працювати в ній; знання застосування комп'ютерів в усіх галузях людської діяльності; знання

алгоритмічних мов та уміння їх практичного застосування (не тільки для фізико-математичного циклу навчальних дисциплін); уміння працювати з програмними продуктами, в тому числі й з текстовими редакторами; знання архітектури ЕОМ. У свою чергу комп'ютерна грамотність дає учням знання про вплив загальної автоматизації на діяльність суспільних інститутів.

Інтенсивне впровадження комп'ютерних технологій в освіту актуалізує питання змісту, форм самостійної роботи учнів. Від викладачів вищої школи вимагається якомога більше уваги приділяти різним видам самостійної роботи, бо вони є важливим чинником культури навчальної праці. Практично всі програмні засоби (комп'ютерні електронні підручники, гіпертексти, експертні системи) мають деякі елементи самостійної роботи. Організація самостійної роботи учнів із застосуванням ЕОМ може водночас здійснюватися по кількох напрямках. Наприклад, Н.Полякова виділяє такі: «розробка часткових алгоритмів розв'язання типових задач; розробка евристичних приписів нетипових задач; розробка навчаючих програм, як більш вищий ступінь алгоритмізації; індивідуалізація самостійних робіт; спеціалізація самостійної роботи з урахуванням практичних задач спеціальності; розробка систем рефератів з окремих розділів лекційних курсів; розробка спеціальних методів навчання; забезпечення спеціальною й довідковою літературою, застосування ПЕОМ і т. ін.».

Практичне використання комп'ютерної техніки в навчальному процесі надає можливість вибирати різні організаційні форми роботи — індивідуальну, групову або колективну. Це ще раз спростовує помилковість думки про те, що комп'ютерне навчання за організаційною формою завжди має тільки один певний вид. Використання колективної навчальної діяльності в навчальному процесі також має переваги. Завдяки такій діяльності учень може виступати і як викладач, і як учень. В цьому випадку також передбачається виховна мета, а саме: в учнів формуються навички спільної діяльності, з'являється почуття відповідальності за всю групу, уміння поважати думку інших при обговоренні рішення певної задачі. Колективна форма навчання дає можливість

спілкування, обговорення ходу рішення задачі (при цьому кожний учень висловлює свою точку зору,) її кінцевого результату. Відбувається врахування всіх пропозицій, висловлених у процесі обговорення. Підвищується ефективність навчального процесу, особливо індивідуального навчання. Учень стає об'єктом педагогічного впливу й активним суб'єктом навчальної діяльності. Отже, можна зробити висновок, що ЕОМ із точки зору використання організаційних форм у навчальному процесі є універсальною.

Серед першочергових завдань, що стоять перед вищими навчальними закладами, є збільшення кількості ЕОМ з використанням сучасних обчислювальних технологій (інтегровані бази даних та знань, текстові, експертні системи, сучасне програмне забезпечення, візуальне та графічне подання інформації).

Можна сформулювати такі переваги навчання за допомогою персональних комп'ютерів:

— комп'ютер дозволяє контролювати індивідуальний темп навчальної діяльності: в учня є можливість самостійно вибрати момент переходу до наступної порції навчальної інформації, задавати швидкість подання задач, час розв'язання задачі. При цьому комп'ютер може служити певним регулятором, попереджаючи учня, що його темп роботи або занадто повільний, або необґрунтовано швидкий;

— комп'ютер дає змогу учневі самостійно переходити від більш високого ступеня важкості навчального матеріалу до більш легкого і навпаки. Учень вибирає й визначає для себе ступінь складності навчальної задачі. Є можливість пробувати свої сили на різних рівнях складності і враховувати прогалини в засвоєнні навчального матеріалу;

— образотворча роль комп'ютера — унікальна. Велике значення мають образотворчі ефекти комп'ютера, із допомогою яких можуть демонструватися процеси і явища, що вивчаються. Комп'ютери роблять Інформацію видимою, завдяки цьому, фундаментально змінюючи людську комунікацію, бо носієм інформації стає картина (анімація, графіка);

— комп'ютер може «особисто» спілкуватися з учням і певним чином мотивувати їхню діяльність (похвалою, оцінкою, зважаючи на помилкові дії та ін.)- Але необхідно чітко враховувати комплекс характеристик ученьів при використанні цієї можливості комп'ютера: застосування цього навчального впливу може мати як позитивні, так і негативні наслідки.

Аналіз досвіду застосування комп'ютерної техніки у вузі, зокрема, на практичних заняттях, дає змогу зробити деякі висновки з проблем застосування нових інформаційних технологій у процесі активізації пізнавальної діяльності учнів при впровадженні індивідуалізованого підходу в навчанні. Використання персональних комп'ютерів дає змогу:

1) інтенсифікувати й раціоналізувати час, що витрачається при традиційному навчанні на демонстраційні моменти, пошук інформації, на індивідуальний темп роботи з теоретичним матеріалом;

2) підвищити ефективність процесу навчання внаслідок переробки великого обсягу навчальної інформації, необхідної для кожного учня;

3) розвивати пізнавальну активність, підвищувати ефективність самостійної роботи, інтерес до теми, що вивчається;

4) установлювати взаємозв'язок, необхідний для управління навчальним процесом, систематично контролювати знання та вміння, отримані під час опрацювання навчального матеріалу, підвищувати рівень перевірки знань;

5) удосконалювати форми й засоби організації самостійної, диференційованої, індивідуальної, індивідуалізованої роботи учнів;

6) індивідуалізувати процес навчання, застосовуючи окрему методику, заставлену в ЕОМ для моделі кожного учня, що дозволяє враховувати індивідуальні особливості учня, розвивати здібності, підвищувати якість засвоєння знань та необхідних умінь.

Працюючи на персональному комп'ютері та освоюючи певний навчальний курс, учень водночас набуває й відпрацьовує навички роботи з електронно-обчислювальною технікою, що в останні десятиріччя набуває всезростаючої ролі в усіх сферах діяльності людини. Використання комп'ютерної техніки в

навчальному процесі може виступати не тільки засобом навчання, а й предметом вивчення. Однак не слід усі проблеми й питання в педагогічному процесі вирішувати шляхом застосування комп'ютерів. Необхідно враховувати принцип педагогічної доцільності при використанні певних засобів. Маючи відправною точкою конкретну дидактичну мету певного навчального матеріалу і конкретної моделі учня, необхідно вибирати форми й засоби навчання, які б стимулювали пізнавальну активну діяльність учнів.

Основними чинниками, що сприяють ефективному проведенню індивідуалізованої роботи з учням, є:

- проведення викладачем чіткої настанови на систематичне використання індивідуалізованої форми навчання під час занять;

- розробка теоретичного навчального матеріалу відповідно до дидактичного принципу доступності;

- розробка системи спеціальних завдань для індивідуалізованої форми навчання, які відповідають навчальним планам та програмам з урахуванням специфіки кожної навчальної дисципліни;

- проведення систематичного контролю при індивідуалізованій формі навчання;

- врахування комплексу індивідуально-психологічних особливостей учнів при організації індивідуалізованого навчання;

- врахування психолого-педагогічних умов у розробці комп'ютерних програмних продуктів, які будуть використовуватися при індивідуалізованому навчанні.

При цьому критеріями ефективності індивідуалізації навчального процесу є:

- стійке засвоєння знань учням із певного предмета;

- уміння використати теоретичні положення навчальної інформації у практичній роботі;

- уміння самостійно працювати з навчальною та додатковою літературою;

- уміння індивідуально проробляти та освоювати запропонований

навчальний матеріал.

Складність визначення дидактичних умов ефективного використання комп'ютерів при індивідуалізованому підході до навчання зумовлена відсутністю єдиної точки зору на визначення поняття індивідуалізація навчання, специфічних особливостей, функцій та шляхів її реалізації в умовах комп'ютерного навчання.

Також відсутня єдина точка зору і на визначення специфічних ознак індивідуалізації, її навчально-виховних результатів.

Ядром теоретичної концепції індивідуалізації навчальної діяльності О.Кірсанов називає «три основні категорії: особистість, діяльність, розвиток». Основні складові, що творять процес навчання, є передача знань, досвіду діяльності, емоційно-ціннісних відношень.

Жодна з існуючих нині психолого-педагогічних теорій навчання, в тому числі П. Гальперін та Н. Талізінної (поетапне формування розумових дій), не можуть у повній мірі і належним чином пояснити закономірності комп'ютерного навчання. Одною з причин такого положення, на думку В. Панченко, є те, що комп'ютерна діяльність «накладає відбиток» на всі види пізнавальних процесів, тому цей аспект проблеми дає великі можливості для дослідницької роботи психологів і педагогів.

Серед різних аспектів вивчення комп'ютеризації (філософського соціального, методологічного та ін.) основним і істотним називають психолого-педагогічний.

Поява нових інформаційних технологій навчання, в тому числі й використання комп'ютерної техніки, зумовила і появу нових форм організації навчального процесу, за яких учень виступає об'єктом навчальної діяльності. Комп'ютер вирішує досить складні задачі за допомогою евристичних засобів. При цьому змінюється статус учня, він усвідомлює себе людиною, яка самостійно приймає певні рішення і відповідає за одержаний результат. А це розкриває нові горизонти в практичній реалізації принципу гуманізму та гуманізації у навчанні при використанні комп'ютерних технологій.

Комп'ютер як засіб навчання вже використовується майже 45 років. І за цей час його використання змінювалися функції і дидактичні можливості, більш серйозний підхід отримала психологічна сторона експлуатації такого засобу. Зазнали змін і прикладні комп'ютерні продукти (навчальні системи, оболонки, середовища), змінилися і їхні можливості та функції.

Використання техніки в навчальному процесі традиційно характеризується двома шляхами:

— взаємодією учня з комп'ютером. Навчання відбувається, як правило, без втручання викладача, оскільки учень одержує всю необхідну інформацію тільки від комп'ютера, спілкуючись із ним у зручному режимі. Єдино можливе втручання викладача тоді, коли через недосконалий програмний продукт виникає непередбачена ситуація й комп'ютер неспроможний будь-що зробити;

— взаємодією викладача з комп'ютером. У цьому разі комп'ютер може застосовуватися як помічник в організації навчального процесу, або ж виконувати контролюючу функцію. Він може видавати результати після виконання учням певної роботи, порівнювати їх, вести підрахунки.

Нові інформаційні технології навчання вносять істотні зміни і в діяльність педагога. Його місце в навчальному процесі значною мірою визначається вибраною навчальною системою, і залежно від функцій, що виконуються цією системою, зумовлюються і функції викладача. Успішне використання комп'ютерних технологій ставить нові вимоги до професійної майстерності викладача.

РОЗДІЛ ІІІ. ПЕДАГОГІЧНЕ І ПСИХОЛОГІЧНЕ ТЕСТУВАННЯ

3.1 Переваги та недоліки тестування

Серед продуктів системи освіти - знання, уміння і навички, передані учням і сприйняті ними. Якість цієї освітньої продукції відіграє в розвитку суспільства першочергову роль.

Без виміру і порівняння якості продукції неможливий розвиток виробництва. У освіті для виміру якості застосовують поурочний, рубіж, підсумковий та інші види контролю. Кожен вид контролю характеризується своєю технологією.

Найбільш об'єктивна технологія масового контролю, на думку багатьох, - автоматична, одноманітна, не залежна від будь-яких суб'єктивних впливів. Хоча, у цієї точки зору є і супротивники, які вважають, що контроль повинен проводитися "очі в очі".

До слабких моментів першої точки зору можна віднести: неможливість автоматичними засобами перевірити і гарантовано оцінити уміння:

- викладати власні думки;
- знаходити нестандартні вирішення;
- учити і вчитися.

Друга точка зору також має слабкі місця:

- висока собівартість;
- відсутність гарантії виключення суб'єктивізму;
- відсутність в регіонах країни достатньої кількості кадрів високої кваліфікації, здатних гарантовано неупереджено та однаково оцінити властивості великої кількості претендентів.

У суперечці між точками зору слід займати "філософську" позицію, знаходити золоту середину. Наприклад:

- контролювати автоматичними засобами знання рутинних(буденних) відомостей, умінь виконувати стандартні дії - тобто те, що у ВНЗ відносять до "заліку";

- контролювати силами фахівців уміння мислити і викладати думки, а також інші властивості, які неможливо контролювати засобами автоматизації.

Атестація ЗУН - багатосторонній контроль, що проводиться найбільш об'єктивною з можливих технологій і супроводжується оформленням документу, який визнається в суспільстві.

Атестація ЗУН - найважливіший елемент навчального процесу. Об'єктивність атестації - одна з гарантій розвитку і підтримки життєдіяльності суспільства.

У світі розроблена і впроваджена безліч підходів до атестації ЗУН, заснованих на застосуванні комп'ютерів і телекомунікаційних технологій. Багато з них визнано достатньо об'єктивними.

Проте серед відомих комп'ютерних навчальних і тестуючих комплексів важко знайти такі, що абсолютно витримують критику з точки зору:

- об'єктивності;
- зручності;
- привабливості;
- повноти набору необхідних функцій контролю і представлення результатів;
- об'ємності змісту.

Одне з виключень - професійні військові тренажери (доречно зауважити, часто створені за кордоном руками наших співвітчизників).

Це пов'язано з багатьма причинами:

- по-перше, у нас педагогічне тестування довгий час було заборонене нормативними документами і, в результаті, нам доводиться наздоганяти інші країни в цій технологічній області. А також, в кібернетиці і генетиці, котрі більше "на слуху";

- по-друге, період первинного розвитку технологій комп'ютерного тестування припав на період економічної кризи;

- по-третє, розвиток якостей вітчизняних комп'ютерних комплексів гальмує прагматичне ставлення до результату їх застосування, що має просту

економічну підоснову. Щоб продукт наділити перерахованими вище якостями – потрібно чималих фінансових затрат;

- по-четверте, культура виробництва у нас значно нижча, ніж в інших країнах, де підприємцеві не все одно, яких працівників він приймає на роботу, і де підприємці давно застосовують автоматичні методи для контролю ЗУН. Зрідка вітчизняні працедавці переймаються важливістю цих заходів.

Перерахуємо недоліки традиційних систем оцінки і атестації ЗУН :

- суб'єктивізм;
- нерегулярність;
- неузгодженість вимог і режимів контролю;
- складність забезпечення режиму секретності зберігання екзаменаційних матеріалів і захисту від підробки;
- наявність випадкових, несистематичних помилок вимірів;
- відсутність чітких математичних критеріїв оцінки.

Багато що з перерахованого цілком може бути подолано за допомогою впровадження комп'ютерних технологій. Проте, поки це - тільки побажання, підтвержене теоретично. Бездоганної системи автоматичного контролю ЗУН нині не існує.

Пов'язано це з багатьма причинами:

- різномірність, часта зміна і відносна незрілість (молодість) як ІТ, так і відповідного технічного забезпечення;
- відсутність належної державної підтримки розробки і впровадження;
- відсутність фахівців;
- відсутність адекватної теоретичної бази.

Шлях подолання цих проблем пов'язаний з новими управлінськими рішеннями і заходами, зокрема такими:

- державна підтримка розвитку тестології як галузі знань і науки;
- створення індустрії систем тестового контролю.

У сучасному науковому світі склалася суперечлива ситуація з визнанням тестології як повноправної науки. З одного боку, загальне визнання тестового

контролю дієспроможним напрямом дослідження широкого спектру педагогічних і психологічних проблем; з іншої - тривала стриманість, неповнота інформації, нерозуміння суті та можливостей.

Можливо, слід більш конкретизувати поняття освіти, наочніше продемонструвати його як товар? Там, де товар, там і сертифікація, там і тестування.

В Україні вважається нормальним тестувати горілку, шпроти, огірки, квас і багато іншого. Нормально, коли увесь товар має сертифікат і виробник товару користується захистом якості свого товару з боку держави. Відповідні гарантії про тестування, що даються незалежними центрами, дозволяють виробникам, покупцям і продавцям цивілізовано захищати свої права. Ті ж гарантії дозволяють об'єктивно оцінювати продукцію, вести стратегічне планування виробництва.

Чому такого нема на українському ринку освітніх послуг? Причин тому багато:

- не має чіткого математичного визначення освіти;
- управлінці в системі освіти не мають достатньої освіти;
- проникнення математики в педагогіку - далека перспектива;
- існують в суспільстві сили, що перешкоджають впровадженню комп'ютерного тестування (з комп'ютером важче домовитись про підміну, чим з екзаменатором).

- відсутнє в Законі "Про освіту" поняття та роз'яснення про захист якості.

Система освіти озброєна лише поняттями акредитації, сертифікації, атестації - тобто тими, що відносяться до формального опису якості.

3.2 Педагогічне та психологічне тестування

У багатьох областях практичної психології широко використовуються вимірювальні психодіагностичні методики, до яких відносяться тести на вимірювання здібностей, досягнень, апаратурні методики, засновані на стандартизованих самозвітах - опитувальниках та техніці суб'єктивного шкалювання.

Коректність у застосуванні цих методик забезпечується не тільки змістовними уявленнями, а й виконанням особливих вимог психометрики. Тестові методики покликані вирішити певне обмежене коло завдань. Це завдання масової експрес-діагностики. Тут не виключені помилки в індивідуальних випадках, діагноз і прогноз подаються лише з ймовірнісною точністю. «...Перш ніж застосувати будь-який тест-опитувальник на особливому контингенті осіб, психолог повинен постаратися подивитись на кожне питання очима суб'єкта тестування. Це вимагає чималої професійної інтуїції ...» (Шмельов А. Г., Похилько В. І. Аналіз пунктів при конструюванні та застосуванні тест-опитувальників: ручні та комп'ютерні алгоритми. - Питання психології, № 4, 1985. - С. 126134.)

Цей текст дозволяє зануритися в коло проблем, які вирішують психологи за допомогою тестування. Психологу, на відміну від педагога, доводиться працювати не з групами учнів, а з абсолютно різними людьми. Тести психологів відрізняються від педагогічних дивовижністю, несподіванкою і простотою завдань. За результатами виконання завдань, психолог вирішує складне завдання - визначає тип людини та її характеристики, беручи до уваги безліч шкал і визначень, у тому числі суперечливих, винайдених як ним самим, так і його колегами попередниками.

Перед психологом, що складає тест, завжди стоїть завдання, яке важко вирішується: створити мінімальну кількість тестових завдань, що дозволяють отримати максимально достовірну і повну інформацію про людину. Тестові завдання психологів багатофункціональні - результат виконання їх «говорить» частково і відразу про декілька характеристик людини. Їх складання - це шлях проб і помилок. Кожне завдання перевіряється на великих групах суб'єктів тестування і лише потім робиться висновок про включення або не включення його в тест. Оскільки можливості експерименту завжди обмежені, немає гарантії, що запитання однаково надійно буде виконувати своє завдання завжди. Крім визначення типу людини, психолог визначає, з часткою вірогідності:

- професійну придатність;
 - ділові якості;
 - спосіб мислення;
 - схильності;
 - психічні хвороби і розлади;
 - сумісності,
- інші характеристики.

До освіти ці характеристики мають віддалене відношення

3.3 Бланкове тестування

Ранні спроби автоматизувати контроль ЗУН в багатьох країнах на початку ХХ століття привели до створення бланкових технологій тестування та «індустрії тестів», а потім - до появи міжнародно визнаних організацій, які проводять щорічно десятки мільйонів сеансів бланкових тестувань.

Недосконалості бланкових технологій тестування стали очевидними після появи достатньо якісних комп'ютерних дисплеїв (моніторів) у 80х роках минулого століття.

До технічних недосконалостей бланкових технологій можна віднести:

- тривалість тиражування;
- висока ймовірність і простота підробки;
- складність збереження конфіденційності;
- висока ймовірність вгадування вірних варіантів шляхом простого логічного порівняння;
- трудомісткість тиражування, зберігання і транспортування бланків і відповідних анкет;
- складність забезпечення належного рівня секретності екзаменаційних матеріалів;
- складність забезпечення правильного сприйняття правил претендентами і їх виконання;
- трудомісткість вилучення і обробки результатів;
- дорожня сеансу та інші.

До методичних недосконалостей бланкових технологій ми віднесемо ще одне, на наш погляд, найсуттєвіше. Суть цієї недосконалості легко зрозуміти, якщо уважно порівняти дві популярні телевізійні гри - «Хто хоче стати мільйонером?» та «Своя гра». У першому випадку один грає вказує вірну відповідь на запитання в пред'явленому списку. У другому - гравці змагаються один з одним у швидкості відповіді на питання, знаючи лише його тему і використовуючи лише свою пам'ять.

З методичної точки зору, ці два сценарії кардинально різні. У першому випадку вказати правильну відповідь незрівнянно легше:

- по-перше, бачачи чотири варіанти відповіді, серед яких лише один вірний, можна здійснити вибір просто логічним шляхом, не маючи спеціальних знань;
- по-друге, навіть якщо немає можливості згадати вірну відповідь або здогадатися шляхом логічного перебору варіантів, можна вибрати будь-яку відповідь і з ймовірністю 25% добитися успіху.

У другому ж випадку, тобто в «Своїй грі», ні логіка, ні удача не допоможуть перемогти - допоможуть лише пам'ять і знання.

Застосовувані тестові завдання називають:

- в першому випадку - завданнями закритої форми;
- в другому випадку - завданнями відкритої форми.

Треба сказати, що, згідно з математичної теорії ймовірностей, наведену вище величину - 25% - можна зменшити, тобто зменшити ймовірність простої удачі. Наприклад, чотирма способами:

- збільшення кількості варіантів виконання завдань;
- розміщення серед продемонстрованих варіантів не однієї, а декількох вірних відповідей, які потрібно вказати всі;
- створення дистракторів;
- застосування технологій психофізіологічного контролю.

Однак, навіть при застосуванні цих способів, при тестуванні:

- все одно зберігається, хоча і зменшується, ступінь можливості досягнення успіху за допомогою простої логіки, оскільки суб'єкт тестування бачить одночасно всі варіанти і має можливість їх порівняти;
- багато праці укладачів тестових завдань виявляється марною (іноді навіть понад третину), у зв'язку з наявністю в результатах тестувань певного високого відсотка успіху, списаного «на удачу»;
- значно «роздувається» обсяг паперових бланків, і, відповідно, збільшується ймовірність збоїв при перевірці;
- збільшується непродуктивний час знайомства випробовуваних із завданнями, відповідно має бути зменшено кількість завдань у зв'язку з об'єктивно накопиченою втому і обмеженим часом сеансу.

У технологіях автоматизованого бланкового тестування зазвичай закладений саме варіант вибору зі списку (завдання закритої форми) з його головним недоліком - можливістю побачити одночасно всі правильні та неправильні варіанти відповіді і скористатися цим, проявивши не знання, а вміння логічно мислити.

На погляд організаторів таких тестувань, всі зазначені вище недоліки тестів із завданнями закритої форми окупаються простотою обробки результатів.

Для «обману» людської логіки творці бланкових тестових матеріалів придумали безліч прийомів. Ці прийоми дозволяють конструювати варіанти відповідей на питання так, що невірна відповідь здається правдоподібною, а вірна - неправильною. Сукупність цих прийомів іноді називають теорією дистракторів.

Творці цих прийомів, подібно майстрам рукопашного бою, зберігають їх у таємниці.

Найдивніше, те що нечасто вдається зустріти серед педагогів бажаючих пізнати таємниці «теорії дистракторів». Не дивлячись на те, праця по створенню тестових матеріалів стає в Україні все популярнішою. Український викладач сам собі голова. Говорять, що українська система освіти така хороша і

така магічно стійка та не зруйнована реформи, бо влаштована вона на дивовижних традиціях:

- не вірити авторитетам;
- вчити так, як прийде в голову в даний момент, не озираючись на поурочні плани;
- підкорятися наказам начальників «по-італійськи»;
- очікувати від учнів лише найкращого, підозрюючи гірше;
- сподіватися на «авось».

Процедура масового бланкового тестування слабо захищена від підробки. Одна справа, коли педагог проводить тестування для полегшення власної праці і сам спостерігає за виконанням процедури, і зовсім інша - коли тестування ведеться великим колективом людей, які незацікавлені в якості результатів. Або, ще гірше, які зацікавлені лише в успіху «вибраних» суб'єктів тестування. Такі явища можна побачити, наприклад, при проведенні Централізованого тестування організованого по всій території України вже декілька років.

Можливих шляхів підміни у бланковому тестуванні досить багато:

- «вибраного» учня можна разом із завданням помістити в окремий кабінет, де він все правильно виконає під наглядом, наприклад, репетитора;
- завдання може виконати репетитор без участі самого суб'єкта тестування і потім положити бланк в загальну пачку;
- в процесі виготовлення бланки проходять через сотні рук і очей, тому тестові ключі «просочуються» на ринок. Їх можна купити заздалегідь. Тестові ключі «пливуть» одночасно з тоннами бланків у регіони України.

При бажанні та досвіді можна придумати ще більше варіантів тіньових шляхів до успіху при бланковому тестуванні.

Переклад тестувань на комп'ютеризовану основу здатний різко підняти об'єктивність результатів. Над вирішенням цієї проблеми працює багато вчених та організацій. Наприклад - Центр тестування МОНУ, де об'єднуються кращі тестологи України. Тим не менш, незважаючи на всю серйозність даного міністерства пов'язані з людським фактором недосконалості бланкової

технології ось уже кілька років зберігаються і тут. Вони охороняються нормативними документами.

Далі будуть детально розглянуті відомі технологічні особливості, що дозволяють забезпечити рівень інформаційної безпеки масових комп'ютерних тестувань на рівні закладів, територій, регіонів, держав.

3.4 Переваги комп'ютерного тестування

Поява персонального комп'ютера, а також супутніх йому принтера і сканера, призвело до значної модернізації бланкової технології та різкого підйому масовості бланкових тестувань. Комп'ютер став використовуватися для підготовки (роздрукування) бланків і для статистичної обробки інформації, яку випробовувані заносили у бланки у вигляді хрестиків, галочок і навіть літер. У порівнянні з ранніми бланковими технологіями, де результат оброблявся вручну, знизилася:

- трудомісткість підготовки матеріалів та обробки результатів;
- ймовірність внесення випадкових помилок в результат.

Разом з тим підвищилися вимоги до акуратності суб'єктів тестування, оскільки:

- програмне забезпечення не завжди вірно дозволяло ідентифікувати нанесені олівцем у бланку знаки;
- помятість бланка могла призвести до його деформації в зчитуючому механізмі та появи додаткових огріхів.

Важко пригадати, кому першому прийшла в голову думка про тестування ЗУН, минаючи бланки, а лише за допомогою комп'ютера. Швидше за все, це відбулося одночасно в багатьох країнах. Так само, як і думка про комп'ютеризацію ігор. Геніальна ідея комп'ютерного тестування ЗУН для педагогіки настільки ж важлива, як і винахід парової машини для техніки.

За останні два десятиліття у світі окреслилася тенденція до переходу від бланкових технологій до повністю електронних технологій тестування. Комп'ютерні технології тестувань дозволили «забути» про недоліки бланкових технологій. За допомогою комп'ютерної техніки виявилось можливим:

1. Швидко порівнювати результати для великої кількості випробовуваних на великих територіях.

2. Впроваджувати і розробляти нові технології адаптивного тестування.

3. Підвищити диференціюючу здатність тестів.

4. Оперативно використовувати методи статистичної обробки для підвищення рівня об'єктивності результатів.

5. Об'єктивізувати результат вимірювання шляхом застосування нових способів демонстрування тестових завдань, наприклад:

- виключенням можливості для суб'єкта тестування бачити всі варіанти відповідей на завдання закритої форми і порівнювати їх логічним перебором;

- заміною фіксованих варіантів наборів завдань випадковими вибірками;

6. Оптимізувати тривалість тестових вимірювань.

7. Усунути деякі можливості, що існували в бланкових технологіях, для виникнення випадкових і технічних помилок;

8. Усунути необхідність перевезень бланків та пов'язані з цим втрати, запізнення, зовнішні впливи;

9. Усунути деякі можливості підробки;

10. Знизити кількість персоналу, необхідного в момент проведення заходів;

11. Оперативно стежити за обсягом та якістю ЗУН великого контингенту учнів, зберігаючи при цьому всі індивідуальні характеристики, підібрані для кожного.

12. Вести моніторинг якості освіти в масштабі країни, використовувати його результати для цілей атестації установ та працівників системи освіти.

13. Збільшити швидкість пошуку і надання завдань відповідно до алгоритмів адаптивного тестування.

14. Спростити зберігання, сортування і формування великих тестових масивів у відповідності до алгоритмів.

15. Збільшити швидкість обробки результатів.

16. Реалізувати можливість автоматичної масової корекції параметрів завдань.

17. Реалізувати можливість інтерактивної адаптивної взаємодії системи з претендентом.

18. Спростити введення алгоритмічних обмежень на дії претендента;

19. Підвищити технологічність проведення тестувань і знизити рівень вимог до персоналу центру тестувань.

20. Спростити ведення архіву тестувань.

Застосування комп'ютерних технологій може цілком позитивно вплинути на рівень якості української освіти в найближчий час.

3.5 Загальні переваги і недоліки тестових технологій

З впровадженням тестових технологій в педагогіці з'явилася можливість використання точних статистичних методів аналізу якості, що дозволяють підвищити об'єктивність суджень про те, якою мірою зусилля викладачів та учнів досягають мети.

Впровадження тестового контролю:

- веде до вдосконалення прийомів навчання, дозволяючи оперативно оцінювати їх результативність на основі об'єктивних критеріїв;
- перетворює педагогіку в (таку, що допускає вимір) точну науку;
- спричиняє появу нових виховних і навчальних методик;
- перетворює викладача з транслятора знань у розробника нових програмно-педагогічних засобів, у організатора самостійного навчання.

Як бланкове, так і комп'ютерне тестування ЗУН нерідко піддаються критиці з боку педагогів. Виготовлені в Україні тестуючі програми мало відрізняються від тих, що пропонують закордонні виробники. До основних недоліків автоматизованих систем тестування часто відносять:

- двозначність (не багатозначність) логіки варіантів відповіді в найбільш популярних тестових завданнях закритого типу;
- неможливість контролю навичок усного мовлення;
- жорсткість контролю, яка призводить до психічного напруження та наднормативної втоми претендентів;
- неможливість діагностики навичок спілкування (комунікації);

- неможливість діагностики винахідницьких якостей;
- неможливість діагностики оригінальності мислення у вирішенні навчальних проблем і завдань;
- неможливість врахування всіх випадковостей та факторів, які завжди врахує викладач-професіонал на іспиті - настрій, втома, темперамент, вік, стать;
- неможливість діагностики асоціативного та образного мислення, здатності до навчання, бажання навчатися.

Тому спільно з використанням тестової форми контролю необхідно практикувати і традиційні форми - семінари, конференції, диспути, обговорення, ділові ігри. Загальноприйнято, що оптимального поєднання форм контролю можна досягти так:

1. За допомогою комп'ютерних тестів проводиться попередній контроль (залік) знань визначень і відомостей, навичок вирішення завдань, навичок сприйняття мови на слух, грамотного написання слів і формул і т.д.;

2. Зустріч учнів з викладачем після залікових випробувань, де вони демонструватимуть вже не формальні знання, а навички комунікації та усного мовлення, гостроту і оригінальність мислення, здатність до навчання.

Більш докладно проблеми двозначності логіки комп'ютерних тестерів і тренажерів ми розглянемо нижче. Поки ж наведемо такі міркування. Оскільки творцями сценаріїв комп'ютерних тренажерів, переважно, були і залишаються чоловіки, вони привнесли в практику створення властивий собі логічний стиль спілкування, де відповіді на питання можуть бути лише правильними або неправильними.

У реальному навчальному процесі, як правило, такої ситуації не буває. Навіть однозначну відповідь учня грамотний учитель здатний ранжувати за п'ятибальною шкалою, спираючись на свій досвід спілкування з ним і в залежності від контексту уроку.

Не буває однозначної наукової істини. Не буває абсолютно точних відповідей на всі випадки життя. Так влаштоване життя. Грамотний фахівець, винахідник, вчений може з'явитися лише в такому навчальному середовищі, де

логіка спілкування багатозначна (з відповідями, наприклад: «так», «ні», «може бути», «не зовсім точно», «не цілком»), де є суперечності в знаннях, де є місце для фантазії. Це одна з причин, за яких сучасна, створена за традиційною логічною схемою комп'ютерна навчальна програма не зможе його навчити і виховати.

Парадоксальність ситуації із застосуванням багатозначної логіки в комп'ютерному навчанні та контролі пов'язана ще і з наступним: проводячи урок, учитель, без будь-яких зусиль задає класу запитання, припускає багатозначні відповіді на питання, розпізнає і оцінює ці відповіді, а висловити їх письмово - може тільки з великими труднощами.

Створювати тест, що містить завдання з багатозначними питаннями і відповідями, виявляється, дуже складно!

Звичайно, відомі технологічні розробки, присвячені застосуванню техніки для контролю знань, не вичерпують всіх можливостей комп'ютера. Можливості комп'ютера в сфері освіти нам ще належить відкривати і вивчати.

3.6 Культура та філософія адаптивного тестування

Розвиток кожної науки супроводжується філософським осмисленням. Тестологія - наука молода. Особливу увагу в ній приділяється адаптивним технологіям, які не тільки оптимізують процес тестування, а й зближують процеси створення і вимірювання освітньої продукції. Нижче наведена сукупність формулювань, що дозволяє зробити висновок про сучасний рівень розвитку філософського осмислення розвитку тестології.

Сутність, способи, форми та зміст комп'ютерного дидактичного оцінювання рівня навчальних досягнень, інтелектуальних умінь і практичних навичок культурологія вивчає у зв'язку з відповідями на питання:

- Який сенс комп'ютерного тестування?
- Чому потрібен перехід до адаптивного оцінювання? та інші.

Дефіцит культурних підстав у традиційних системах контролю навчальних досягнень пов'язаний з мізерним урахуванням рефлексії особистості. Цьому раціональному представленні відповідає культурологічний підхід, що пов'язує

міру і критерій прогресу в сфері освіти з масовою і об'єктивною оцінкою особистості.

Філософія культури адаптивного тестування (КАТ) - це система методологічних і теоретичних знань, основним змістом якої є:

- закономірності встановлення дійсного стану навченості суб'єктів тестування;
- правила і стандарти продукування культурних дидактичних систем;
- встановлення механізмів функціонування і розвитку систем тестування.

На перший план філософія КАТ висуває проблеми духовного і соціального змістів діяльності, гуманізму та моральності.

Філософія КАТ формується під впливом двох підходів: методологічного та технологічного. Об'єктами методологічного підходу є концептуальна модель тестування, система правил конструювання феноменів культури, стратегія оцінювання та подання результатів тестувань, конкретні форми подання тестових завдань, припустимі межі класифікації та атестації випробуваних і т.д.

Емпіричний та теоретичний підходи до проектування інтелектуальних систем тестування розрізняють за методами. У першому випадку розробник оперує змістовними поняттями в образі ідеальних об'єктів, що володіють набором ознак.

Концепція культурології тестування розуміється як стислий опис соціальної інформації, яка створюється, перетворюється, накопичується, передається та модифікується за допомогою створених учасниками освітнього процесу знакових та інструментальних засобів, що утворюють культурний простір.

Елементарним феноменом тестової культури (ЕФТК) вважається тестове завдання, яке створене із застосуванням правил та стандартів, ранжовано за ступенем складності, має задане значення валідності та надійності, характеризується універсальними, загальними і специфічними рисами.

Універсальність відбувається тому, що в культурі проектування тестових завдань існує стійкий порядок, утворений основними положеннями тестології.

Спільні риси, характерні для тестових завдань, полягають у тому, що тестові твердження з певної галузі знань, створені різними тестологами, пов'язані з культурою мислення розробників та можливістю їхнього спілкування в процесі конструювання.

Має місце культурне різноманіття в різному поданні одного і того ж змісту навчального матеріалу з допомогою унікальних будівель тестових тверджень. Це несе в собі творчий потенціал, характерний для мислення кожного розробника проблемної тестової ситуації. Проте головним тут є застосування стандартних схем конструювання, що надає завданням специфічних рис та робить їх такими, що легко впізнаються. Впізнаваність форми тестового твердження має принципове значення, оскільки вона сприяє концентрації уваги саме на змісті проблемної ситуації.

Поняття стандартизованої форми завдання включає такі зв'язки і відносини між елементами, які характеризуються ознаками об'єктивності та повторюваності. На відміну від методологічних правил, стандарт не має винятків і його діям підпорядковані проблемним тестовим ситуаціям з будь-якої галузі знань.

Культурний простір тестування являє собою відкриту систему, що розвивається на основі об'єктивних соціальних законів, форму інтроспекції випробовуваних, спосіб їх раціональної організації.

Наявність банків тестових завдань (БТЗ), що охоплюють окремі напрями або спеціалізації, - одне з джерел різноманіття культурних форм тестів, що надає їм забарвлення різних наукових шкіл. Неповторність кожного БТЗ означає, що відносно навчального змісту БТЗ створені в різних навчальних закладах, еквівалентні.

Є більш розвинені, більш потужні і менш розвинені культурні форми і культурні системи. У рамках культурного простору створюється можливість безперервного взаємозбагачення територіально роз'єднаних культурних систем, синтезу різних культурних форм та модифікацій елементарних феноменів тестової культури.

У цьому просторі можна виділяти координати знання, цінності і регулятиви, тобто вимоги та правила, за якими учасники будують свою поведінку. Ці координати утворюють три площини, що перетинаються - духовну, технічну та соціальну, на які «проектується» зміст культурного простору.

Тестологія характеризується комплексом різноманітних змістів -

- духовних (наприклад, в особистісних планах);
- технічних (наприклад, у використанні тих чи інших інструментальних середовищ);
- соціальних (наприклад, у відносинах з інформаційними службами).

Предмет культурології адаптивного тестування включає: вивчення об'єктивних закономірностей розвитку культурного простору стандартизованого комп'ютерного адаптивного оцінювання рівня навчальних досягнень у будь-яких аспектах довільної творчої діяльності, створення і реалізація символічно відмічених і суспільно значущих для всіх учасників процесу тестових випробувань цінностей і змістів.

Об'єктивність оцінки досягається:

- однаковим змістом норм і правил для всіх претендентів;
- застосуванням єдиної автоматичної шкали оцінювання;
- можливістю подання змісту ТЗ в різних формах;
- стандартизацією тестових матеріалів, технології та умов;
- усуненням впливу чинників, що заважають і артефактів;
- науковою обґрунтованістю структури культурних форм тестів;
- валідністю БТЗ;
- захистом від несанкціонованого доступу;
- достатньою величиною генеральної вибірки ТЗ для формування множини культурних форм тестів;
- застосуванням механізмів адаптації;
- жорстким зв'язком між результатом і змістом.

З філософської точки зору процес тестування не дозволяє досягти абсолютно точного вимірювання рівня досягнень. Це пов'язано з наступними причинами:

- не всі поняття областей знань, можна подати у вигляді ТЗ;
- будь-яке спостереження за поведінкою об'єкта не може служити доказом встановлення рівня його абсолютної навченості;
- кінцева кількість ТЗ накладає обмеження на точність оцінки.

Об'єктом культурології є учасники створення моделей, методологічних правил і стандартів, культурних систем і артефактів, що поєднують різні феномени культурного простору в цілісну систему.

В якості власних програм культурологія виділяє наступні підрозділи:

- історію становлення культури адаптивного тестування;
- філософію культури адаптивного тестування;
- методологічну і прикладну культуру дидактичного оцінювання досягнень.

Культурний простір КАТ являє собою впорядковані елементарні феномени і форми культури, виконані із застосуванням методологічних правил і стандартів, що зберігаються в культурних системах і утворюють цілісну структуру завдяки наявності інтелектуальних інструментальних засобів і систем комунікації.

РОЗДІЛ ІV. ТЕХНОЛОГІЇ МАСОВОЇ КОМП'ЮТЕРНОЇ АТЕСТАЦІЇ ЗУН. ЦІЛІ, ЗАВДАННЯ, ХАРАКТЕРИСТИКИ, ЕТАПИ, ШЛЯХИ, РЕЗУЛЬТАТИ

Тестовий контроль - надійний і об'єктивний спосіб оцінювання. Він орієнтований на якомога більш точне вимірювання ступеня (рівня) засвоєння понять та розділів навчальної програми, умінь, навичок.

В Україні педагогічне та психологічне тестування довгий час було заборонено. Історія виникнення ідей тестування сягає корінням у педологію, наукову течію, що виникла серед педагогів на початку ХХ століття і офіційно затавроване в СРСР як «лженаука» подібно кібернетики та генетики. Зараз Україні доводиться доганяти світ.

4.1 Мета і завдань педагогічного тестового контролю

Процедура атестації є важливим елементом освітнього процесу. Переходячи до автоматичної атестації, мають на увазі багато цілей, які, в кінцевому рахунку, зводяться до однієї головної:

Мета педагогічного тестового контролю:

- опосередкована підтримка та підвищення рівня якості освіти.

Кожен учасник навчального процесу, а також споживач освітньої продукції, використовуючи тестовий контроль, або виконуючи тестовий контроль, вирішує досить різні завдання. Перерахуємо деякі з них.

Завдання тестового контролю, як ефективного інструменту управління якістю, з точки зору керівника навчального закладу це:

- підвищення рівня якості навчального процесу;
- звільнення викладачів від рутинної праці (ні один викладач не зможе так, як це зробить комп'ютерний комплекс, за 15-20 хвилин поставити всім учням групи по 40-50 запитань і об'єктивно оцінити відповіді);
- об'єктивна диференціація абітурієнтів, як мінімум, на три категорії: на тих, хто вміє, і хто буде вчитися, на тих, кого не слід вчити, і на «напівпрохідних»;

- об'єктивна диференціація учнів на тих хто успішно завершили навчання, другорічників, і тих, кого слід виключити з навчального процесу;

- у випадку функціонування в навчальному закладі кількох навчальних програм, а також інших нестандартних ситуаціях (відрядження кращих учнів на олімпіаду, підготовка до складної атестації установи) вирішення завдання тонкої диференціації учнів;

- об'єктивне виявлення вад навчального процесу (недбальство викладачів, неузгодженість навчальних планів та інші);

- підвищення аудиторного потенціалу, зниження неефективного аудиторного навантаження викладачів;

- зниження потоку апеляцій;

- вимірювання рівня освітньої підготовки учня з дисципліни;

- діагностика досягнення цілей навчання, встановлених відповідними нормативними документами;

- забезпечення привабливості навчального процесу за рахунок мотивації, спонукаючи та підтримуючи бажання учнів підвищувати якість особистої освіченості, зниження аудиторного навантаження;

- залучення контингенту учнів з числа зайнятих на виробництвах;

- генерація та подання рекомендацій щодо управління навчальним процесом;

- зниження собівартості навчального процесу;

- підвищення рівня престижності навчального закладу і привабливості його для інвестицій.

Завдання тестового контролю з точки зору викладача:

- полегшення рутинної праці з перевірки стандартизованих знань. За допомогою комп'ютера можна посилити контроль не тільки знань фактів і визначень, але і навичок вирішення типових завдань;

- прогноз успішності проходження учнями вихідного тестового контролю, отримання оперативної інформації про вади в якості навчальних досягнень учнів;

- підвищення рівня якості навчання (за рахунок, наприклад, включення елементів гри та змагання) шляхом застосування автоматизованого контролю.

Завдання тестового контролю з точки зору учня:

- виявлення вад у власних знаннях і навичках, дати рекомендації;
- вирішення психолого-педагогічних проблем (неприйняття особистості викладача, боязнь екзаменаційної лотереї та необ'єктивності викладача та ін.)

Завдання тестового контролю з точки зору керівника управління освітою:

- вирішення проблеми контролю якості освіти;
- об'єктивізація атестації закладів освіти та їх працівників;
- об'єктивізація порівняльної картини якості освіти у всіх установах, моментальний знімок, «зріз».

Завдання тестового контролю з точки зору батьків учня:

- усунення екзаменаційних стресових ситуацій з життя дитини;
- підвищення комфортності навчання;
- спрощення профорієнтації (оцінки проф. придатності) і прогнозу напрямів майбутнього навчання шляхом об'єктивного співставлення рівня досягнень учня з дисциплін та окремих розділів з рівнем досягнень колег;
- зниження вартості навчання.

Завдання тестового контролю з точки зору роботодавця майбутніх випускників:

- об'єктивне ранжування майбутніх співробітників за рівнем ЗУН з метою полегшення вибору наданого ним робочого місця, а також мотивованого відмови в наданні робочого місця;
- відкриття відомостей про результати тестового контролю учнів вузів - майбутніх випускників для спрощення пошуку нових молодих перспективних співробітників;
- об'єктивна орієнтації у виборі гідного вузу в якості базового (наприклад, для масової підготовки фахівців у регіональному проекті) шляхом знайомства з результатами порівняльного моніторингу якості освіти.

4.2 Функції ОІТ контролю навчальної діяльності

Функції автоматизованого контролю:

- Інтегративна - сприяння виробленню уявлень про науку як єдиний комплекс знань і методів;
- Дослідницька - автоматична обробка інформації про досліджувану систему з метою вироблення рекомендацій щодо управління якістю;
- Інформативна - представлення інформації про педагогічну систему;
- Контролююча - здійснення контролю якості навчального процесу;
- Оцінююча - зіставлення результатів контролю з еталоном;
- Організуюча - представлення автоматизованих засобів якісного і своєчасного виконання навчального плану;
- Реорганізуюча - реорганізація праці педагогів, усунення рутинної праці, реорганізація навчального процесу, збільшення частки занять із застосуванням ОІТ;
- Прогностична - прогноз успішності діяльності учнів;
- Регулятивна - надання рекомендацій з регулювання навчального процесу;
- Стимулююча - заохочення творчої діяльності, спонукання учнів і педагогів до самоосвіти та пошуку нових форм і шляхів навчання.

Час додасть до цього переліку нові функції, які поки що знаходяться за горизонтом нашого сприйняття.

4.3 Види, типи, етапи, вимоги, характеристики комп'ютерного тестування

Контроль засвоєння навчального матеріалу і оцінка ЗУН є складовою частиною системи оцінки якості. Автоматизація контролю тісно пов'язана з автоматизацією всього навчального процесу і забезпечує зворотний зв'язок в автоматизованих системах управління освітнім процесом.

Види автоматизованого контролю:

1. Вхідний контроль призначений для:

- визначення початкового рівня навчальних досягнень учнів з метою відбору кандидатів для подальшого навчання;

- підтримання високого рівня пізнавального інтересу під час навчання на основі визначення стилю навчання, оптимального для студента;

- складання індивідуального шляху навчання.

2. Поточний контроль призначений для:

- з'ясувати рівень засвоєння учнями змісту навчання.
- ступеня готовності учнів до вивчення наступного матеріалу;
- корекції індивідуальних шляхів навчання;
- забезпечення регулятивної функції.

3. Періодичний контроль призначений щоб:

- з'ясувати рівень засвоєння учнями змісту навчання. Проводиться рідше, ніж поточний контроль, охоплює цілі розділи навчальної програми. Основна форма - залікові тести. На додаток використовуються курсові роботи, контрольні-залікові завдання і т. д. Для періодичного контролю, як правило, використовуються ті ж тести, що і для підсумкового контролю.

4. Підсумковий контроль призначений для:

- перевірки якості виконання учнями навчального плану, проводиться у формі іспиту по всій дисципліні;

- вирішення питання про присвоєння випускникам відповідної кваліфікації.

5. Самоконтроль призначений для:

- визначення претендентами успішності свого навчання;
- встановлення претендентами прогалів у вивченні дисципліни;
- корегування претендентами індивідуального стилю навчання.

6. Взаємоконтроль призначений для:

- формування у претендентів оціночних алгоритмів і якостей;
- підвищення об'єктивності оцінювання за рахунок участі в цьому процесі самих претендентів.

Автоматизація дозволяє значно підвищити роль всіх видів контролю, особливо самоконтролю і взаємоконтролю, відкрити нові аспекти їх

застосування. Не всі методи контролю можуть бути в рівній мірі піддані автоматизації, але з розвитком ІТ з'являються нові можливості для цього.

4.4 Характеристики ОІТ атестації ЗУН

Важливими характеристиками процедури комп'ютерної атестації є:

1. Диференціююча здатність;
2. Валідність;
3. Надійність;
4. Обсяг результату (кількість корисних відомостей, які можна з нього зробити);
5. Оптимальність співвідношення часу і точності вимірювань;
6. Оптимальність співвідношення собівартості процедури (матеріали, техніка, оплата праці) та ступеня досягнення мети (обсягу і об'єктивності результатів);
7. Привабливість;
8. Простота (рівень необхідної попередньої підготовки персоналу);
9. Технологічність і простота (зрозумілість) подання результату;
10. Зручність (простота) застосування технології.

Характеристики ці, природно, взаємопов'язані. Досвід показує, що при прагматичному підході організатори тестувань найменше звертають увагу на характеристики № № 3,4,7.

4.5 Загальні етапи комп'ютерної атестації

Процедура комп'ютерної атестації - досить тривалий багатоетапний процес, що складається з етапів:

1. Реєстрація учнів у спеціалізованій БД;
2. Тренувальне тестування з елементами навчання, що включає автоматичну генерацію проміжних результатів та рекомендацій учням;
3. Основне тестування;
4. Статистична обробка результатів груп учнів;
5. Публікація рейтингів;

6. Підготовка атестаційної документації (відомостей) на затвердження викладачу;

7. Підготовка і подання викладачу інформації про якість засвоєння тими, хто навчається, розділів дисципліни;

8. Поповнення банку результатів новими даними;

9. Статистична обробка результатів, у тому числі - порівняння їх з отриманими раніше та паралельно, тобто з результатами тестування інших колективів досліджуваних;

10. Подання результатів у вигляді, що сприяє зростанню рівня якості освіти (у вигляді статей, доповідей, висновків, наказів, звітів тощо).

4.6 Шляхи досягнення мети комп'ютерної атестації

Для досягнення цілей процедури атестації, необхідне проведення певних заходів:

1. Забезпечення валідності тесту;

2. Забезпечення варіативності тесту;

3. Забезпечення поновлюваності бази тестових завдань;

4. Забезпечення максимальної роздільної здатності тесту;

5. Забезпечення достовірності результатів (виняток зовнішніх впливів);

6. Забезпечення повного і доступного подання результатів;

7. Включення елементів ділової гри в процедуру атестації та інтерфейс комп'ютерного тестового комплексу;

8. Забезпечення ефективності сеансу тестування (максимально повний і достовірний результат при мінімальних витратах часу і ресурсів);

9. Включення в процедуру атестації елементів змагання (перманентна публікація списків претендентів, що одержали найвищі рейтинги у пресі та на Інтернет-сайтах);

10. Забезпечення потреби рекомендацій, отриманих учнями та викладачами за результатами тестувань.

Звернемо увагу на пункти № 9 і № 10. У разі масових профорієнтаційних та освітніх заходів, це - публікація результатів, нагородження учасників, надання переможцям престижних робочих місць тощо.

4.7 Типи сценаріїв комп'ютерного тестування

Комп'ютерні комплекси для тестування ЗУН можуть бути побудовані із застосуванням різних сценаріїв. В літературі не зустрічається загальноприйнятої їх класифікації. Наведемо одну з них:

1. «Стандарт». Найпоширеніший, що став стандартним сценарій, де завдання з фіксованого списку подаються на екрані послідовно, всі варіанти виконання завдання видно одночасно, і претендентові в кожному випадку пропонується вибрати один з наведених варіантів виконання. Завдання часто «зважені», рейтинг претендента обчислюється як приведена до 100-бальної шкали сума ваг вірно виконаних завдань. Тут велика ймовірність

- простого списування;
- випадкового вгадування;
- вгадування шляхом порівняння варіантів.

Застосовуються ТЗ І форми.

2. «Стандарт. Випадковий вибір ». Сценарій, ускладнений і відрізняється від «Стандарту» тим, що завдання з фіксованого списку подаються претендентам у випадковому порядку, а їх варіанти виконання також перемішуються при кожному пред'явленні. Такий підхід значно знижує ймовірність

- простого «списування».

Застосовуються ТЗ І форми.

3. «Стандарт. Множинний вибір». Ускладнений, в порівнянні з «Стандартом», сценарій, де потрібно вказати в списку не один, а всі правильні варіанти виконання завдання. Такий підхід значно знижує ймовірність

- випадкового вгадування.

Застосовуються ТЗ І форми.

4. «Стандарт. Множинний випадковий вибір». Ускладнений, в порівнянні з «Стандарт. Множинний вибір» сценарій, де завдання подаються претендентам у випадковому порядку, а їх варіанти виконання також перемішуються при кожному поданні. Такий підхід значно знижує ймовірність

- випадкового вгадування;
- простого «списування».

Застосовуються ТЗ I форми.

5. «Єдиний вибір». Ускладнений, в порівнянні зі «Стандартом», сценарій, згідно з яким разом з текстом завдання демонструється лише один з можливих варіантів виконання. Від претендента потрібно оцінити цей єдиний варіант по двозначній шкалі (правильно - не правильно). У цьому варіанті дуже низька ймовірність вгадування шляхом порівняння, проте велика ймовірність простого вгадування. Для виправлення останнього нестачі застосовується система «штрафів». Такий підхід дозволяє не тільки констатувати знання й контролювати їхню глибину, але і значно знижує ймовірність:

- випадкового вгадування;
- вгадування шляхом порівняння варіантів;
- простого «списування».

Застосовуються ТЗ I форми.

6. «Пошук однозначної відповідності». Сценарій, де від учня потрібно одне - однозначно зіставити терміни (зображення, звукозапис) з двох списків (подібно до того, як у математиці визначається ізоморфну відповідність множин). У такому підході дуже низька ймовірність:

- вгадування шляхом порівняння варіантів;
- випадкового вгадування.

Застосовуються ТЗ III і I форм.

7. «Пошук неоднозначної відповідності». Це ускладнений варіант, де потрібно побудувати таку відповідність, коли одному терміну (зображенню, звукозапису) з першого стовпця можуть відповідати 2-3 терміна з другого, і

навпаки. Такий підхід дозволяє не тільки констатувати знання й контролювати їхню глибину, але і робить нікчемною ймовірність:

- вгадування шляхом порівняння варіантів;
- випадкового вгадування.

Застосовуються ТЗ III і I форм.

8. «Створення повідомлення». Сценарій, в якому від учня потрібно ввести у вказаному місці слово, словосполучення або сукупність цифр і літер - тобто закодувати повідомлення, що є розв'язком поставленої в завданні задачі. Цей сценарій складний як для претендента, так і для автора, оскільки часто такий введений претендентом фрагмент може бути представлений в кількох різних, але абсолютно правильних варіантах. Закодоване повідомлення може вказувати закон взаємної відповідності елементів двох або трьох множин, а також послідовність елементів множини.

У такому підході дуже низька ймовірність

- вгадування шляхом порівняння варіантів;
- випадкового вгадування.

Застосовуються ТЗ II, III, IV форм.

9. «Пошук і виправлення помилки». Сценарій, ускладнений порівняно з восьмим тим, що місце введення послідовності знаків не визначено, і учневі пропонується його знайти (немає явної вказівки на місце в таблиці або тексті на місце розташування помилки). Такий підхід дозволяє не тільки констатувати знання й контролювати їхню глибину, але і робить нікчемною ймовірність:

- вгадування шляхом порівняння варіантів;
- випадкового вгадування.

Застосовуються ТЗ II форми.

10. «Диктант». Розширений і ускладнений, в порівнянні з 9, мультимедійний сценарій, в якому потрібно написати диктант, слухаючи звук у навушниках, або перевести прослуханий текст на іншу мову. Такий підхід позбавлений більшості недоліків, пов'язаних з випадковістю і дозволяє

- контролювати знання правил написання математичних виразів, словосполучень, хімічних формул тощо.

Застосовуються ТЗ II форми.

11. «Подорож». Ігровий мультимедійний сценарій, де для виконання завдання потрібно проробляти складні маніпуляції з клавіатурою та мишею. Наприклад, «Рухатися» по анімованому приміщенню, вводити з клавіатури словосполучення в спеціальних «вікнах» і «розшукувати» об'єкти на підставі інструкцій. Такий підхід дозволяє не тільки констатувати знання й контролювати їхню глибину, але і робить нікчемною ймовірність:

- вгадування шляхом порівняння варіантів;
- випадкового вгадування.

Застосовуються мультимедійні ТЗ II форми.

12. «Визначення оцінки». Сценарій близький сценарієм «Єдиний вибір» та ускладнений тим, що претендент повинен оцінити ступінь вірності єдині пред'явленої відповіді з якої-небудь багатозначної шкали. Наприклад:

- «так, - не знаю, - ні »;
- «вірно, - не точно, - не вірно»;
- «вірно, - не повно, - не точно, - не вірно».

Такий підхід дозволяє не тільки констатувати знання й контролювати їхню глибину, робить нікчемною ймовірність вгадування шляхом порівняння варіантів; Сценарій може бути ігровим, імітувати навчальні дії вчителя та учнів. Застосовуються градуйовані ТЗ I форми.

13. «Вчитель» або «Дидактор». Ігровий мультимедійний напівінтерактивний сценарій, ускладнений порівняно з «Постав оцінку», де претенденту пропонується провести усне опитування у віртуальному класі та поставити оцінки віртуальним учням. Ускладнення полягає в наступному:

- варіанти виконання ТЗ є дворівневими, після вибору варіанта і його оцінки, демонструються варіанти, уточнюючі вже оцінений варіант, які так само пропонується оцінити;

- серед виконуваних завдань присутні дворівневі завдання II форми (відкрита форма).

Сценарій відрізняється вкрай малою ймовірністю «вгадування» вірного варіанту.

Застосовуються мультимедійні дворівневі градуйовані ТЗ I і II форми.

14. «Діалог». Ігровий мультимедійний сценарій, в якому претендентові у визначеному порядку подаються висловлювання персонажів у стверджувальній формі, які потрібно оцінити за двозначною або багатозначною шкалою (відповіді персонажу). Можна використовувати і пряме введення короткої відповіді з клавіатури.

Такий підхід дозволяє контролювати глибину знань, проте є велика ймовірність випадкового вгадування.

Застосовуються ТЗ I форми або градуйовані ТЗ I і II форм.

15. «Адаптивне тестування». Тестування ведеться формально по одному із наведених вище сценаріїв, але послідовність подання завдань залежить від того, як претендент впорався з попередніми завданнями. Психологічний стан претендента оцінюється автоматично, в залежності від оцінки стану змінюється складність і трудомісткість завдань, включаються блоки психологічного розвантаження. Інформація надходить переважно по слуховому каналу. Цей сценарій вважається найбільш перспективним для застосування в практиці індивідуальних і масових тестувань.

Застосовуються ТЗ I, II, III, IV форм.

Багато відомих і поширені БТЗ, практично без корегування, можуть бути застосовані у будь-якому з описаних сценаріїв.

4.8 Запитання теоретику і практику

Зараз навколо проблем технологій тестування сконцентрувалися думки багатьох вчених і практиків - прихильників різних, майже антагоністичних поглядів. Зокрема, в сучасних суперечках піднімаються питання:

1. Чи може (повинен) процес тестування приносити учням задоволення, бути захоплюючим?

2. Чи може (чи обов'язково) результатом тестування стати моральне задоволення?

3. Чи можна (чи обов'язково) публікувати результати тестування?

4. Чи може (повинен) процес тестування носити змагальний відтінок?

5. Чи слід обмежувати час тестувань? А якщо обмежувати, то, який час - загальний час роботи чи час виконання конкретних завдань?

6. Яка оптимальна тривалість сеансу тестування у різних випадках?

7. Чи може (повинен) процес тестування бути повчальним?

8. Чи є тестування обов'язковим учасником освітнього процесу? Який час слід відвести тестуванню в навчальних планах? Як тарифікувати час тестування і час підготовки тестування для персоналу та викладача?

9. Чи потрібні (чи обов'язкові) психологічні розвантаження в процесі тестування? Чи слід організувати ці розвантаження за допомогою комп'ютера, чи комп'ютер на цей час слід вимкнути?

10. Чи може (повинен) процес тестування спиратися на особистісно-орієнтовану сюжетну лінію?

11. Чи слід дозволяти претенденту вносити виправлення у вже виконані завдання?

12. Чи слід виділяти претенденту ліміт підказок?

13. Слід зазначати в процесі тестування учневі на вірно і невірно виконані завдання?

14. Чи може (повинен) процес тестування бути м'яким, який дозволив би учневі самостійно вибирати план дій - послідовність виконання завдань, повертатися і виправляти результати виконання завдань?

15. Чи слід дозволяти претенденту бачити весь спектр можливих варіантів виконання тестового завдання одночасно?

16. Чи можуть (чи повинні) тестові завдання містити елементи сюрпризу, гумору?

17. Який сценарій тестування ефективніший для освітнього процесу в цілому – серйозний чи ігровий?

18. Який тестуючий комплекс краще - той, де результати претендентів зростають з кожним новим сеансом, чи той, де результати залишаються незмінними від сеансу до сеансу?

19. Чи повинен тест, будучи вимірювальним інструментом, бути однаковим для всіх або його слід адаптувати? Як потім враховувати те, що для різних претендентів тест був адаптований по-різному?

20. Чи повинні претенденти тестуватися «всі відразу», чи можна використовувати режим «Живої черги»? Як це вплине на реалізацію дидактичних цілей і завдань?

Відповіді на ці питання досить важливі. Роль тестувань стрімко зростає. Від цих відповідей, без перебільшення, залежить рівень майбутніх фахівців і, відповідно, майбутнє нашої держави.

Мета введення педагогічного тестового контролю в освітньому закладі можна сформулювати й так: підвищити ефективність управління якістю освіти. Це стисле формулювання об'єднує багато аспектів впливу на кадрову політику, на вкладення коштів, на партнерські стосунки тощо. Відповіді на перераховані питання можуть залежати від розуміння того, що таке ефективність управління і в чому вона виражається:

- в малому штаті;
- у малих витратах;
- у високій прибутковості;
- в швидкості навчання;
- у затверженому стандарті навчального процесу;
- у високій варіативності;
- в частці учнів, що перейшли на вищі ступені навчання;
- в частці учнів, що працюють за одержаною спеціальністю.

4.9 Загальні та часткові вимоги до комп'ютерного тестового комплексу

Комп'ютерний тестовий комплекс - це сукупність компонентів:

- база тестових завдань, з яких за будь-яким алгоритмом створюється тест;

- комп'ютерна програма, що реалізує сценарій тестування і забезпечує процес тестування піддослідних в мережі або на одному персональному комп'ютері;

- пакет комп'ютерних програм, що забезпечують збір, статистичну обробку та представлення результатів тестування;

- комп'ютерна база даних, яка акумулює результати тестування;

- необхідні для процедури тестування і не включені в традиційну офісну конфігурацію ПК пристрої (аксесуари).

До тестових комплексів висувають, наступні вимоги:

1. Ефективність (оптимальність):

- точність вимірювання;
- швидкість вимірювання.

2. Диференціююча здатність:

- об'єктивність;
- адаптованість до умов.

3. Мобільність:

- швидкість доставки та розгортання до готовності;
- кількість, доступність, вартість, швидкість надання необхідних для сеансів і їх обробки матеріалів та послуг (доступ до мережі, розмноження інструкції та довідкових матеріалів, додаткові аксесуари).

4. Дружність інтерфейсів:

- користувача;
- керівника;
- обслуговуючого персоналу;
- автора;
- редактора.

5. Якість БТЗ:

- наповненість (кількість завдань);
- повнота (відповідність наявних ключових завдань у всіх розділах дисципліни);

- варіативність, фасетність (кількість можливих варіантів);
- сепарабельність (роздільність завдань відповідно до тем і умов).

6. Діапазон вимірюваних характеристик:

- широта діапазону;
- варіативність характеристик відповідно до умов.

7. Собівартість:

- створення;
- наповнення;
- редагування;
- валідизації;
- обслуговування.

8. Привабливість:

- процедури вимірювання;
- інтерфейсу користувача.

9. Швидкість і простота:

- наповнення;
- редагування;
- корегування шкали і умов завдань.

10. База результатів:

- представництво (наповненість, зрозумілість) результатів;
- конвертованість результатів у зв'язку із змінами системи оцінки.

Таким чином, якість результату тестування, тобто ступінь відповідності його дидактичним цілям і завданням, залежить від багатьох факторів.

РОЗДІЛ V. НЕДОЛІКИ ІНТЕРФЕЙСУ КОМПЛЕКСІВ ТЕСТОВОГО КОНТРОЛЮ

5.1 Групи основних недоліків інтерфейсу

Перерахуємо чотири групи основних недоліків інтерфейсу відомих комп'ютерних тестуючих комплексів.

1. Недоліки інтерфейсу учня:

- відсутність свободи вибору режимів самонавчання (недружелюбність);
- відсутність змагального компоненту (знижує ефективність навчального процесу);

- відсутність елементів сюрпризу і привабливості (нудність);

- одноманітність подання інформації (що викликає ефект «рівної дороги »).

2. Недоліки інтерфейсу укладача тестових завдань:

- складність освоєння;
- великі витрати часу на введення інформації і конструювання завдань;
- складність введення графічної інформації;
- необхідність технічного супроводу, тобто присутності додаткового фахівця;

- відсутність анімаційних засобів;

- складність переведення тестових завдань в електронний вигляд;

- складність редагування тестових завдань.

3. Недоліки інтерфейсу екзаменатора:

- складність освоєння;
- відсутність засобів оперативного контролю успіхів учнів;
- відсутність засобів інтегрального спостереження за успіхами групи в часі;

- складність порівняння результатів тестування груп учнів з різних дисциплін.

4. Недоліки інтерфейсу керівника навчального закладу:

- складність освоєння;

- відсутність засобів систематизації та зберігання результатів тестувань;

- відсутність засобів зіставлення успіхів навчальних груп з різних дисциплін в часі;
- складність оперативного контролю успішності учнів закладу в цілому.

5.2 Зауваження користувачу.

Практика показує, що найменш за все люди, які проводять тестування, звертають увагу на виконання наступних вимог:

- привабливість процедури тестування, психологічно правильної процедури інструктажу тощо;
- дружність інтерфейсу;
- показовість і конвертованість результатів;
- підбір форми завдань (неправильна форма призводить до неправильного висловлювання змісту і розуміння суті завдання претендентами);
- відповідність інструкції формі та змісту завдань;
- конвертованість результатів.

Це, скоріш за все, пов'язано з існуючою серед «зайнятих» людей прагматичністю поглядів щодо необхідності дотримання дидактичних цілей.

Якщо тестування в освітніх установах проводяться не епізодично, і люди, які їх проводять, розраховують на дидактичну корисність заходів, то, крім наведених зауважень, абсолютно необхідно врахувати наступні. Процедура тестування стане привабливішою і кориснішою для системи освіти, якщо надати їй елементи:

- сюрпризу;
- змагання;
- публічності.

Проводьте тестування як тривалу олімпіаду (чвертьфінал, півфінал, фінал тощо). Маленькі і великі успіхи претендентів повинні бути одразу доступними всім учням, батькам, вчителям. Відповідальні працівники адміністрацій в обов'язковому порядку повинні згадувати прізвища переможців тестувань у своїх доповідях і звітах. Портрети кращих учнів та їх батьків повинні висіти в актових залах. Інформація про минулі тестування повинна відображатися в

газетах з обов'язковою публікацією таблиць рейтингів. Для того, щоб переможців було більше, робіть, як у спорті - розширюйте номінації: тривале, швидкісне, тематичне тестування тощо.

ПАМ'ЯТАЙТЕ! Освітня олімпіада важливіша для розвитку освіти на території регіону та держави, ніж спортивна.

У спортивній олімпіаді перемагає сильний і спритний, а в освітній - розумний та знаючий.

5.3 Результати комп'ютерної атестації ЗУН і фактори впливу

Результатом процедури комп'ютерної атестації знань (але не практичних умінь!) є чисельні значення наступних чотирьох параметрів навченості учня:

1. Знання загальних відомостей з дисципліни і вміння серед них орієнтуватися;
2. Знання визначень, фактів, дат, назв, формулювань;
3. Знання способів вирішення та оформлення вирішень типових навчальних завдань, навички роздумувати і доводити;
4. Уміння вирішувати складні навчальні проблеми.

На підставі значень цих параметрів, для кожного учня будується рекомендована оцінка за п'ятибальною шкалою і рейтинг. Після обробки результатів будується таблиця рейтингів учнів у межах групи. Рейтинг, звичайно, - одне число, що не дає уявлення про конкретні речі в структурі знань претендента. Такий рейтинг називається скалярним. Більш точно відобразити ситуацію може векторний рейтинг, що складається, наприклад, з 4-х чисел, кожне з яких відповідає значенню одного із зазначених вище параметрів навченості. Відповідно, пропоновані студенту тестові завдання діляться на 4 групи.

Провівши тестування одних і тих же суб'єктів тестування кілька разів, переміщаючи їх, визначеним і відповідним чином облаштованими перервами на відпочинок, з використанням того ж (такого ж) тесту, можна отримати додаткову інформацію, динамічну:

1. Здатність суб'єкта тестування до самостійного навчання;

2. Здатність суб'єкта тестування до адаптації в режимі тестування;
3. Розвиненість колективізму серед суб'єктів тестування.

Відправляючи суб'єктів тестування на перерву можна дозволяти і не дозволяти їм читати підручники, спілкуватися. Можна варіювати тривалість перерви. Можна перерву влаштовувати в бібліотеці, повної навчальної літератури. Можна варіювати стимул. Варіантів перерви, як бачимо, багато і, відповідно, варіюються результати повторних вимірювань. Це - не менш цінні результати, ніж ті, що отримані в одноразовому вимірі.

Іноді здатність випробуваного до самостійного навчання цінується значно вище, ніж той багаж ЗУН, який випробуваний приніс на сеанс тестування.

Після закінчення сеансу тестування проводяться заходи атестаційного блоку. Серед них:

1. Претендент отримує оцінки своєї підготовки з кожного з параметрів, рекомендуємо інтегральну оцінку і рейтинг за 100-бальною шкалою, а також перелік своїх помилок;

2. Викладач знайомиться з характерними помилками претендентів, які зберігаються у відповідній БД, коригує параметри підсумкової інтегральної оцінки;

3. Керівник установи отримує статистичні дані, що характеризують результат у порівнянні з аналогічними (попередніми та паралельними);

4. Керівництво територіального управління освітою отримує інтегральне зведення результатів, що дозволяє оцінити характеристики якості системи освіти, що придатні для побудови прогнозів;

5. У разі якщо таке ж тестування паралельно проводилося в іншому регіоні або за кордоном, наприклад, оформлене у вигляді міжрегіонального або міжнародного турніру, відповідна комісія порівнює результати, підводить підсумки і нагороджує переможців у міжрегіональному або міжнародному масштабі.

Результати комп'ютерної атестації повинні висвітлюватися у ЗМІ та друкованій пресі з обов'язковим зазначенням переможців і:

- набраних ними балів;
- отриманих ними нагород і привілеїв (від школи і вузів);
- директорів їхніх шкіл.
- їх вчителів;

Загальновідомою повинна бути і адреса центру, де протягом всього навчального року можна спробувати свої сили в порядку живої черги.

Це значно підвищить популярність не тільки тестувань, а й освіти в цілому, як серед учнів, так і серед їхніх батьків. Не зайвим буде відзначити і кращих працівників освіти.

Погодьтеся, освітня олімпіада, з точки зору розвитку регіону та держави, не менш важлива, ніж спортивна. У спортивній олімпіаді перемагає сильний і спритний, в освітній - розумний і знаючий. У відсталих державах культ сили завжди вище культу розуму. Власне, тому вони й відсталі.

5.4 Інтерпретації результатів тестувань

Виділяють 3 типи інтерпретації результатів тестування мовою тестології:

1. Інтерпретація результату тестування предметно-педагогічна;
2. Інтерпретація результату тестування критеріально-орієнтована;
3. Інтерпретація результату тестування нормативно-орієнтована.

Охарактеризуємо їх коротко, за В. С. Аванесовим.

Інтерпретація результату тестування предметно-педагогічна - інтерпретація головним завданням якої є: з'ясувати, які елементи навчальної дисципліни засвоєні суб'єктом тестування. При цьому аналізується виконання великої кількості завдань, порівнюється зміст атестаційних матеріалів з результатами тестування, з урахуванням прийнятих правил і робиться висновок: чи можна атестувати суб'єкт тестування на підставі отриманих даних.

Інтерпретація результату тестування критеріально-орієнтована - виникає при вирішенні завдання атестації, прийому до ВНЗ, профвідбору тощо, де важливо визначити відповідність піддослідних вибраному заздалегідь критерію. Для атестації випускників, їм дають тільки завдання з області

мінімально допустимої компетентності. Особливу увагу звертають на визначення залікового рівня складності. Зміст завдань принципово полегшено. Завдання повинні виконувати всі випускники, допущені до атестації. Такі тести нерідко застосовують органи управління освітою, що стоять перед необхідністю: у короткий час перевірити стан освіти у великій кількості навчальних закладів і не дозволити їм опуститися нижче допустимого рівня. Від піддослідних вимагають виконати невелику кількість завдань, щоб визначити - що випробуваний знає, і що не знає з заданого стандарту. Інтерпретація результатів ведеться працівниками органів управління освітою і тими педагогами, на думки яких управлінці спираються при атестації.

Інтерпретація результату тестування нормативно-орієнтована на інтерпретацію з опорою на прийняті в математиці статистичні показники, такі як середнє арифметичне, процентна норма тощо. Тестування без інтерпретації результатів не має наукового сенсу. Головні питання такої інтерпретації - не «Хто що знає?», а «Хто вище норми, і хто на якому місці?» У цьому суть нормативно-орієнтованої інтерпретації тестових результатів. Для відповіді на ці питання не потрібний тест з великою кількістю завдань. Потрібно з використанням мінімуму завдань отримати максимум дисперсії надійних тестових балів, розрахувати процентні норми і рейтинг.

5.5 Тривалість сеансу тестування.

Одна з основ тестології - ідея створення інструменту для швидкого і відносно точного оцінювання характеристик великих контингентів досліджуваних. Вимога економії часу природна в масових процесах. Напрямок сучасної організації тестового педагогічного контролю – адаптація, яка значно скорочує часу сеансів. Від часу сеансу істотно залежить якість результатів. Кожен тест характерний оптимальним часом сеансу, зміна якого знижує рівень якості вимірювань. Оптимальний час сеансу визначається емпірично. З часом сеансу пов'язано кількість завдань тесту. Тривалість тестування обмежується виходячи з міркувань:

- зручності процедури тестування;

- темпу накопичення втоми досліджуваних;
- досягнення оптимального рівня диференціації претендентів.

Більшість тестів, що застосовуються в педагогічній практиці, є швидкісними.

Часовий режим задається:

- на весь сеанс тестування;
- на субтести;
- на окремі тестові завдання.

Активно обговорюється проблема методів узгодження завдань на швидкість, необхідних для підвищення надійності тестів.

Найвідоміший тест «на час» - американський TOEFL. Обмеження в часі виконання там цілком обґрунтовані. Відомі швидкісні тести для водіїв. Однак те, що добре для відбору пілотів і водіїв, не завжди підходить для атестації ЗУН учня. Сучасні концепції побудови педагогічного процесу вимагають індивідуалізації всіх ланок навчання, в тому числі і контролю.

Там, де тест безпосередньо не пов'язаний з функціональними обмеженнями для видів робіт, немає необхідності вводити фактор часу як лімітуючу характеристику оцінювання діяльності випробуваного. При контролі трудових навичок визначення ліміту часу доцільно, так як виконання трудових операцій має встановлені нормативи. Час виконання завдань важливий при контролі практичних умінь в електротехніці тощо. Але в таких дисциплінах, як історія чи література, введення часових обмежень у багатьох випадках проблематично, тому для подібних тестів слід експериментально встановити такі нормативи, в які могли б укладатися і сангвініки, і флегматики - адже у кожного з них свій індивідуальний стиль діяльності.

Важливий фактор часу при проведенні масових тестувань - це питання економії ресурсів і обмеження рамок дослідження. Не менш важливий він і при проведенні видовищних заходів - вікторин, конкурсів де використовуються тестові підходи. Наприклад - телевізійні шоу типу: «Хто хоче стати мільйонером?» тощо.

Багаторічні експерименти на тисячах школярів, студентів, вчителів показали:

- Школярі: результати 15-ти та 30-ти хвилинних комп'ютерних сеансів, де кількість завдань ставилися як 1:2, відносяться, в середньому, як 1:2. При збільшенні ж часу сеансу до 45 хвилин (з відповідним збільшенням кількості завдань) призводить до відчутного «завалу» залежності на 15-20%. У експериментах використовувалися гомогенні тести з однотипними завданнями на загальні знання і знання визначень;

- Школярі: якщо тести включають завдання на вирішення чисельних або графічних завдань, час сеансу має бути збільшено до 2 годин. Однак при цьому комп'ютер використовується всього кілька хвилин - для реєстрації та введення результату. Комп'ютер тільки заважає випробуванню - шумить, обмежує площу столу, поширює випромінювання. Тому такі завдання краще використовувати в іншому режимі - отримати варіант, виконати роботу у спокійній обстановці в аудиторії, а потім перейти в комп'ютерний клас і ввести результат;

- Студенти: після 40 хвилин сеансу тестування загальних знань, навичок вирішення простих завдань, знань визначень - включається ефект «рівної дороги» і більшість тестованих починають працювати автоматично, не замислюючись і прагнучи до швидкого завершення сеансу. Прагнення до кращої оцінки для них змінюється прагненням до трійки;

- Вчителі: не витримують тривалості сеансу більше 20 хвилин. Відбувається зрив на обговорення завдань, автоматичне «доклацування» тесту тощо;

- Школярі, студенти, вчителі: Описана вище ситуація докорінно змінюється при включенні в сеанс тестування елементів ділової гри, несподіваності, сюрпризу. Наприклад, шляхом застосування незвичних для практики тестування інтерфейсних елементів і формулювання завдань у розмовному тоні з використанням «незвичних для серйозної роботи» словникових обертів. Це дозволяє значно продовжити період активності уваги,

як у студентів, так і в школярів, а отже, підвищити точність вимірювань за рахунок включення більшої кількості завдань.

Таким чином:

1. При використанні комплексів з «прямокутними інтерфейсами» без елементів психологічного розвантаження, при тестуванні загальних знань та знань визначень та навичок рішення простих завдань обмежувати тривалість сеансу:

- для учнів 2-8 класів 10-20 хвилинами,
- для учнів 9-11 класів 20-30 хвилинами,
- для студентів - 30-40 хвилинами;

2. Під час тестування умінь і навичок вирішення складних чисельних і графічних задач доцільно основну роботу проводити у навчальній аудиторії та використовувати комп'ютер лише для введення результатів та реєстраційних даних. У цьому випадку час роботи можна обмежити кількома годинами, тобто періодом настання втоми контролюючого персоналу. Можна навіть влаштовувати великі перерви, але із застосуванням засобів інформаційної безпеки - варіативність тесту, контроль шпаргалок, обмеження спілкування тощо;

3. Можливо, «золота середина» тут: тренувальні комп'ютерні сеанси з використанням завдань, подібних контрольним, слід проводити без жорсткого обмеження часу, а контрольні - обмежувати;

4. Слід ширше впроваджувати в практику комп'ютерних тестувань елементи ділових ігор та психологічного розвантаження. Це дозволяє значно продовжити час сеансу, запропонувати випробуванним більше завдань без побоювання ефектів типу «рівної дороги» і, тим самим, підняти рівень надійності вимірювань та якості освіти вцілому.

Питання про зв'язок інтелекту зі швидкістю виконання операцій при тестуванні не простий і в даний час обговорюється психологами. Щомісяця з'являються нові наукові та методичні публікації на цю тему.

5.6 Кількість завдань у тесті

Від кількості завдань у тесті прямо залежить якість вимірювання, оскільки кожне завдання дозволяє оцінити знання окремого інформаційного блоку навчальної дисципліни.

У тестології немає чітких обмежень на кількість завдань дидактичних тестів: воно доходить в стандартизованих американських тестах до 200 і більше, не обмежуючись будь-то зручним для обробки числом (особливо з появою і застосуванням для тестування комп'ютерної техніки). Завдань потрібно стільки, щоб по-можливості повно відобразити основний зміст діагностованого обсягу знань. (Є. А. Михайлович, 2001).

Досвід показує, що для усвідомленого виконання понад 40 (навіть простих) завдань учню необхідно мати дуже високу мотивацію. Якщо сеанс вимірювання не призводить до можливості отримання учнями поваги, бажаного робочого місця або відчутного призу - можна очікувати, що через 10-15 хвилин сеансу учень переключиться на виконання тестових завдань методом «випадкового тикуну». Зберегти увагу і виконати 100 завдань багатьом не під силу навіть при дуже високій мотивації.

Те, що сказано вище, відноситься до поширених формальних процедур тестування, де застосовуються комплекси з «не відволікаючими» інтерфейсами, демонструються на екрані одноманітні прямокутники з одноманітними текстами. Інша справа - коли сеанс тестування представлений у вигляді ділової гри з сюрпризами, діючими персонажами тощо. У цьому випадку вдається досягти значного продовження уваги і збільшити кількість завдань у сеансі. Ігрові технології дозволяють підвищити рівень об'єктивності вимірів.

5.7 Специфікація тесту

Специфікація тесту є обов'язковою операцією емпіричного аналізу тесту і необхідною умовою створення нового тесту. Специфікація полягає в тому, що будується таблиця, в якій зазначають:

- номери завдань та їх приналежність субтестам;
- спрямованість завдань;

- складність, трудомісткість, вага завдань;
- яку з характеристик кожне із завдань конкретно діагностує.

У специфікації повинні бути сформульовані діагностичні цілі тесту. Спектр діагностичних завдань відображено в номенклатурі і типології тестів і тестових завдань.

При створенні складних дидактичних тестів необхідно скласти більш докладну таблицю, в якій будуть, поряд із зазначеною вище інформацією, перераховані дидактичні цілі (засвоєння формул та алгоритмів, розуміння таблиць і графіків тощо), а також представлені в БТЗ розділи, параграфи, теми. Це пояснить, на які діагностичні цілі тест орієнтований, і що він вимірює.

РОЗДІЛ VI. РЕЗУЛЬТАТИ ПЕДАГОГІЧНИХ ДОСЛІДЖЕНЬ.

6.1 Артефакти

Великий вплив на результати тестового контролю надає так званий ефект «Звикання». Він виражається на початку сеансу тестування, де застосовується тестовий комплекс з незвичним для користувача інтерфейсом: уповільнення реакції, підвищення метушливості, підвищення ймовірності арифметичних та орфографічних помилок, неправильні маніпуляції. Це може бути пов'язано і з тим, що користувач:

- незнайомий з інструкцією і порядком проведення процедури;
- бере участь у тестуваннях часто, але з використанням інших комплексів, і йому потрібен деякий час для освоєння нового інтерфейсу;
- вже тестувався з використанням цього комплексу, але база тестових завдань змінилася - створена іншими авторами, оформлена в іншому літературному стилі або містить незвичні жаргонні вирази.

Виключити вплив «звикання» на результати тестування нескладно – потрібно лише провести попереднє тестування, дати користувачам звикнути. Згідно з дослідженнями для цього достатньо від 5 до 10 хвилин.

Крім цього, результати можуть бути перевернуті за наведеними нижче причин.

- Група користувачів попередньо ознайомила із завданнями тесту. Це може статися випадково, коли їм просто «попалася» книга з подібною інформацією або репетитор провів тренування, спираючись на власну інтуїцію, та навмисно, коли стався витік інформації. В обох випадках на кривій розподілу результатів тестування може «виросити» дивовижний «горб». Методи нейтралізації цього артефакту відомі і будуть розглянуті у лекціях присвячених інформаційній безпеці.

- Тест не варіативний і правила допускають повторні сеанси. У цьому випадку користувач, що пройшов сеанс два-три рази, а також використання тривалих перерв на спілкування з книгою і викладачем, знаходиться у виграшній ситуації. Крива розподілу «оживе», з кожним сеансом в кінці її буде

збільшуватися «горб». Для усунення перекосу, слід поставити всіх користувачів у рівні умови: всі тестуються разом, варіанти тесту змінюються, бажано із застосуванням процедур випадкового вибору. У зростанні «горба» повинні мати можливість прийняти участь всі користувачі. Такий багатосеансний «марафон» може принести і позитивні плоди, природним чином диференціювавши користувачів на «завзятих», «байдужих» і «середняків».

- Серед претендентів виявилася досить велика група, для якої цей тест не призначений. Це можуть бути учні спеціальної школи «з нахилом», або, навпаки, учні школи, де дана дисципліна не вивчалась. У першому випадку «горб» на кривій розподілу виростає в її кінці, в області вищих рейтингів, у другому - на початку, в області «випадкового тику».

- Серед користувачів є нездатні витримати досить тривале напруження протягом сеансу. Збільшення тривалості сеансу діє гнітюче на всіх користувачів і по-різному на ступінь об'єктивності їх результатів. А для забезпечення необхідного рівня якості вимірювань необхідно включення в тест досить великої кількості завдань, на виконання яких необхідний достатній час. Для усунення цього протиріччя є два шляхи: застосування технологій адаптивного тестування, які дозволяють оптимально розпорядитися часом сеансу та застосування ігрових технологій, що дозволяють за рахунок трансформації інтерфейсу підвищити привабливість процедури.

- У формулюваннях завдань допущено використання розмовної лексики, іншомовних або специфічних для якоїсь групи суб'єктів тестування виразів, які підвищують ймовірність неоднозначного сприйняття умов.

- У ряді завдань зустрічаються неоднозначно сприйняті користувачами активні елементи, неадекватно описані в інструкції, а користувачам поставлено жорстку умову - нічого ні в кого не питати під загрозою припинення сеансу. У такій ситуації відбувається зіткнення безграмотності персоналу та низькою фаховою підготовкою укладачів тесту, що реально впливає на розподіл результатів.

- Фахівець, який проводив тренувальні заняття, не цілком адекватно сприймає дисципліну і, в процесі тренування, допустив зневажливі зауваження на адресу укладачів, свої коментарі про якість завдань, свої фантазії відносно вірних і невірних варіантів. Якщо цей фахівець користується достатнім авторитетом серед учнів, він завдасть такими діями шкоди знанням своїх підопічних і об'єктивності результатів тестування. Боротися з цим важко, така ситуація трапляється, наприклад, коли викладачі всю відповідальність за тренування перекладають на інженера комп'ютерного класу, кажучи: «я все одно там нічого не розумію». Для зменшення впливу цього артефакту потрібно поступати за правилом: у процесі тренувань в класі обов'язково повинен знаходитися досвідчений викладач дисципліни.

До чинників, що ставлять під загрозу чистоту тестових вимірювань, відносять:

- реактивний ефект - зменшення або збільшення сприйнятливості досліджуваних до експериментального впливу під впливом попереднього тестування;

- взаємодія вибірки та експериментального впливу;

- умови організації експерименту, що викликають реакцію досліджуваних, яка не дозволяє поширити отримані дані про вплив на осіб, які зазнають такого ж впливу в не експериментальних умовах;

- інтерференція експериментальних впливів, що виникає, коли одні й ті ж випробовувані піддаються кільком впливам, оскільки вплив більш ранній, як правило, не зникає.

Розглянемо ці чинники докладніше:

Реактивний ефект. Проявляється у двох варіантах.

Перший варіант. Учні тренувалися на проходження тестів по-різному. На тренування ходили не всі, а для тих, хто тренування відвідував, різні викладачі акцентували увагу на тестах різного виду. При валідизації абітурієнтських тестів без урахування цього чинника ефект гарантований. Ефект можна контролювати якщо:

- попередньо опитати учнів про тренувальні заняття;
- паралельно проводити візуальне спостереження за поведінкою користувачів і зіставляти дані, що накопичуються за програмою такого формалізованого спостереження з даними анкетування. Не виключено використання відеокамери при умові, що вона буде добре замаскована (інакше отримаємо додатковий «візит-ефект» з непередбачуваними наслідками - від скарг до істерик).

Це дозволить зробити і підвищити точність поправки.

Другий варіант. Це хороший психодіагностичний супровід, який налаштовує обстежуваного на успішну роботу з дидактичним тестом. Варіант зробити такий супровід для учнів звичним - привабливий. Але в контрольних групах можлива відмова від застосування супроводу, що впливає на результат. Після стандартизації тесту - його валідизації, суворої перевірки надійності, коли будуть прораховані всі основні поправочні коефіцієнти (в тому числі на реактивність до психодіагностичного супроводу за даними контрольних груп або ретест, паралельної форми тесту), психодіагностику можна буде спокійно підключати до дидактичного тесту і використовувати як стимулюючі засоби навчання. Адже з позиції гуманістичної педагогіки головне діагностичне завдання - не зловити учня на незнанні, а допомогти йому засвоїти матеріал, розвинути мислення і здібності.

Взаємодія вибірки та експериментального впливу. Тут небезпека викривлення результатів полягає в тому, що експериментальна і контрольна групи, за допомогою яких відбувається валідиція тесту, істотно відрізняються від вибіркової сукупності, для якої він призначений.

Перший варіант. Суть проблеми - в репрезентативності вибірки. Особливо якщо вибірка складається з добровольців. Їх демографічні характеристики, успішність, психологічні особливості можна прорахувати і врахувати. А як бути з тими, хто відмовився бути добровольцем при адаптації тесту? Наскільки їхні особистісні характеристики адекватні учасникам експерименту?

Другий варіант. Розглянемо реальну ситуацію. Тест у вузі проходить адаптацію на контингенті чотирьох груп з двох факультетів, у той час як на цих факультетах існує лише по 4 групи на тому ж курсі, а всього факультетів у вузі 5; вузів ж аналогічного профілю 10. Питання про результативність та обґрунтованість мотивації відбору груп для участі в експерименті неоднозначний. Зміщення вибірки вплинуть на результати валідизації і дослідник ризикує отримати тільки дані про локальну валідності тесту саме для тих груп, які брали участь в експерименті. При валідизації тесту у вузі необхідно або прагнути не до підбору зручних для експериментаторів послушних кураторських груп, в яких працює сам розробник тесту, а до можливо більш репрезентативної вибірки, та проводити масовий експеримент на всьому контингенті курсу, що вивчають в цей час дану дисципліну за даною програмою. Якщо по цій же програмі працюють в інших вузах, то на репрезентативних вибірках бажано організувати перевірку тесту, суворо дотримуючись зовнішніх умов процедури, а потім враховуючи ці дані при зіставленні результатів і підрахунку валідності.

Третій варіант. Можливий «відсів» суб'єктів тестування під час експерименту з експериментальної та контрольної груп. Якщо експеримент ретельно підготовлений і збір даних проходить швидко, то така небезпека для валідизації практично незначна.

Четвертий варіант. Негативна реакція піддослідних на експеримент. Реакція випробовуваних на експеримент може бути викликана будь-яким аспектом процедури. Наприклад, самим фактом запрошення до участі в експерименті, якого багато хто з однокурсників не отримав. Таємнича поява незнайомих експериментаторів лякає учнів і підвищує ймовірність стресових ситуацій. Рекомендується замаскувати експеримент різними засобами: наприклад, не оповіщати про нього, якщо він зовні схожий на звичні заходи. Найбільш складна проблема - забезпечення випадкового відбору для отримання репрезентативної (по відношенню до контингенту, для якого призначений тест) вибірки. Якщо експеримент передбачає індивідуальне навчання (парну роботу

діагноста з випробуванням), то в розподілі класу, навчальних груп по аудиторіях (експериментальна - ліворуч, решта - праворуч) необхідності немає. При бланковому варіанті дидактичного тестування цього можна уникнути, але витрати часу зростуть у таку кількість разів, скільки буде учасників експерименту. Можна, поступово скорочувати необхідну кількість самих експериментаторів - вони стануть швидко інструктувати і оперативніше реагувати на зворотній зв'язок. Але це може стати побічним фактором: так, першого піддослідного інструктували 10 хв, п'ятдесятого - 2 хв, останнього - скоромовкою. Іноді для пом'якшення впливу чинника учням дають змогу записуватися у вузі на різні факультативи, які читаються в різних приміщеннях в одні і ті ж години. Контроль запису дозволяє скласти за цими даними еквівалентні експериментальні групи і нетрадиційно використовувати процес природної випадковості. Найкращим вирішенням вважається використання в якості одиниці випадковості цілих класів - щоб експериментальна і контрольна групи склалися тільки випадковим чином. Слід також звертати увагу на те, як піддослідні сприймають саму ситуацію тестування і завдання різного типу. Ігнорування відносин піддослідних до ситуації експерименту і ступеня сприйняття ними завдання, зазвичай сформульованого в інструкції, веде до помилкової інтерпретації результатів і до низької толерантності розроблення методик. Цілком ймовірно, що учні, які звикли до закритих тестових завдань, будуть гірше виконувати їх, ніж ті, хто не мав справи з тестами. Треба знати аудиторію.

П'ятий варіант. Інтерференція експериментальних впливів. Під інтерференцією в психології розуміється погіршення збереження в пам'яті матеріалу. У контексті експериментальної термінології під інтерференцією маються на увазі незафіксовані в експерименті накладання факторів. Інтерференцію впливів потрібно передбачати і враховувати при аналізі результатів вибірки (випадковості) навчальних груп, що беруть участь в експерименті. Наприклад, проведення експерименту в одному з вибраних трьох вузів може збігтися з впровадженням саме в цьому вузі нової системи

стимулювання успішності. Це суттєво підвищить їх інтерес до тесту як засобу діагностики та коригування власних знань. Або ж, серед груп одного з факультетів, відібраних для апробації тесту, буде активно поширюватися інформація, що створене в місті спільне підприємство за участю великої іноземної фірми збирається оплатити навчання за кордоном переможцям хімічної олімпіади, або ж що ректорат вирішив різко скоротити контингент студентів у зв'язку з відсутністю замовлень на фахівців, і невдахи у тестуванні можуть бути відраховані. Природно, в перших випадках несподіваний позитивний інтерес до тесту експериментаторів викличе здивування, а останній буде стимулювати зростання стресових очікувань, що відіб'ється і на експерименті. У будь-якому випадку попереднє опитування кураторів, доброзичливо налаштована бесіда для з'ясування можливості впливу таких факторів абсолютно необхідні, так само як і гарантія того, що ніяких негативних селекцій для невдах не буде тощо. І, звичайно ж, обов'язкова фіксація всіх «підозрілих» зовнішніх факторів у звіті з висуненням і подальшою перевіркою гіпотез про цей вплив, його причин і можливих наслідків.

6.2 Динаміка результатів педагогічних вимірювань

Прочитавши методичний посібник з тестології, яких чимало, починаєш вірити, що одна з головних характеристик педагогічного тесту - надійність. Тестологи вважають надійним (тобто гарним) тестом той, результати проходження якого не змінюються (або мало змінюються), скільки б разів піддослідних не тестували. Як в техніці: хороша та лінійка, яка не змінює товщину цегли в процесі вимірювання. Нижче показано, що висока надійність - це якраз той ідеал, якого не можна досягти в педагогічних вимірах в умовах масових тестувань.

Основний метод вимірювання надійності – ретестінг (повторне проходження тесту), яке дублює тестування, яке проводять протягом одного чи двох сеансів.

Протягом одного сеансу повторне тестування влаштовують шляхом:

- або дублювання завдань у складі тесту, представлення їх у різній формі і різному формулюванні;

- або поділу тесту навпіл і прийняття твердження, що половинки еквівалентні.

Якщо повторне тестування влаштовують у два сеанси, то застосовують в обох випадках один тест або тести з подібними, зміненими завданнями (варіанти тесту). При цьому групи піддослідних можуть бути різними, але повинні бути «статистично рівними». «Статистичну рівність» встановлюють шляхом попереднього опитування або попереднього тестування.

При повторному тестуванні протягом одного сеансу, учні часто виявляють кмітливість і, наперекір винахідливим укладачам тесту, знаходять дублі - однакові завдання. Далі вони роблять так: якщо є впевненість у точності рішення, обидва рази вибирають один (вірний на їх погляд) варіант. Якщо ні - вибирають різні варіанти. Цей простий алгоритм, який застосовується мотивованими піддослідними, впливає на чистоту експерименту однозначно: значення «надійності» завищується.

Як би не старалися експериментатори дотриматись чистоти експерименту при тестуванні в два сеанси, але все ж:

- учні чомусь раптом починають консультиватися з викладачем, читати підручники і підвищувати свої знання в перервах між сеансами;

- учні виходять після сеансу та розповідають, тим хто стоїть в черзі, зміст завдань та ймовірні відповіді;

- учні навчаються в різних школах і неможливо передбачити заздалегідь, яку з дисциплін вони знають краще чи гірше;

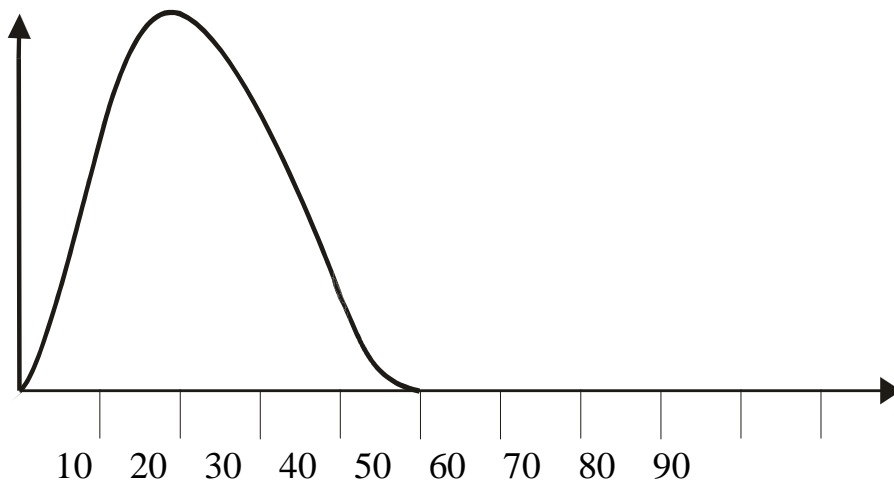
- учням може все набриднути і вони почнуть тестуватися методом «тику» (до речі, найкращий метод отримання надійних результатів).

Технологія тестування, той антураж, який створюється навколо сеансів, обов'язково призводить до посилення або послаблення мотивації піддослідних до успіху. По-іншому не буває, адже тестування - освітній акт. Мотивація

відіграє найважливішу роль у вимірах надійності. Мотивація безпосередньо впливає на величину коефіцієнта надійності.

Якщо піддослідні не мотивовані - результати вимірювання сконцентруються у точки випадкового «тику» області визначення результатів тестування:

Рис. 1.



Скільки б сеансів не проводив експериментатор - результати ті самі й залишаться. Це - тривіальний варіант; тут, після вимірювання, надійність тесту буде дуже високою.

Якщо піддослідні мотивовані, у другому випадку середній результат буде вище. Це сутність людини. Змінити їх важко. Вони - не машини. Однакового результату не виходить. Зупинити зміни результатів не вдається, яким би чудовим не був використаний тест. Проведіть третє тестування, і Ви побачите - результати «пливуть», причому, для різних піддослідних - у різні сторони. Чим вище мотивація піддослідних - тим не надійніший тест. Надійних, в розумінні прямого слідування визначенням, дидактичних тестів не існує.

Парадокс: чим вищі дидактичні якості технології тестування, чим більше вона підсилює мотивацію піддослідних до додаткового навчання, тим вона не надійніша. Тестування, як частина навчального процесу, має нести дидактичне навантаження. Яку технологію тестування вибере педагог: підсилюючу чи ту, що знижує мотивацію учнів?

І що ж - не буває добрих, в традиційному розумінні надійності, тестів? І тестологія - не наука? Або ми не правильно сприймаємо поняття «чистоти експерименту»? А може, тестологічне поняття «надійності» просто не адекватне дійсності?

Розглянемо визначення надійності, дане В. С. Аванесова [Композиція тестових завдань. Навчальна книга. З видавництва., Доп. М.. Центр тестування, 2002р. -240с.]:

«Надійність тесту - показник точності і стійкості результатів вимірювання при його багаторазовому застосуванні. Надійність тесту тим вища, чим «однаковіші» результати його застосування при тестуванні однорідних груп претендентів.

Тепер розглянемо обговорення надійності, дане тим же автором (текст наводиться у скороченні):

Надійність характеризує ступінь адекватності відображення тестом відповідної генеральної сукупності завдань. Раніше передбачалося, що міра надійності являється стійкою характеристикою тесту. На надійність сильно впливає ступінь гомогенності груп випробовуваних, рівень їх підготовленості, а також інші фактори, пов'язані не стільки з тестом, скільки з умовами його проведення. Тому в останні роки стали частіше писати про надійність вимірювання, маючи на увазі тест як результат тестування в конкретних умовах, а не загальну характеристику тесту як методу.» (за В.С. Аванесовим)

Докладно ситуацію з визначенням надійності тесту в тестології описав Є. А. Михайличев [Дидактична тестологія. М.: Народна освіта, 2001. - 432 с.].

Виявляється, недовіра до поняття «надійність» давно існує в середовищі тестологів. Так чи потрібна вона - «надійність»? Може, вона повинна поступитися місцем іншим характеристикам, що більш адекватно відображають властивості тесту?

При уважному розгляді робіт по тестології можна помітити, що самі значення «коефіцієнта надійності», ретельному виміру яких присвячується так багато сил, після вимірювання і обчислення більше ніде реально не

застосовуються. Використовується лише думка, породжена в процесі вимірювання - «цей тест кращий, а той - гірший». Але хіба не можна для зміцнення думки придумати інші, менш спірні критерії?

Поняття «надійність», виявляється, не самотнє, в розумінні нашого критичного розгляду. При повторенні сеансів, так само, як значення коефіцієнта «надійності», «пливуть» і значення ваги завдання. При цьому навіть не важливо - ті ж претенденти прийшли на повторний сеанс чи інші. Важливо лише, щоб сеанси були послідовними, а претенденти мали можливість спілкуватися.

Вага завдань займає одне з центральних місць в парадигмі IRT. За допомогою значень ваг темперують тест, впорядковують завдання за критерієм складності.

Розглянемо реальну ситуацію. Візьмемо тест, складений у повній відповідності з канонами навчальної дисципліни та тестології. Візьмемо дві групи претендентів, які добре підготовлені за двома різними підручниками цієї дисципліни відповідно. Нехай групи будуть різної чисельності. Підручники беремо різні, бо питання дисципліни там висвітлені по-різному, в тому числі й самі підручники недосконалі. Відповідно, частину завдань гарантовано правильно виконають претенденти з першої групи, а іншу частину - з другої. Тоді, обчислена після сеансу вага кожного завдання виявиться залежною від співвідношення чисельності груп. Скільки б ми не міркували про «репрезентативності», від цієї визначеності нікуди не дінешся. Чим менше претендентів виконали конкретне завдання, тим більша його вага. Отже, за умови рівної ймовірності відповідності завдань тому чи іншому підручнику, ймовірність перемоги вища у тих претендентів, які належать до меншої групи.

Вибір підручника для навчання - справа суб'єктивна. Виходить, що значення ваги також суб'єктивне. Але тоді навіщо витратити стільки сил на розрахунок цих значень?

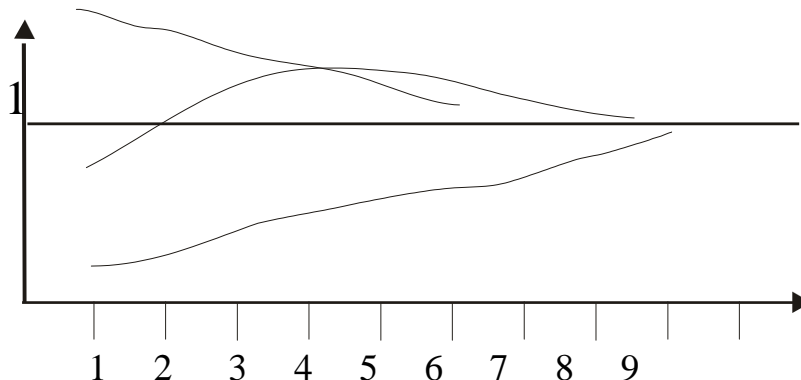
Період обчислень ваг довгий. Поки проходять тестування сотні і тисячі учнів, минають багато тижнів і місяців. Але це не вся складність. За цей час у

навчальних закладах багато чого змінюється, з'являються нові підручники, змінюються навчальні програми, самі навчальні дисципліни стають іншими. «Пливе» ґрунтовність знань учнями тих чи інших розділів дисципліни. При всьому цьому, природно, повинні змінюватися і бази завдань. Разом з базами, повинні змінитися і значення ваг завдань. І це - ще не всі аргументи.

У тестології пояснюють таку ситуацію, тобто результати вимірювань, що «пливуть», «не репрезентативністю» вибірки претендентів. Це можна було б так і залишити, якщо б не було видно явні закономірності «запливів» цих чисел. Можна стверджувати: якщо претенденти мотивовані, а тимчасові інтервали між сеансами досить великі й претенденти мають можливість спілкування, то після кількох сеансів ми побачимо – що:

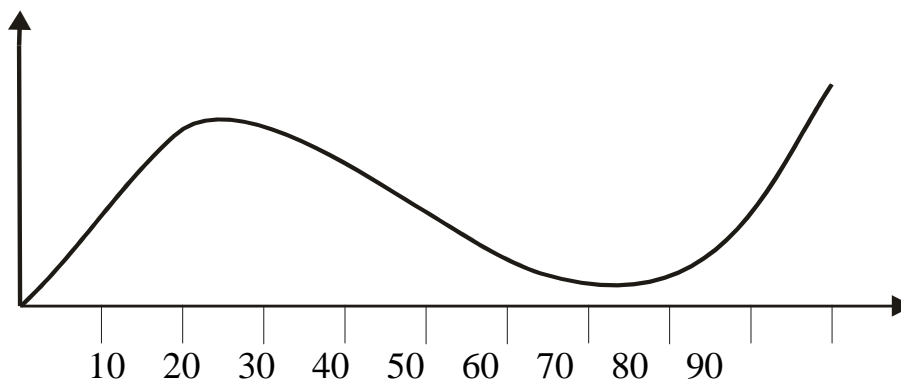
А) ваги завдань намагаються вирівнятись:

Рис. 2.



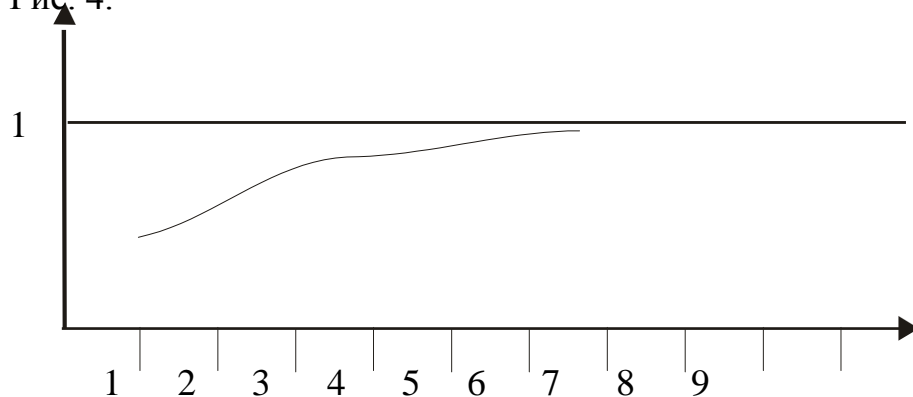
Б) результати претендентів мають тенденцію до групування в двох районах області визначення кривої розподілу - поруч з точкою випадкового «тику» і поряд з межею 100%:

Рис. 3.



В) надійність тесту зростає і поступово стає найвищою:

Рис. 4.



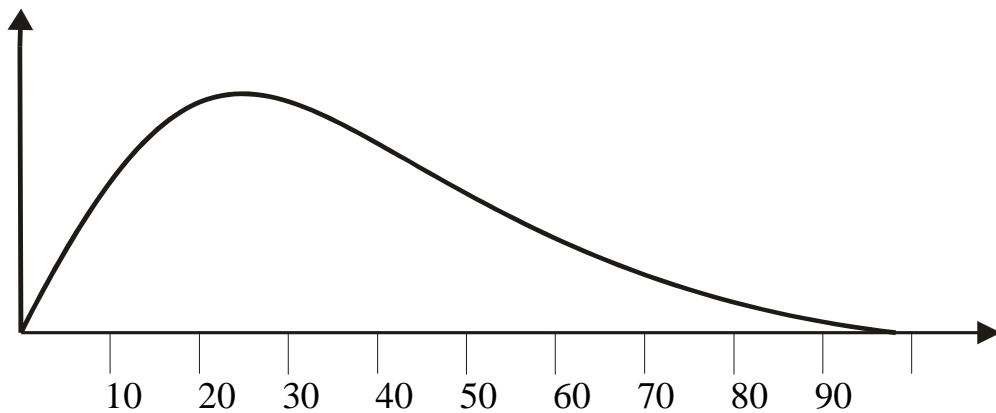
Ваги завдань, за початковим задумом, обчислюються для міфічного «усередненого» претендента. Ситуація з цим обчисленням нагадує притчу про лікарню, де в якості показника успішності роботи обчислювали середню температуру хворих, і середній хворий був здоровий ($T_{CP} = 36,6^{\circ}$). Замість того, щоб порівняти результати учнів двох шкіл між собою та прямо визначити реального лідера, тестологи, в рамках діючої парадигми, порівнюють їх результати з результатом «середнього учня» (який не тільки невідомий, але ще й «тестувався» минулого року). У результаті може статися, наприклад, ситуація, коли якесь легке для «середнього учня» завдання, школярі, які навчалися у конкретного вчителя, не виконали, але це майже не вплинуло на їхній підсумковий результат. І навпаки.

Деякі тестологи, дивлячись, як «плавають» значення коефіцієнтів, придумали вихід для «зміцнення» теорії: взяти і заборонити повторні тестування. Але школярі виявилися хитрішими вчених - вони запам'ятовували завдання і передавали їх, разом з відповідями, своїм колегам, які чекали на сеанс у черзі в коридорі. «Попереду танків по мінному полю йшли тральщики». Кожен наступний сеанс проводився з більш підготовленими претендентами. І ваги попливли. Якби їх значення обчислювалися кожні дві години, то вчені побачили б чудесні метаморфози зі школярами, котрі стають розумнішими на очах, і з одними і тими ж тестовими завданнями, які, протягом дня, стають все простішими і простішими.

Давайте розглянемо динаміку результатів багаторазового проходження одного і того ж тесту, однією і тією ж групою претендентів. Для того щоб, картина не стала тривіальною, нехай тест буде варіативним, тобто його завдання кожного разу будуть змінюватися, але спектр їх типів буде зберігатися. Нехай ваги завдань обчислені заздалегідь так, як це вимагає теорія, і не змінюються.

Після першого сеансу, якщо тест «достить» хороший, а група претендентів «достатньо» представницька й велика, згідно з канонами тестології повинен отриматись наступний розподіл результатів:

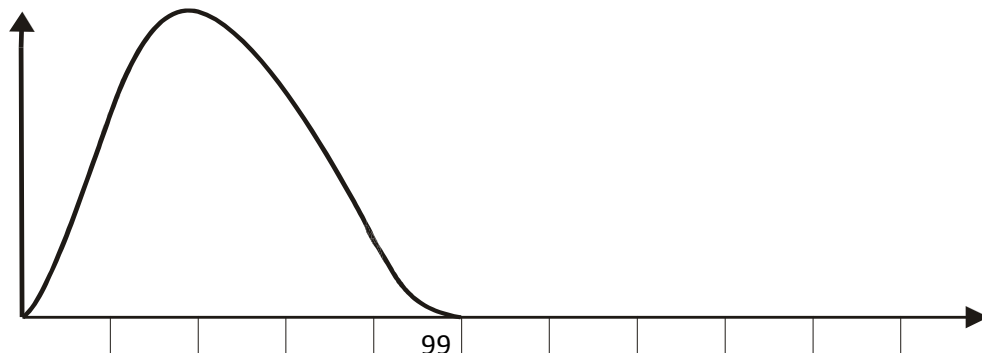
Рис. 5.



Канони тестології (теорії IRT) вимагають, щоб вершина кривої розподілу була одна. Починатися вона повинна від нуля, потім перевалювати через вершину і плавно знижуватися до нуля на межі 100%. Так воно іноді й відбувається. Якщо в групі претендентів 300-400 чоловік і більше, претенденти мотивовані, завдання заздалегідь невідомі, властивості претендентів більшою мірою однорідно розподілені, то графік вийде більш-менш плавним, як на цьому рисунку (Рис. 5.).

Буває й інакше. Скажімо при тестуванні з якої-небудь колективно «нелюбимої» дисципліни або за відсутності у претендентів мотивації:

Рис. 6.

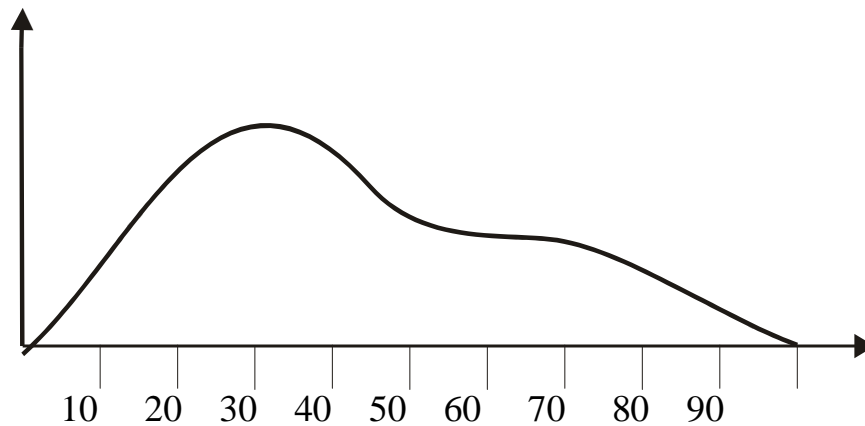


10 20 30 40 50 60 70 80 90

Тут вершина розподілу розташована над якимось середнім значенням, яке отримається, якщо проходити тестування методом випадкового «тику». Ширина кривої в цьому випадку залежить від властивостей самого тесту і може бути обчислена за стандартними формулами теорії помилок.

А буває і так, що до групи претендентів потрапили учні з істотно різним рівнем та спектром ЗУН. Скажімо, коли школярі гуманітарного і фізико-математичного класу одночасно тестуються з тригонометрії. Крива може стати двогорбою:

Рис. 7.



Горб, що ближче до нуля, відповідає гуманітаріям, а інший, подалі - фізикам і математикам. Горби можуть бути розділені різко, а можуть і злитися в широку вершину. Ступінь розрізнення горбів залежить від властивостей тесту та навчальної програми, яку виконали претенденти.

Можна припустити й інші ситуації, коли крива розподілу результатів буде багатогорбою. Тестологи, як правило, «відкидають» такі випадки, кажуть, що вибірка непередставницька, претендентів мало і взагалі - претенденти попалися не ті, що треба.

Причина «любові» тестологів до одногорбих кривих лежить біля витоків IRT - однієї з розповсюджених теорій, що описує результати тестувань. Саме з IRT тестологи «втягують» «технарські» ідеї про надійність, репрезентативність тощо. Саме потоваришувавшись з IRT тестологи починають жити у

фантастичному світі «надійності» та «репрезентативності». У цьому - витoki сюрпризів, які підносить тестологами звичайна «сіра» повсякденність.

Реальність багатша теоретичних уявлень. Рідко кому прийде в голову простежити, як змінюється крива розподілу результатів з плином часу, з кожним новим сеансом. Це - цікаво.

Відбуваються зрушення і їх швидкість залежить від наступних причин:

- ступеня зовнішньої мотивації (якщо претенденти - абітурієнти престижного вузу);

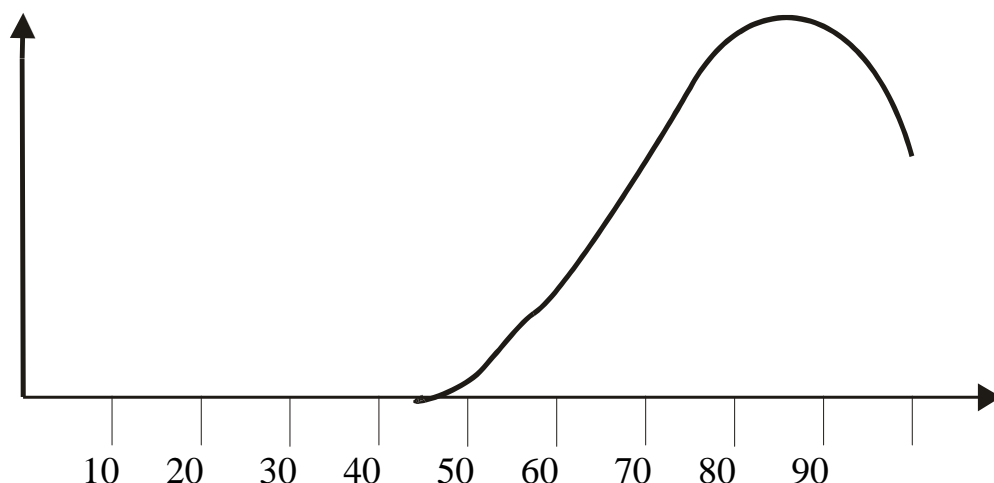
- ступеня внутрішньої мотивації (якщо претенденти стежать за своїми особистими успіхами, намагаються розвивати свій інтелект, прагнуть до знань);

- під впливом мотивації учні оперативно (або не оперативно) консультуються з учителем, читають підручники в перервах між сеансами або ввечері дома;

- при відсутності мотивації учням з плином часу може все набриднути.

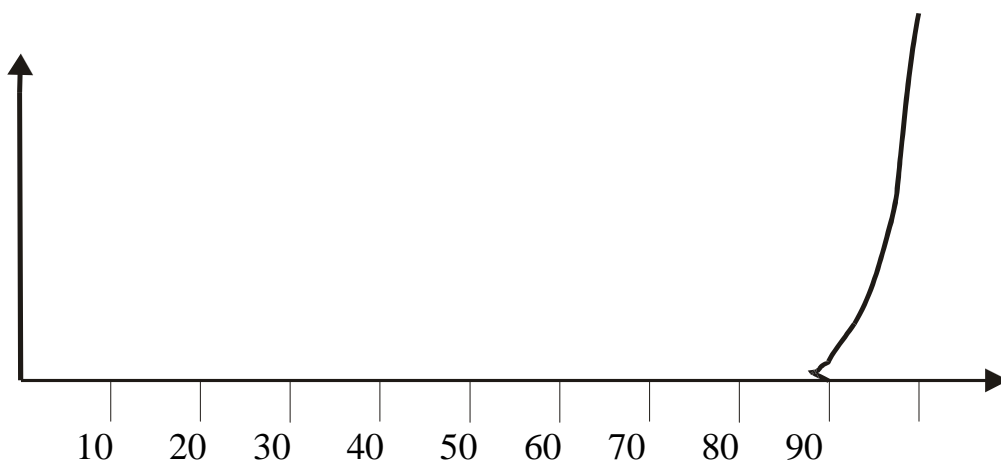
У першому і другому випадку, через 3-5 сеансів крива може стати такою:

Рис. 8.



А потім, ще після кількох сеансів, результати більшості претендентів виявляться сконцентрованими на межі 100%:

Рис. 9.

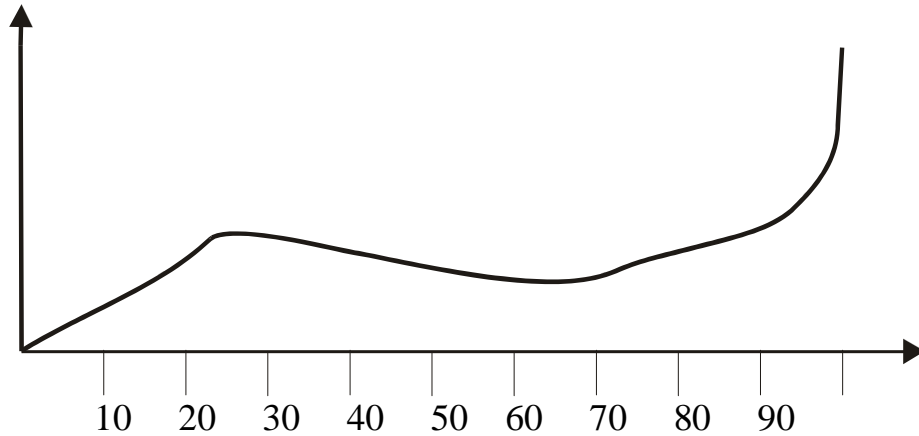


Претенденти - різні. Неоднорідність групи претендентів може проявитися в динаміці кривої розподілу результатів так:

- Претенденти, які втратили інтерес до тестування, не мають достатньої мотивації і просто втомлені покажуть результати поблизу рівня випадкового «тику»;
- Результати сильно мотивованих претендентів, які активно з'ясовували методи виконання завдань і правильні відповіді, виявляться через 2-3 сеанси поблизу відмітки 100%;
- Результати мотивованих, але слабо підготовлених до стресового донавчання претендентів будуть повільно рости, рухатися до позначки 100%.

Картину розподілу тоді можна уявити таким малюнком:

Рис. 10.

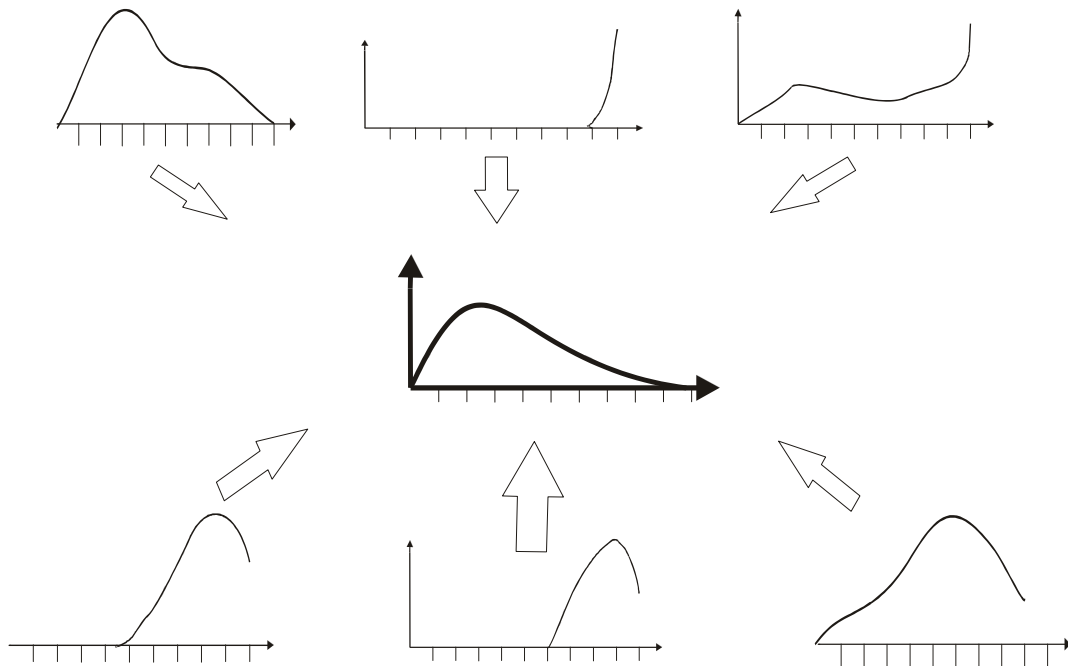


Звичайно, ця крива абсолютно не схожа на ті, що вивчають прихильники ІРТ.

Розгляд, який проведений, буде таким за умови обумовленої вище сталості ваг завдань.

У тестології існують добре розроблені методи перерахунку ваг на підставі отриманих розподілів результатів. Представлені вище криві (Рис.6-10) можуть бути трансформовані до канонічного вигляду (Рис.5) шляхом «перезважування», перерахунку значень ваг:

Рис.11.



Якщо після кожного сеансу тестувань приводити криві розподілу результатів до канонічного виду шляхом «перезважування», то, разом з

«пливучими» кривими розподілу результатів (тепер - повільно пливуть), отримаємо «пливучі» ваги.

Нескладно припустити, куди «пливуть» ваги завдань:

- «зневірені» претенденти натискають клавіші випадково і, відповідно до законів випадковості, всі завдання будуть виконані ними, правильно та неправильно, рівномірно;

- претенденти, що «прагнуть до успіху», рано чи пізно, дізнаються всі вірні варіанти виконання завдань і виконують все вірно;

- «середняки» вносять різнобій в описану рівномірність, однак їх кількість поступово зменшується за рахунок тривалого самонавчання.

Отже, ваги «пливуть» один до одного і поступово стають рівними.

На процес «плавання» можуть впливати артефакти, один з них такий: одна частина учасників повірили своєму вчителю, який навчив їх правильно відповідати на питання тесту, інша частина - сусідові студенту, який вирішив завдання по-своєму. У підсумку, зовсім випадково, частина завдань частиною претендентів буде протягом декількох сеансів виконуватися невірно, тобто виявиться дуже складною для абсолютно конкретної і незмінної кількості претендентів.

Може бути, що ми просто не помічаємо «об'єктивної реальності»? Може, людина не машина, і до неї не можна підходити з «технарськими» мірками і визначеннями? Може, краща тестова технологія саме та, яка чудово виконує дидактичну функцію, де результати претендентів ростуть, і, чим швидше - тим краще? Або та, де ваги завдань, багаторазово перераховуються при багаторазових тестуваннях, стають рівними швидше? Адже головна мета дидактики як раз в цьому і полягає - навчити учнів шляхом тренувань. Чим швидше зростають результати - тим краще застосована технологія (тест + програмний комплекс + сценарій сеансу) пристосована для тренування корпоративних груп та тим вища її дидактична цінність.

Властивість, що відображає швидкість росту результатів претендентів, цілком можна назвати релевантністю, тобто ступенем відповідності тесту дидактичним цілям. А відповідний коефіцієнт - коефіцієнтом релевантності.

Для підвищення надійності тестологи рекомендують провести попереднє опитування претендентів. Можливо повірити, що більшість «претендентів» правдиво відповідатимуть перед сеансом на питання:

- Чи не знає він з будь-яких джерел правильних варіантів виконання завдань?

- Не чи проходив він це тестування раніше і скільки разів?

- Чи буде він проходити тест як слід, або просто потикає в клавіші?

Ні. Вони не зацікавлені в правдивій відповіді з різних причин. Таке опитування має лише умоглядний інтерес, його результату не можна довіряти.

Людина - не машина. Результат вимірювання властивостей людини залежить від застосовуваного інструменту і самого процесу вимірювання. А властивості людини змінюються в залежності від кількості і послідовності вимірювань. Отже, втрачається зміст скрупкульозного вимірювання складності і трудомісткості завдань відповідно до груп претендентів з невідомою передісторією.

Чи є сенс в зміні властивостей інструменту вимірювання відповідно до різних груп претендентів? Так. У всякому разі фізики часто так роблять, вимірюючи властивості різнорідних речовин. Але при цьому повинні бути адекватними процедури перерахунку результатів, чого, як ми бачимо, домогтися досить складно в масових тестуваннях.

Вимірювання надійності в тестології нагадують квантово-механічні експерименти, де прилад неминуче взаємодіє з системою в процесі вимірювання і змінює її. І чим точніший прилад, тим сильніше він змінює систему. Фізики напрацювали непоганий досвід і ним потрібно користуватися. Наприклад, поняття дуальності характеристик і співвідношення невизначеностей, ймовірно, могли б послужити хорошу службу тестології.

Подібно до квантової механіки, тестологічна теорія повинна об'єднати в рамках системи, що вивчається і тест, і технологію тестування, і тих хто проходить тестування, і тих претендентів, які очікують своєї черги, і навіть оточення претендентів (вчителів та інших порадників, продавців шпаргалок тощо.). Не можна адекватно описати поведінки обмеженої системи, розглядаючи лише сукупність не взаємодіючих претендентів і тесту.

Оскільки складність і трудомісткість тестових завдань необхідно враховувати при підведенні підсумків тестувань, а існуючі поняття «надійності» і «ваги завдання» не витримують критики, треба від парадигми, пов'язаної з дивною «надійністю» і частим «перезважування» завдань, перейти до іншої, більш об'єктивної і більш адекватної відповідної реальності.

6.3 ОІТ та менеджмент якості. Принцип безперервного покращення

Прогрес людства прийшов до необхідності використання системи менеджменту якості. Таких систем створено багато. Описів таких систем є безліч, їм присвячено багато томів, які містять в собі схеми підприємств, схеми керівництва, переліки принципів, зразкові інструкції та перелік обов'язків тощо. Це такий зразок, що якщо дотримуватися його можна влаштувати виробництво продукції, та з великою ймовірністю досягнути успіху. Серед загальноновизнаних систем - ISO-9000. Один з принципів ISO-9000 - принцип безперервного покращення, згідно з яким система менеджменту якості вважається хорошою, якщо якість продукції зростає.

У більшості сфер людської діяльності з поняттям якості і вимірюванням якості проблем немає. Інша справа в освіті. У педагогічних вимірюваннях немає місця «надійності» у традиційному технічному розумінні. Результати тестування не повторюються так, як це хотілося б, наприклад, інженеру, який працює з механізмами.

Результати тестування ЗУН, з тією або іншою мірою визначеності, відображають стан якості освіти претендентів. Зростання результатів тестування ЗУН, якщо процес тестування не ускладнений негативними артефактами, свідчать про зростання рівня якості освіти.

Зміни результатів тестувань, проведених з одним і тим же (можливо, варіативним) тестом, можна поділити на загальні та локальні. Під локальними вимірами можна розуміти ті, які відбуваються при багаторазовому тестуванні однієї групи претендентів у рамках однієї теми навчальної дисципліни. Під загальними вимірами можна розуміти ті, які відбуваються при тестуванні різних груп претендентів, або однієї групи, але з різних тем. Локальні зміни характеризують, швидше, особистісні характеристики претендентів, а загальні - навчальний процес. Позитивні загальні зміни свідчать про зростання якості навчального процесу.

Прагнення до «надійності» тесту, рідко призводить до результату, а «надійність», як чисельна характеристика тесту, реально не використовується в тестологічних розрахунках. Логічно замість терміну «надійності» ввести в науковий обіг термін «характеристика», яка відображає ступінь локальних змін результатів тестування. По-скільки зростання освітніх результатів - це дидактична мета, цю характеристику можна назвати «релевантність», розуміючи під цією назвою ступінь відповідності дидактичній меті.

На відміну від «надійності», виміряти «релевантність» нескладно. Наприклад, як різниця усереднених результатів послідовних сеансів тестування однієї й тієї ж групи претендентів, за двома варіантами одного тесту, та в рамках однієї технології.

Позитивна релевантність говорить про зростання результатів, звідси - зв'язок з принципом неперервного поліпшення системи ISO-9000: тест, технологія тестування хороші результати у випадках їх позитивної релевантності. Принцип «позитивної релевантності» краще характеризує тест і технологію тестування, ніж принцип «надійності».

Доцільно застосувати термін «релевантність» не тільки до тесту, 1) але ще до технології тестування, 2) так і до технології комп'ютерного навчання.

6.4 Яким повинен бути комп'ютерний комплекс для навчання і контролю ЗУН?

Якими повинні бути освітні комп'ютерні комплекси, щоб бути потрібними та зручними у користуванні для вчителя, учня, керівника? Як зробити так, щоб проходження сеансу комп'ютерного тестування викликало в учнів зростання інтересу до навчання та до підвищення якості своїх знань? Як «сховати» від учня саму процедуру тестування, зняти атмосферу напруги і хвилювання, сконцентрувати його на розкритті творчих якостей та реалізації знань? Як усунути або звести до мінімуму можливості підробки, списування?

Ділова комп'ютерна гра багатьма сприймається як своєрідна панацея. Можливо, що саме в цій області потрібно шукати шляхи побудови сценаріїв комп'ютерного навчання і контролю ЗУН – привабливих, що викликають бажання досліджувати, вчитися. Вдалі спроби побудови таких сценаріїв уже існують. Можна навести в якості прикладу відомий сценарій і гру Сіда Мейєра «Цивілізація». Однак слід зауважити, що такі ігри зазвичай містять досить мало по-справжньому навчальної інформації і можуть бути розглянуті лише як основа для подальшої методичної і сценарної роботи.

До негативних наслідків таких ігор можна віднести ймовірну неправильну орієнтацію гравців у реальних поєднаннях величин в реальному житті: скільки має бути тих чи інших ресурсів, щоб виробництво інших розвивалося таким-то чином.

Створення професійних комп'ютерних ігор-тренажерів - цілком розроблена і освоєна фахівцями область. Існують прекрасні тренажери і засоби для їх створення - конструкторські програмні комплекси, банки реалізацій. Викликає співчуття та заздрість лише те, що це відноситься виключно до спеціальних тренажерів - для льотчиків, танкістів, судноводіїв, автогонщиків. Ринок же тренажерів для учнів загальноосвітніх навчальних закладів для комп'ютерних лабораторних робіт і навчальних практикумів - поки заповнений виробами порівняно низької якості. Вони малоінформативні, однотипні, не викликають інтересу до навчання.

Як правило, колективи по розробці та виготовленню комп'ютерних тестуючих комплексів складаються з фізиків, математиків, інженерів-комп'ютерників. Педагоги та психологи до робіт практично не залучаються.

Деякі студенти, які розглядають якість навчання як особливу цінність, висловили думку, що обрана форма контролю знань (тренажер «Дидактор») - найкраща з бачених ними, оскільки:

- гра дозволяє продовжити період уважної роботи;
- демонстрація тільки одного варіанту виконання завдання на екрані робить неможливим користування простим логічним порівнянням для відгадування;
- необхідність оцінювання варіантів виконання не по двозначній, а за чотиризначною шкалою, примушує глибше замислитися над навчальним матеріалом.

Такої ж думки дотримуються і багато школярів та вчителів.

РОЗДІЛ VII. ТЕСТОВІ ВИПРОБУВАННЯ І УПРАВЛІННЯ ОСВІТНИМИ РЕЗУЛЬТАТАМИ

7.1 Види мотивації

Одна і та ж група учнів, що проходить одну й ту саму процедуру тестування, може в різний час показати різні результати. Це може бути пов'язано і з поступовим звиканням учнів до технології, і з екологічними умовами, і з іншими факторами, такими, наприклад, як посилення або зниження мотивації під впливом спілкування з педагогами або неконтрольованого спілкування з однолітками.

Саме ця різниця результатів часто не дозволяє вважати той або інший комплекс надійним (згідно з визначенням), не дивлячись на всі його видимі переваги. Це серйозне протиріччя змушує дослідників шукати інші засоби, ніж традиційно зрозуміла надійність, критерії якості тестового комплексу.

Екологічні та погодні умови цілком піддаються контролю. Тому їх вплив на результати тестування може бути експериментально виміряні. Без особливих зусиль можна виключити вплив ефекту звикання, провівши попереднє тренування більш довготривалим у часі.

Зовсім інша справа з мотивацією. Мотивовані люди роблять неймовірні дії, здійснюють подвиги, зрушують гори. Люди, які вміють мотивувати собі подібних і знають, як це робиться, стають відомими і знаменитими організаторами виробництва, педагогами, мільйонерами, тренерами команд, злочинцями.

Результати тестування учнів, які не мають вираженої мотивації до підвищення свого рейтингу, не вийдуть за межі області «найбільш ймовірного» результату. Результати тих же учнів, але які прагнуть отримати приз, очевидно, покращаться.

Мотивація може бути випадковою. Мотивація може бути не тільки позитивною, але й негативною. Наприклад, вона може бути пов'язана з потребою «виділитися» будь-яким шляхом - скажімо, отримавши за підсумками тестування найменший у групі рейтинг.

Виключити вплив негативної і випадкової мотивації можна тільки одним шляхом: дати учням нову, потужну, загальну для всіх мотивацію, яка здатна «перебити» своїм впливом всі інші. Це не тільки виключить випадковості, але і призведе до достовірного отримання «найвищих» результатів, на які здатні учні.

7.2 Мотивація учнів до підвищення якості ЗУН

Мотивація учнів різноманітна. Вона повинна бути відома організаторам тестування до початку тестувань та початку аналізу результатів. Мотивація може бути відкритою(відомою) і прихованою (невідомою організатору тестування). Власне мотивація може бути прихованою і від самих претендентів, в силу якихось психологічних обставин - недостатності навичок самоаналізу тощо. Вплив прихованої мотивації може бути прийнято організаторами тестувань, наприклад, за систематичну помилку процедури вимірювання.

Несподівана поява та несподіване зникнення мотивації може «перетасувати карти» тестологів, які вимірюють надійність і валідність тесту. Мотивацію може створити будь-хто і будь-що. Наприклад, неформальний лідер колективу піддослідних або враження від побаченого кінофільму тощо. Хороший тест буде визнаний непридатним з причини, яка ніколи не буде відомою.

Особливо помітно на результати тестування впливає мотивація, пов'язана з:

- азартом змагання;
- грою;
- прагненням пізнання цікавих відомостей;
- байдужістю до результату.

Причому, гра і змагання тут можуть відбуватися не тільки з колегами, але і з віртуальними персонажами, та навіть з собою.

Для прояву мотиваційних ефектів необхідно, щоб претенденти мали можливість тестуватися кілька разів, а тест повинен володіти варіативністю.

У 1994-1996 році силами колективу Центру нових освітніх технологій ДСДУ були проведені експерименти з виявлення та підвищення мотивації школярів до участі та перемозі у тестуваннях з загальноосвітніх дисциплін. На основі цих експериментів була створена відносно проста технологія мотивації, яка в результаті показала високу ефективність не тільки серед школярів, але і серед студентів та вчителів. Вона названа творцями технологією «м'якого тестування» і описана нижче.

7.3 Технологія м'якого тестування

До недоліків існуючих технологій тестування ЗУН відносять, наприклад:

- двозначність логіки оцінки тестів, що не дозволяє розпізнати, наскільки точно претендент відчуває нюанси визначень і описів явищ;
- відсутність моментів психологічного розвантаження, постійна і невиправдана підтримка у суб'єкта тестування протягом сеансу стану крайньої напруги і відповідальності за кожну неправомірну дію, що приводить до випадкових помилок;
- односторонність рейтингової оцінки, відсутність в результатах тестування детальної інформації про нахили претендента;
- складність визначення ваги завдань, що вимагає попередніх масових експериментів і тимчасових витрат, що виявляються невиправданими після чергової зміни викладача або підручника з дисципліни.

Технологія м'якого тестування (ТМТ) створюється на «противагу» традиційним жорстким:

- коли претендентам забороняють повторне тестування;
- коли завдання можна виконати тільки правильно або неправильно, без диференціації;
- коли у процесі тестування немає місця навчанню;
- коли сеанс тестування нудний;
- коли до успіхів претендентів нікому немає діла;
- коли претендента ведуть тестуватися не тоді, коли він готовий, а коли призначено час сеансу.

Перевести термін «м'яке тестування» на англійську мову можна як «Soft testing» або «Soft grading», відповідно до відомої аналогії (за В.С. Аванесовим).

Творці ТМТ мали вирішити десять наведених нижче завдань.

1. Ввести в ТЗ багатозначні логічні відносини, створити критеріально-орієнтовану техніку оцінки виконання завдань, що включає не тільки полюсні («вірно» і «не вірно») варіанти оцінки, а й більш широкий спектр, у тому числі двовимірну матричну шкалу.

2. Знизити ймовірність логічного вгадування претендентом правильних дій.

3. Ввести в сценарій тестування ігрові моменти, що дозволяли би безболісно збільшити час сеансу і кількість пропонованих завдань, а також здійснили психологічне розвантаження протягом сеансу.

4. Знайти і ввести в сценарій моменти, які спонукали б претендента до самоосвіти.

5. Створити і впровадити техніку підрахунку багатовимірного рейтингу, що виключає необхідність обчислення відносних ваг завдань і дозволяє несуперечливо судити про різні нахили претендентів.

6. Реалізувати в єдиному програмному комплексі можливості однакового створення та застосування ТЗ всіх відомих форм, давши можливість авторів реалізувати їх переваги без залучення додаткових програмних засобів.

7. Максимально спростити працю авторів, переклавши працю з конвертації та сортування ТЗ на плечі комп'ютера.

8. Реалізувати найбільш прогресивні методики забезпечення конфіденційності та інформаційної безпеки.

9. Реалізувати на практиці тестувань переваги мультимедійних технологій.

10. Зробити програмний комплекс мобільним, що дозволяло б негайно приступати до роботи в мережі та на персональному комп'ютері, з компакт-диска і в мережі Інтернет без проходження процедури інсталяції.

Ось основні положення технології м'якого тестування (їх можна вважати вирішенням поставлених вище завдань). Їх десять.

1. Застосовувана у тестуваннях БТЗ повинна бути досить велика для того, щоб за сусідніми комп'ютерами (або повторному сеансі) учасникам не попадалися однакові завдання.

2. Учасникам слід дозволити тестуватися не тільки офіційно, а й у вільний час, відвідуючи комп'ютерний клас «живою» чергою. Існує принцип: «Тестуйся з будь-якої дисципліни, скільки хочеш раз, не ущемляючи цього права своїх колег». Заключна частина цієї фрази виявилася абсолютно необхідною в зв'язку з обмеженістю технічних і часових ресурсів. Всі результати, отримані у присутності відповідальної особи, враховуються. Як підсумковий, зараховується кращий результат. Дозвіл тестуватися в довільний час приводить до підвищення результатів, оскільки навіть на звичайному візуальному іспиті, як відомо, велику роль грає настрій учня.

3. Відмова від сеансу або отримання малої суми балів не призводить до негативних наслідків для учня.

4. Інтерфейс комп'ютерного комплексу повинен бути ігровим, не занудним. Завдання та інтерфейс повинні містити елементи сюрпризу. Спочатку це приваблює, а потім - не відволікає. Це абсолютно необхідна вимога, яка, разом із забезпеченням привабливості процесу контролю, дозволяє

- «відсунути» період початку впливу дії ефекту «рівної дороги»;
- збільшити кількість пропонованих завдань і, відповідно, час сеансу.

5. Час активного комп'ютерного тестування школярів не повинен перевищувати 15-30 хвилин. Кількість завдань не повинно перевищувати 30-60. Це виявлена межа, після якої навіть відмінники і вчителі з досвідом починають тестуватися «методом тику». Якщо завдання короткі, то подвоєння часу тестування від 20 до 40 хвилин не дає якісних змін чи методичних вигод. Найчастіше при такому сеансі просто подвоюється кількість набраних учасниками балів. Тривалість сеансу може відіграти вирішальну роль у привабливості технології, наприклад, сидіння за комп'ютером більше години значно зменшує аудиторію охочих перевірити власні сили.

6. У тому випадку, якщо серед тестових завдань є обчислювальні задачі, то в разі неактивного тестування, час можна збільшити. Важливо, щоб в процесі вирішення завдань учень дивився не на екран (втома зору), а на листок паперу. Краще, якщо завдання учні отримають попередньо, вирішать їх в аудиторії, а потім дадуть відповідь на питання тесту про хід думок та результати.

7. Після закінчення часу тестування, необхідно оперативно проаналізувати результати і голосно, так, щоб це чули в коридорі вчителі та учасники наступної зміни, назвати прізвища школярів, які набрали кращі результати протягом сеансу. Якщо при цьому попросити всіх присутніх влаштувати оплески переможцю зміни, вони це із задоволенням зроблять. Ці оплески - абсолютно обов'язковий атрибут. Це дія, як доведено, «підстібає» і переможців, і переможених. Це характерно в середовищі як студентів, так і вчителів. Такі експерименти проводилися регулярно в різних містах і в різних аудиторіях з однаково позитивним результатом.

8. Тестування слід вести протягом декількох днів. Вранці кожного нового дня на дошці оголошень повинні з'являтися списки вчорашніх учасників, розбиті по класах і дисциплінах та розподілені за величиною набраних балів. Це - найважливіший фактор. Саме ці списки стануть найбільш значущим об'єктом сьогоднішніх обговорень в їдальні, коридорах, класах. До речі, створена в такий спосіб база даних може стати в нагоді приймальним комісіям ВНЗ.

9. На наступний день, після появи списків результатів, у комп'ютерному класі завжди виростає жива черга. Саме ця черга завжди стає своєрідним каменем спотикання. З одного боку - це показник успіху технології, з іншого - присутність в черзі та очікування свого «зоряного часу» для школярів стає значно більш важливим, ніж відвідання лекцій, екскурсій, їдальні. Для «знищення» черги доводиться залучати сторонні сили - чергових вчителів тощо.

10. ТМТ така проста, що важко повірити в її дієвість. Але факти уперта річ. Це експериментально підтверджена інформація, яка говорить сама за себе:

- результати претендентів зростають (в середньому) протягом періоду випробувань при кожному повторному сеансі;

- очікуючи сеанс претенденти уважно вивчають підручники та додаткову літературу, зацікавлено обговорюють і задають пов'язані з навчальними дисциплінами питання вчителям та один одному;

- вчителі, які з'являються поруч з чергою претендентів (спочатку випадково, а потім - постійно), спонтанно влаштовують консультації з усіх дисциплін, так як ці консультації потрібні й бажані.

Так, (п.10) спонтанно створюється продуктивна навчальна обстановка, яку важко зустріти у навчальному закладі в стандартних умовах.

Результати експериментів з ТМТ і тестових комплексів дозволили виробити відносно новий для тестології критерій: технологія організована тим краще, і тестовий комплекс тим кращий, чим вища швидкість росту результатів учасників при повторних тестуваннях. Цей критерій подібний до відомого в статистичній фізиці принципу «виробництва ентропії». Схожість цю підкреслюють щодня отримані криві розподілів результатів тестувань - динаміка диференціації вибірки претендентів.

У 1999 році, коли відзначалося 200-річчя з дня народження О. С. Пушкіна, на базі ВДЦ «Океан» проводилося експериментальне впровадження комп'ютерного ігрового тестера-тренажера «пушкіністика». В один із днів експериментаторам довелося випадково почути розмову бібліотекаря і двох вчителів. Вони, не знаючи, що проводиться експеримент, жваво вихвалялись тим, що в бібліотеці «нарозхват» книги Пушкіна і про Пушкіна, що книги доводиться «притримувати» і дозволяти користуватися ними тільки в читальному залі! Більше того, вони помітили, що творчістю і життєписом Пушкіна абсолютно несподівано і в масовому порядку цікавляться «фізики», «математики», «інформатики».

Підкреслимо, що описане характерно не тільки для середовища переможців районних олімпіад, а й для будь-якої сільської та міської школи, де проводяться тривалі комп'ютерні тестування знань за цією технологією.

Разом з основним результатом експерименту, успішною перевіркою працездатності технології, було проведено вимірювання й отримано такі відомості:

1. Період «звикання» школяра, не знайомого раніше з комп'ютером, до процедури комп'ютерного тестування складає 30-40 хвилин. Іншими словами, якщо у школяра немає підручника, і він користується тільки власними знаннями протягом одного дня, його результати від першого до третього сеансу ростуть, а результати третього і четвертого сеансів однакові. Якщо школяр знайомий з комп'ютером, «звикання» настає після першого сеансу, а результати другого і третього сеансів однакові.

2. У разі якщо школяр має можливість скористатися консультацією вчителя або підручником для самостійного повторення матеріалу, його результати тестування протягом 3-4 днів можуть досягти певного максимуму. Іншими словами, протягом цілком доступного для огляду періоду застосування описаної технології росте і здатність до навчання, і обсяг отриманих учнями знань. Причому, це виходить спонтанно, без застосування будь-яких додаткових стимулів.

Основним результатом описаних тут і багато разів відтворених експериментів є наступний: ТМТ - адекватний інструмент для управління зростанням рівня якості освіти.

7.4 Мотивація студента до підвищення якості ЗУН

Описана вище технологія м'якого тестування «працює» і в студентській аудиторії. Але тут її можна і потрібно видозмінити.

Студент - не школяр. Він не тільки поглинає знання, він ще й змушений шукати майбутнє місце роботи. Він також бажає, щоб майбутній роботодавець також його шукав і вибирав. Давайте поглянемо, як можна в цьому допомогти студенту, а заодно і роботодавцю. А, можливо, вдасться заодно і якість освіти поліпшити?

Зараз все більшої популярності набувають комп'ютеризовані іспити з багатьох дисциплін.

Давайте уявимо собі, що результати тестувань студентів поміщаються на спеціальний сайт в Інтернет, що доступний і студентам, і їхнім друзям, і батькам, і роботодавцям.

Створення сайту - це перший етап впровадження технології.

Напевно, не кожен студент відразу погодиться з широкою публікацією реклами своїх знань. Але це не проблема, можна організувати демонстрацію результатів «по бажанню». Студентів, які заздалегідь шукають контактів з роботодавцями, напевно зацікавить такий спосіб самореклами. Їм не потрібні оплески в аудиторії, їх цікавить власне майбутнє. Студенти ж, поруч з прізвищами яких на сайті немає результатів тестування, швидко переглянуть свої погляди на необхідність публікації. Докази тут прості: якщо результатів немає, значить - нема чим пишатися.

Другий етап впровадження - слід дозволити студентам тестуватися в порядку живої черги довільну кількість разів, публікуючи в Інтернет їх кращі результати.

Третій етап впровадження - залучити до проведених Вами тестувань студентів інших вузів. Навряд чи вони відмовляться прорекламувати себе і свої знання в мережі Інтернет, нехай навіть і за оплату.

Коли в базі даних результатів тестувань з'являться результати тестувань студентів 3-4 вузів, виникне хороша база для робіт з моніторингу якості вищої освіти в місті (регіоні). Публікація результатів моніторингових досліджень якості підштовхне до конкуренції вузів і, в цілому, неминуче призведе до зростання рівня якості.

Це і є той найголовніший результат, до якого ми всі прагнемо - підвищити ефективність управління якістю освіти за допомогою інформаційних технологій.

7.5 Привабливі риси процедури комп'ютерного тестування

У тестологів не прийнято звертати увагу на привабливість процедури вимірювання. Ця особливість перейшла в сучасність з досвіду бланкового тестування, де результат мало залежить від естетичності бланків.

Насправді це досить серйозне питання. Давайте уявимо собі два профтехучилища. У майстернях одного - традиційна затхлість, подряпані пристрої та верстати, вицвілі плакати, розбитий кахель на підлозі. Це легко уявити. У майстернях іншого училища - все навпаки, все як у рекламній телепередачі. Де буде вища кількість бажаючих вчитися, кількість випускників?

Цікавий вплив зовнішнього антуражу не стільки на результат тестування, скільки на спонукання учнів проходити тестування самостійно і неодноразово, прагнучи продемонструвати все більш високий результат. У здатності генерувати прагнення людей до самоосвіти - головна властивість і перевага нововведень у системі освіти.

Привабливість важко і не потрібно описувати числом, різні процедури тестування можна порівняти експериментально, з досить представницькою вибіркою учасників. При цьому слід використовувати один і той же тест. Експеримент слід супроводжувати грамотним опитуванням учасників, в якому буде враховано думку всіх учасників процесу:

- Керівників навчального закладу;
- Виконавців, які наповнювали базу тестових завдань;
- Викладачів;
- Персоналу комп'ютерного класу;
- Претендентів.

Думка претендентів буде багато в чому залежати від того, наскільки вдало пройдений тест. Другий тест вони майже завжди пройдуть більш вдало, оскільки отримали досвід при першому тестуванні. Щоб виключити цей артефакт, слід варіювати послідовність двох процедур тестування для різних претендентів, або розділити тестування на етапи відповідно до процедури.

1. Керівникам навчального закладу сподобається та процедура, після якої їм продемонструють більш представницькі, зрозумілі таблички звіту.

2. Творцям бази даних тестових завдань сподобається та процедура, де їм потрібно менше часу для проведення робіт (за умови однаковості винагороди).

3. Викладачам сподобається та процедура, де б вони менше «псували свої нерви», і де вони отримали більш чітку інформацію про претендентів.

4. Персоналу комп'ютерного класу сподобається та процедура, де групи претендентів були більш дисциплінованими і на підготовку якої (інсталяцію, інструктаж) пішло менше часу.

5. Претендентам сподобається та процедура тестування, де:

- простіше досягти більш високого особистого результату;
- менше доводиться відволікатися на діалоги з персоналом;
- менша тривалість сеансу;
- результати об'єктивні;
- зрозуміле представлення результатів;
- коротка зрозуміла інструкція;
- присутність психологічного розвантаження, що відповідає віку учня і ситуації;
- особисті успіхи відразу стають відомі, а невдачі - ні;
- можна бути успішним як особисто, так і в складі команди, з однієї дисципліни чи комплексно;
- реакція оточуючих на особисті успіхи емоційна і негайна;
- є можливість позмагатися у знаннях з віртуальним супротивником і колегами - учнями;
- в інтерактиві присутній сюрприз;
- менше психологічної втоми накопичується до кінця сеансу;
- менші непродуктивні втрати особистого часу на простоювання в черзі, повторне вислуховування і без того зрозумілих інструкцій тощо;
- менше комп'ютерних збоїв;
- вища якість монітора, миші, меблів, повітря тощо;
- можна повернутися до виконаних завдань і виправити свої помилки;
- учасники знаходяться в рівних умовах (у всіх однакова можливість користування шпаргалками, ніхто нікому не підказує, ніхто не може ззовні вплинути на результат тестування).

Як бачимо, на думку претендентів про привабливість процедури тестування впливає величезна кількість факторів, які можуть бути несподіваними і про які рідко замислюються організатори тестувань.

Ви, напевно, бачили телегру «Хто хоче стати мільйонером?». Що приваблює сюди телеглядачів? Радість за чужий успіх або енциклопедичні відомості, які можна прочитати і без того? Ні! Привертають два моменти: почуття гумору ведучого і несподівані варіанти відповідей – сюрпризи, які зустрічаються зрідка. А що привертає глядача до гри «Поле чудес»? Те ж саме – гумор ведучого і витівки гравців. Хто вигадав ці сценарії? Є підозра, що це – хороші педагоги і психологи. Ці ігри будять бажання змагатися у знаннях, копатися в енциклопедіях.

У цих ігор є аналоги – прекрасні комп'ютерні тренажери. Діти та дорослі з задоволенням працюють з ними. І поодиночі, і групами. Пряма дорога їм – у сферу освіти. Потрібно лише включити туди необхідні для навчального процесу тести. А потім проводити тренаж, скажімо у формі змагань.

Підвищити привабливість комп'ютерного навчального посібника чи процедури тестування не так просто, це завжди – область професійної діяльності. Але гра вартує свічок, оскільки є висока ймовірність позитивного впливу привабливості тестування на якість освіти в навчальному закладі.

РОЗДІЛ VIII. ІНФОРМАЦІЙНА БЕЗПЕКА ПРОЦЕДУР МАСОВОГО КОМП'ЮТЕРНОГО ТЕСТУВАННЯ

8.1 Основні вимоги інформаційної безпеки тестувань та забезпечення

їх виконання

- Секретність баз тестових завдань;
 - Секретність переліку осіб, які мають доступ до тестових завдань;
 - Анонімність осіб, що мають доступ до інформації про тести, в період конструкторських робіт;
 - Заборона несанкціонованого доступу до баз тестових завдань і тестових програмних оболонок під час їх переміщень, установки, зберігання, використання;
 - Заборона несанкціонованого доступу до результатів тестувань під час їх створення, пересилання, обробки та зберігання;
 - Недопущення використання несанкціонованих довідкових матеріалів;
 - Застосування спеціальних методик розміщення претендентів в аудиторіях, яка виключає можливості взаємного підказування, підглядання тощо.
 - Недопущення несанкціонованих підказок і пояснень з боку персоналу;
 - Недоступність сторонніх осіб до аудиторій, коридорів, туалетів тощо, в т.ч. завчасна перевірка приміщень на наявність шпаргалок;
 - Запобігання та ліквідація наслідків порушення названих вище пунктів.
- Саме ж забезпечення передбачає:
- створення документального супроводу (інструкції, положення, накази);
 - відбір персоналу;
 - навчання та інструктаж персоналу;
 - підвищену ставку винагороди персоналу за забезпечення особливого режиму;
 - проведення інспекцій, в тому числі і шляхом залучення анонімних осіб, котрі будуть працювати «інкогніто»;
 - оснащення технікою відеоспостереження і сейфами;

- залучення сил міліції або позавідомчої охорони;
- заходи по виявленню порушень і адміністративного покарання винних осіб.

8.2 Дивовижні явища з досвіду масових комп'ютерних тестувань

Впровадження технологій тестування в Росії виявилось пов'язаним з низкою обставин, що мають відношення до правових аспектів. З'явилися, наприклад, хакери-«доброзичливці», які намагаються впливати на результати тестування - заліки, іспити, оцінки вступних випробувань.

Хакери змогли вплинути на ритмічність навчального процесу вузів, що беруть участь в подібних експериментах:

- проводилися спроби розкриття засекречених файлів;
- поширювалися комп'ютерні віруси;
- публікувалися відповіді на тестові запитання в мережі Інтернет.

Хакери намагалися розкривати захищену інформацію не тільки «з інтересу», але і з метою наживи. Це є елемент злочину.

Боротися з інформаційним шахрайством правовим шляхом важко, особливо в Росії, де хакерство давно перетворилося на своєрідний національний вид інтелектуального спорту. Однак слід звернути увагу на специфіку системи освіти.

Дослідники зіткнулися з новим, цікавим суперечливим явищем. Разом з активізацією «злочинної» хакерської діяльності, паралельно, незвичним шляхом ініціюється масовий процес підвищення рівня якості освіти:

- хакери самостійно осягають нові, навіть ті, які не вивчаються в вузі, знання в області комп'ютерних наук;
- «споживачі» хакерської продукції цілком свідомо і старанно вивчають «куплені» тексти з відповідями на запитання тестів, причому, вивчають так, як не вивчали ні один з своїх або чужих конспектів.

У згаданих текстах просто міститься повний текст курсу, викладений дещо у іншій формі, ніж у підручнику - у вигляді запитань та відповідей.

Цією позитивною властивістю нового явища, безумовно, слід скористатися в благих цілях, паралельно усунувши можливість виникнення негативних явищ.

8.3 Шляхи забезпечення інформаційної безпеки масових комп'ютерних тестувань

Перерахуємо нетрадиційні шляхи усунення негативних наслідків інформаційних «злочинів» і забезпечення інформаційної безпеки масових комп'ютерних тестувань.

1. Не слід «засекречувати» відповіді на запитання комп'ютерного тесту. Більшу частину текстів бази тестових завдань слід видавати друкарським способом. Це приведе до зникнення об'єкту несанкціонованої купівлі-продажу.

2. Запитань і відповідей в базі тестових завдань повинно бути так багато, щоб неможливо було швидко і повноцінно скористатися паперовою чи електронною шпаргалкою. Для відображення в базі матеріалу обсягом 36 навчальних годин, як показує досвід, досить 300-400 завдань (10 завдань на годину). Зазубрити їх неможливо.

3. Претенденти, які сидять поруч, повинні бачити на своїх моніторах різні тестові завдання. При повторному тестуванні кожний претендент отримує новий варіант. Екзамен - це лотерея; претендент не повинен заздалегідь знати варіанти завдань, які «випадуть» саме йому. Виконання цих умов досягається шляхом випадкової генерації варіантів з бази тестових завдань.

4. Варіанти тестів не повинні повторюватися, однак програма - генератор варіантів повинна складати списки завдань для кожного претендента не абсолютно випадковим чином, а виходячи з принципу рівномірності покриття навчального матеріалу - по-трішки з кожної теми курсу.

5. Слід вживати заходів захисту не тільки на програмному рівні. Наприклад, поточні результати тестування необхідно контролювати в інтерактивному режимі, а програмні модулі повинні зберігатися на фізично незалежних носіях і оновлюватися з періодом меншим, ніж час сеансу. Цим підвищується рівень захисту від несанкціонованого проникнення.

6. Слід періодично оновлювати бази тестових завдань. Можна не тільки доповнювати базу кількісно, але й, наприклад, просто видозмінювати словникові звороти в завданнях. Цим усуваються багато з можливостей застосування шпаргалок.

7. «Змова» претендентів з представниками персоналу, які «закриють очі на шпаргалки», буде виключена, якщо дозволити використання будь-яких паперових навчальних посібників під час тестування. При цьому слід обмежити час сеансу тестування так, щоб було неможливо встигнути скористатися цією інформацією в повній мірі. Цим не тільки усувається можливість «злочинної» змови претендентів з персоналом комп'ютерного класу, а й створюється стимул додаткового вивчення матеріалу саме тих посібників, які рекомендовані викладачем курсу.

8. Контрольне тестування має відбуватися для всієї навчальної групи одночасно. Захист від зовнішніх впливів можна забезпечити присутністю зацікавленого викладача, наприклад того, який буде вести наступні навчальні курси для цих студентів в наступному семестрі.

9. Слід виключати на час сеансу тестування зв'язок класу із зовнішніми мережами.

10. Слід використовувати процедуру випадкового вибору номера комп'ютера і псевдоніма для кожного претендента. Претенденти в класі повинні сидіти у випадковому порядку і мати достатньо довгі псевдоніми, якими вони не зможуть швидко помінятися.

11. Тестування повинне вестися на рейтинговій основі. Це впливає як на якість освіти, так і на встановлення внутрішньої системи контролю. Якщо студенти будуть знати, що кількість п'ятірок і четвірок, які вони можуть отримати обмежена, поступово усунуться підказки.

12. Результати рейтингу (кращі) повинні публікуватися в мережі Інтернет і бути доступними для ймовірних роботодавців випускників, що стане додатковим стимулом до навчання. При цьому відповідно до законодавства,

кожен тестуючий комплекс повинен надати студентові право вибору - дозволити чи не дозволити таку публікацію.

Останній пункт може здатися спірним.

Досвід, накопичений протягом проведення масових заходів у 1994 -2004 рр.. в Приморському краї, свідчить про наступне: публікація результатів тестування у вигляді рейтингових списків призводить до виникнення феномену змагальності, примушує учнів проходити знову і знову сеанси тестування, читати навчальну літературу, консультиватися один з одним і вчителями з питань навчальних дисциплін.

Дозвіл учням безперешкодно проходити тестування знову і знову, в сукупності з публікацією результатів, призводить до:

- зростання результатів;
- зростання популярності тестувань;
- зростання зацікавленості вчителів;
- зростання зацікавленості батьків.

Змагальність може бути підтримана локальними заходами. Наприклад, якщо голосно похвалити одного з учасників за щойно отриманий високий рейтинг, і попросити всіх присутніх поаплодувати йому, відбувається дивне: претенденти починають конкурувати за місце в комп'ютерному класі, придумуючи неймовірні причини, щоб повторити тестування. Такі «локальні» впливи на аудиторію, крім того, знімають частину миттєвої втоми, підвищують увагу.

Не менш цікавий ефект впливу на піддослідних інтерфейсу тестуючої програми. Якщо інтерфейс нудний і одноманітний, увага учасників розсіюється швидше, ніж у разі ігрового інтерфейсу, що поєднує елементи сюрпризу, ігри, новизни. Так само розподіляється і ступінь бажання претендентів проходити тестування за допомогою того чи іншого комплексу. Ігровий інтерфейс спонукає до повторного тестування.

Перераховані заходи не тільки усувають негативні чинники, а й позитивно впливають на рівень якості освіти.

8.4 Канали витоку конфіденційної інформації. Забезпечення секретності БТЗ

Очевидно, учень, який отримав несанкціонований доступ до екзаменаційних тестових завдань до іспиту, має чималі переваги перед іншими. Наявність таких учнів призводить не лише до появи помилкових оцінок ЗУН, але і:

- впливає на об'єктивність результатів регіональних моніторингових досліджень;
- робить необ'єктивним аналіз властивостей тестових завдань в екзаменаційній базі.

Недоступність тестових завдань учням до іспиту - один з важливих компонентів технології контролю ЗУН.

Поширені в Росії комп'ютерні тестові комплекси, що застосовуються для масових тестувань школярів і студентів, не відповідають вимогам таємності. Це пов'язано з наступними обставинами:

- ТЗ складаються викладачами, що працюють в момент складання завдань в оточенні небажаних свідків - колег, друзів;
- Коло укладачів ТЗ мале;
- Кількість ТЗ, що включаються в тест, невелика. Учні, що вже пройшов тестування не складає труднощів передати всю побачену і запам'ятовану інформацію вчителям і друзям, які будуть тестуватися пізніше;
- ТЗ подаються учням, які сидять в комп'ютерному класі в єдиному пронумерованому порядку. Це спрощує створення шпаргалок;
- У комп'ютерних мережевих класах, як правило, розставлені комп'ютери так, що сусіди бачать відразу кілька екранів. Якщо тест представлений одним варіантом, учні, які бачать дії сусідів, мають можливість списувати;
- Неваріативність. Учень, який стане тестуватися повторно, отримає той же варіант завдань, з яким він познайомився кілька годин тому і який він мав можливість обговорити з репетитором;

- Крім авторів ТЗ, до них мають доступ оформлювачі (сканування та включення ілюстрацій, таблиць), програмісти тощо - кілька десятків фахівців. «Витік» інформації від такого величезного колективу важко контролювати;
- Країна «не знає своїх героїв». Договори, які підписані центрами тестування зі своїми співробітниками і між собою, як правило, не включають конкретних умов щодо забезпечення конфіденційності інформації;
- Немає відомостей про те, що когось якимось чином покарали за розсекречення інформації про шкільні або студентські тести.

Перераховані обставини знижують рівень секретності (конфіденційності) тестів. Інформація про завдання майбутніх тестувань, незважаючи на титанічні зусилля Міністерства освіти, задовго до початку заходів стає доступною всій країні, з'являється на сайтах в мережі Інтернет. Покупець не бачить продавця: інформація оплачується охочими придбати тестовий матеріал шляхом електронних переказів і розсилається їм за допомогою електронної пошти. Як кажуть, не купує завчасно ключі шкільних тестів лише ледачий.

Загалом, вважається, що продавець тестових ключів - злочинець. Проте не завжди вдається зрозуміти, як його знайти, але і за що, і як його карати. Загалу не відомі випадки серйозних покарань осіб, що допускають витік конфіденційної інформації із центрів тестування і продавців, котрі торгують ключами шкільних тестів.

Продаж інформації про завдання майбутніх шкільних тестувань перетворилася на своєрідну індустрію. Продаж невірної, підробленої інформації по тестах стала точно такою ж індустрією, і це, як не парадоксально, лише компенсує втрату конфіденційності.

В інших країнах світу ситуація з секретністю тестів вважається іншою. Пакети з тестовими завданнями поширюються по країні під охороною озброєних поліцейських. Відомості по укладачів тестових завдань засекречені, і секрет цей охороняється законом. Розсекречення тестової інформації - злочин, підлягає суворому покаранню у судовому порядку.

Більш того, всі ці заходи - верхівка айсберга. Є ще невидима частина суспільних взаємовідносин, що захищає таємність тестів не гірше озброєної поліції. У США, наприклад, давно вихований середньостатистичний добропорядний громадянин, який, помітивши злочинні дії кого-небудь, негайно повідомляє про це «куди потрібно». Добропорядний американський школяр ніколи не буде списувати на очах у кого-небудь контрольну роботу. І нікому не дасть списати. І сам перший розповість вчителю або поліцейському про те, що хтось із його друзів списує.

Ця «чесність» має, звичайно ж, підґрунтя - систему суспільно визнаних стимулів. Розглянемо її з різних точок зору зацікавлених сторін.

8.5 Стимули збереження конфіденційності тестових матеріалів у США

Стимул № 1

Так склалося, що можливість отримання високої оцінки в американських школах і вузах - запитання конкурсу. Вищих оцінок не може бути більше норми. Учень знає: якщо хтось поруч з тобою списує, значить, він прагне потрапити до списку кращих разом з тобою, а кількість місць обмежена. Кожен учень особисто зацікавлений в тому, щоб вказати вчителю на колегу, що списує.

Стимул № 2

В університетах існує обмежена кількість вакансій для безкоштовного або пільгового навчання. Є система престижних іменних стипендій, кількість яких обмежена. Заняття пільгової вакансії або отримання престижної стипендії - не тільки предмет гордості, але й питання полегшення життя зараз і вдалого старту кар'єри в майбутньому. Отже, кожен однокурсник - твій конкурент, а найпростіший спосіб позбутися конкурента - знайти і надати викладачеві або адміністрації докази його непорядності в навчанні.

Стимул № 3

Адміністрація навчального закладу зацікавлена в якості навчання. Це не лише питання престижу, це питання державне і приватних субсидій. Тому, адміністрація зацікавлена у заохоченні учнів, які виявлятимуть екзаменаційні

підробки, вчинені їхніми друзями та викладачами. На відміну від пострадянської школи, такий «шпигун» в американській школі - герой!

Стимул № 4

Адміністрація штату зацікавлена у підвищенні конкурентноспроможності своїх університетів на рівні держави. Це і додаткова реклама, і опосередковане джерело надходження коштів до бюджету (наприклад, оплата навчання для нерезидентів штату завжди значно вища). Адміністрація жорстко відноситься до фальсифікації. Для осіб, помічених у фальсифікації, значно ускладнюється подальша кар'єра - відомості про них заносяться в спеціалізовану державну базу даних. Адміністрація заохочує кращі навчальні заклади зниженням податків, додатковим фінансуванням тощо.

Стимул № 5

Батьки прагнуть підготувати кращий «плацдарм» для старту кар'єри дитини. Вони зацікавлені у виявленні несумлінних однокурсників дитини, які потенційно можуть її у майбутньому «обійти», і повідомляють про них адміністрації.

Стимул № 6

Батьки зацікавлені не лише в навчанні дитини у більш престижному (і дорожчому - оплата за навчання) ВНЗ, але й у збереженні бездоганності репутації дитини.

«Ставки» високі. Тому батьки самі контролюють процес її навчання з точки зору недопущення непорядності.

8.6 Стимули збереження конфіденційності тестових матеріалів у Росії

Вище ми познайомилися з роботою системи забезпечення конфіденційності тестових матеріалів і, відповідно, захисту якості освіти в США.

Росія - інший світ. Російський студент і школяр завжди готовий допомогти колезі. Ризикуючи власною репутацією і майбутньою кар'єрою, він допоможе колезі за сусідньою партою під час тестування - вирішить його завдання і дасть списати свої.

Іншими словами, піднімаючи якість освіти, росіянам доводиться вирішувати ще й завдання перелому менталітету, ширше вводити в шкільний і вузівський ужиток елементи змагання і конкуренції.

Традиційний для Росії «витік» інформації перешкоджає успішному впровадженню централізованих тестувань. Міністерство освіти робить відчайдушні заходи щодо забезпечення секретності тестів (навіть пакетики для зберігання тестових бланків, і ті замовляє в Німеччині). Однак на практиці всі ці заходи виявляються дорогими і неефективними.

Заходи щодо забезпечення секретності можуть призвести і до зниження якості тестових завдань. Будучи створеним одноосібним автором і зберігаючись в таємниці, вони виявляються закритими для рецензування, редагування та експериментального контролю в референтних групах.

Де ж вихід? Де знаходиться «гордіїв вузол» і як його розрубати?

Вихід є. Політики кажуть, що для того, щоб припинити витік інформації на рівні держави, треба або все засекретити, або - все розсекретити. Простіше і дешевше піти другим шляхом.

Оптимальна технологія забезпечення секретності (конфіденційності) тестових завдань може бути викладена на 1-2 сторінках нормативного документа. Опишемо її у вигляді алгоритмічної послідовності дій.

8.7 Забезпечення конфіденційності тесту в масштабі регіону і країни

На підставі досвіду участі в організації всеросійських централізованих тестувань в Приморському краї («Телетестінг» і ЦКТ), а також досвіду участі в інспектуванні таких тестувань в Республіці Саха (Якутія) і Хабаровському краї, дослідникам вдалося виділити і систематизувати технологічні недоліки і розробити достатньо бездоганну і економічну технологію забезпечення конфіденційності.

У цій схемі поєднані два варіанти сценарію процедури тестування, розроблені для повної та часткової комп'ютеризованої форми тестування.

8.8 Принципи безпечної технології створення БТЗ

1. Список колективу авторів, рецензентів і редакторів тестових завдань є відкритим і може бути опублікований.

2. До електронної бази вводиться не менше 2-3 тисяч завдань з кожної дисципліни, що включають, наприклад, матеріал всіх опублікованих з грифом Міністерства освіти підручників і задачників. Це число обумовлено наступним:

- Саме стільки завдань (разом з варіантами) з кожної навчальної дисципліни можна знайти у відкритій пресі, і щось нове придумати досить складно.

- Якщо середній учень просто вивчить хоча б чверть вірних варіантів виконання такої кількості завдань - він гідний високої оцінки.

- Створити таку базу цілком можливо втрих - учотирьох за 1-2 місяці.

- Така кількість завдань, що супроводжуються ілюстраціями і таблицями, цілком поміститься на одному компакт-диску.

3. Завдання сортовані за принципом подібності (рівня складності, рівня трудомісткості, відповідності одній темі, рівня важливості). Груп подібних завдань для кожної дисципліни вийде близько сотні. У кожну групу увійде два-три десятки рівних за складністю завдань. Процедура створення варіантів ґрунтується на принципі: по одному завданню з кожної групи. Математичне пояснення: нехай тест включає 100 завдань, по одному з кожної групи. У кожній групі нехай буде 20 завдань. Якщо перерахувати всі можливі варіанти поєднань завдань, які створить комп'ютер, вийде 20^{100} . Це астрономічна кількість, що гарантує неповторність варіантів.

4. «Ні» традиційному засекречуванню завдань! Завдання публікуються у вигляді збірок для підготовки. Обсяг такої допомоги складе 500-800 сторінок, тобто може перевищувати обсяг підручника. Страшного нічого не трапиться, якщо школяр вивчить значну частину опублікованої в посібнику інформації. У цьому реалізується головна дидактична мета. Публікація збірок принесе дохід центрам тестування і, в той же час, усуне ґрунт для небажаних тіньових явищ.

5. «Ваги» завдань обчислюються завчасно, шляхом експериментальних тестувань. Тестуються всі бажаючі, експерименти проходять багаторазово протягом року в багатьох регіонах. Це можливо в умовах відсутності секретності, оскільки всі можливі варіанти завдань і запитань заздалегідь опубліковані. Вивчити їх всіх - рівносьильне повторному неодноразовому проходженню курсу навчання.

- Завчасне визначення ваг прискорить процедуру статистичної обробки результатів реальних тестувань, усуне цей загальновідомий об'єкт багаторічної критики.

- Економічну підтримку експерименту можна здійснити введенням оплати участі. Охочих потренуватися буде багато.

6. Алгоритм формування єдиного і неповторного варіанта тесту (списку завдань) для кожного претендента заснований на коді, що включає

- паспортні дані;
- точні дату і час створення коду;
- чотирьохзначне число - ключ, що міняється на єдиному Інтернет-сайті через визначені проміжки часу і транслюється по радіо чи телебаченню одночасно на всій території протягом усього періоду тестування;

- параметри (зі значеннями, наприклад, 0 і 1) відображають бажання або небажання претендента опублікувати результат. Відповідно до обраної цифри, претендент дізнається про свої результати на наступний після тестування день в загальних списках, або – по завершенню всіх заходів з надісланого йому іменного сертифікату.

Блок-генератор варіантів орієнтується на ці дані і дає можливість дозволити претендентам повторні тестування з будь-якої дисципліни, з гарантією неповторності варіантів, у тому числі і платні попередні тренування. Варіант тесту може бути згенерований знову і в іншому місці при повторному введенні перерахованих даних, наприклад, при роботі апеляційної комісії.

7. Алгоритм формування варіанту тесту включає вибір форми представлення варіанту - бланкової (якщо центр тестування не обладнаний

достатньою кількістю комп'ютерів) або комп'ютерної. Вибір здійснює керівник центру тестування. Блок вибору форми дозволяє роздрукувати тест на принтері для кожного претендента. Після закінчення сеансу тестування претенденти повинні ввести результати виконання завдань у комп'ютер.

8. Нормативні документи та інструкції включають офіційний дозвіл претендента тестуватися і тренуватися довільну кількість разів з будь-якої дисципліни протягом періоду тестування, якщо це не зачіпає прав доступу до інших претендентів. Перший шанс (сеанс тестування) надається претендентам безкоштовно, наступні - після оплати.

9. Тестування проводяться на базі спеціалізованих центрів, відповідальні особи яких сертифіковані.

10. Керівники спеціалізованих центрів несуть персональну відповідальність за дотримання правил тестування - забезпечення доступності тестових матеріалів, високу робочу віддачу друкованих пристроїв тощо. Законодавством визначено адміністративну відповідальність за фальсифікацію.

8.9 Десять кроків безпечної технології тестування

1. Затверджується та інструктується персонал територіального Центру тестування. Керує тестуванням спеціаліст, сертифікований Міністерством освіти.

2. Програмне забезпечення, що надійшло завчасно з Міністерства освіти, встановлюється в комп'ютерному класі територіального Центру тестувань. Воно використовується як для тренувальних, так і для атестаційних тестувань. Доступ претендентів на тренування обумовлений інструкцією. Тренувальні та атестаційні тестування проходять по одному і тому ж технологічному плану.

3. Адміністратор мережі Центру тестування регулярно, за встановленим розкладом, отримує по телебаченню, по радіо або з Інтернету ключ, необхідний для формування коду претендентів. Введення числа синхронізовано з часом початку і закінчення сеансів.

4. Початок реєстрації синхронізовано за часом з отриманням ключа. Контроль документів претендентів при реєстрації проводиться незалежним

представником міліції. Незалежний контроль перекриває один з ймовірних шляхів підроблення - підміну претендента.

5. Претенденти, що приступають до виконання, протягом нормативного проміжку часу вводять в комп'ютери свої паспортні дані. На основі цієї інформації формується код претендента і відповідний варіант завдань тесту, який можна виконувати сидячи перед екраном комп'ютера або, якщо час сеансу тестування великий і комп'ютерів мало - в аудиторії, маючи на руках роздрукований варіант і бланки для роботи. Рішення про форму тестування приймає керівник центру на основі відомостей про кількість претендентів, наявності техніки тощо.

6. Оголошується початок сеансу. Виконання завдань в аудиторіях центру тестування і в комп'ютерному класі контролюється черговими і представниками вузів регіону. Контроль в даному випадку спрощено, оскільки:

- варіанти завдань у всіх учасників різні;
- вага й обсяг шпаргалки, наявної в продажу в найближчому магазині,

достатня для того, щоб її вчасно помітити.

7. Оголошується закінчення сеансу.

8. Учні, які виконували завдання письмово, повертаються в комп'ютерний клас і протягом декількох хвилин вводять в комп'ютер свої коди і свої результати. Документи претендентів контролюються представником міліції. Якщо результат введено в комп'ютер після закінчення нормативного часу, наприклад, після отримання чергового ключа, завдання вважається невиконаним.

9. Результати тестування відправляються на спеціалізований сайт Міністерства освіти для опрацювання. Вони залишаються невідомими протягом обумовленого інструкцією періоду. Учні мають право пройти сеанс знову в той же день або в інший день періоду тестувань.

10. Після закінчення технологічного періоду обробки, публікуються результати. Претенденти з ними ознайомлюються на стендах. Після закінчення

періоду тестувань на всій території, претенденти отримують сертифікати, виготовлені на бланках форми затвердженої Міністерством освіти.

8.10 Безпечна технологія обробки результатів і публікації

1. Результати претендентів у вигляді зашифрованого файлу з територіального Центру тестування направляються на спеціалізований сайт тестувань для опрацювання. Опрацювання проводиться протягом нормативного терміну.

2. Опрацьовані результати тестувань повертаються до територіального Центру тестування у вигляді зручному для друку і публікації. Результати претендентів, які не побажали публікації, у публікованих списках відсутні.

3. Принцип формування зашифрованого файлу результатів і його обробки відомий тільки двом фахівцям - керівнику та заступнику керівника спеціалізованого сайту. Обидва вони є і творцями сайту.

4. Введення на сайт будь-якої іншої інформації, крім файлів-результатів блокується спеціальними методами. Обробка проводиться автоматично без участі людей. Визначена міра особистої відповідальності творців сайту за фальсифікацію результатів, та створені методи захисту таких сайтів.

5. Вся вхідна та вихідна інформація спеціалізованого сайту тестувань дублюється незалежним Центром зберігання у затвердженому порядку, що виключає можливість фальсифікації. Дубльована інформація може бути використана при роботі апеляційної комісії.

6. Результати тестувань оперативно публікуються на стенді територіального ЦТ.

7. Претенденти мають право повторити спробу тестування кількість разів за умови оплати і наявності можливості. У новому сеансі код, який буде сформований для претендента комп'ютером, буде іншим, і варіант завдань - теж.

8. Таблиці рейтингів учнів, які пройшли тестування, без зволікання паралельно публікуються на спеціалізованому сайті, захищеному від посягань.

Цей захід не тільки підвищує популярність і масовість тренувальних тестувань, але й додатково «перекриває» шлях фальсифікації результатів.

9. Після закінчення тестувань всі учні отримують завірений печаткою сертифікат, що включає його рейтинги за 100 - або 1000-бальною шкалою з усіх дисциплін.

10. Іменний сертифікат претендент може подати до ВНЗ при вступі. Оцінки за п'ятибальною шкалою абітурієнти отримують в конкретному вузі на підставі

- його рейтингу;
- рекомендацій Міністерства освіти;
- наказу керівника установи або органу управління освітою, який встановлює шкалу перерахунку.

8.11 Напрями зниження загальної собівартості та собівартості заходів безпеки

Наведена схема зниження загальної собівартості та собівартості заходів безпеки оптимальна для застосування її в регіонах, які забезпечені технікою, так і тих, які не мають технічної бази в достатній кількості. В даному випадку немає необхідності:

- використання дорогих сканерів;
- залучення великого персоналу для обробки результатів;
- залучення персоналу для вторинного контролю обробки результатів;
- тиражування і розподілу великих кількостей стандартних бланків;
- авіаційної пересилки паперових матеріалів;
- одночасного використання площ і технічних засобів;
- одночасної доставки великої кількості претендентів;
- охорони зберігання тестових матеріалів;
- виключені неефективні втрати часу роботи комп'ютерів.

Розрахунки показали, що можна за такою схемою провести тестування знань випускників усіх шкіл протягом трьох-чотирьох тижнів. Собівартість заходів з врахуванням оплати праці і транспортних витрат, витрат на

інспекторські перевірки та інструктивні заходи склала близько 35 грн./сеанс. Ця сума повністю окупається за рахунок оплати тренувань самими претендентами. Одноразове атестаційне тестування має бути безкоштовним.

8.12 Способи фальсифікації, які відомі в практиці тестування

1. Підміна претендента при реєстрації в ЦТ шляхом підробки документів або змови.

2. Підміна претендента при тестуванні в ЦТ шляхом зміни місця або часу виконання завдань.

3. Підміна бланків з результатами роботи претендента у процесі транспортування.

4. Підміна бланків з результатами роботи претендента в ЦТ шляхом змови.

5. Підміна файлу результатів у територіальному ЦТ шляхом змови.

6. Злом та виправлення файлу результатів протягом часу зберігання в ЦТ шляхом змови. Для злomu потрібно досить багато часу.

7. Підміна файлу результату на сервері ЦТ або в момент пересилання між серверами шляхом хакерської атаки.

8. Шпаргалка. Купівля і таємне використання шпаргалки протягом сеансу тестування. При цьому якщо варіантів тесту достатньо багато, використовується змова.

9. Списування. Відбувається, якщо претенденти сидять в аудиторії так, що бачать результати дій один одного.

10. Обмін файлами. Відбувається, коли один претендент реєструється з паспортними даними і паролем іншого претендента.

РОЗДІЛ ІХ. КІЛЬКІСТЬ ТА ЯКІСТЬ ЗУН

9.1 Іспит як форма виміру ЗУН

Багато термінів і понять, що застосовуються педагогами на практиці, не визначені як такі, що можна виміряти, тобто не конструктивні. Знання, уміння і навички - саме такі поняття.

Оперувати невимірними властивостями, коли мова йде про реалізацію ІТ, не можна. Для того щоб виміряти властивість (або навіть явище в цілому, якщо його властивостей не визначене), тобто надати йому конкретне число, необхідно вміти робити одне з двох:

1. Розкласти властивість (явище) на складові частини, які можна перерахувати (порахувати), кожна з яких може бути окремо описана і порівняна з іншими;

2. Впорядковувати властивість (явище) в ряді йому подібних; ранжувати властивості (явища) за формальним алгоритмом, що дозволить однозначно визначити, чому дана конкретна властивість (явище) повинно існувати між двома подібними властивостями чи явищами (попереду чи позаду, ліворуч або праворуч, внизу або вгорі).

Складно однозначно сказати - яку математичну теорему потрібно знати краще, чи біографію якого письменника поставити на перший щабель. Сьогодні спеціалісти однозначно стверджують одне, завтра - інше.

Люди поки не визначили іншого загально визнаного способу вимірювання (співвідношення) обсягу і якості своїх ЗУН, як іспит. Іспити прийнято ділити на формальні (процедура описана нормативними документами) і неформальні.

Іспит називають неформальним, якщо він відбувається в реальній ситуації, наприклад, коли потрібно терміново знайти ліки або їжу в лісі, усунути аварію, вирішити складну задачу тощо. Такий іспит дозволяє виділити кращого, на цей момент, спеціаліста або знавця в групі. Результат неформального іспиту не гарантує, що в інший час, в іншій подібній ситуації найкращим не виявиться інший претендент (може бути, що в момент іспиту у нього боліла голова, або на

нього впливали дії інших членів групи). У цьому розумінні, неформальний іспит не об'єктивний.

До неформальних можна віднести і такий іспит, коли екзаменатор задає питання «на вмання» і ставить оцінку, орієнтуючись за кількома словами, не дочекавшись повної відповіді.

Публічний захист випускної роботи учням також може вважатися неформальним іспитом, оскільки в заліковій книжці оцінка випускної роботи відокремлена від інших, і процедура захисту передбачає перевірку знань лише по одній вузькій темі.

Формальний іспит проводять у формі тривалої співбесіди або тестування. Під час співбесіди претендентові задають кілька основних запитань (частіше два або три) з різних областей навчальної дисципліни і кілька додаткових. Вислуховує відповіді комісія або, частіше, один викладач. На основі вислуханих відповідей викладач, що опирається на власний досвід, робить свою оцінку кількості знань. Ця формальна оцінка - всього лише одне число, яке варіюється в рамках оціночної шкали. Досвідчені викладачі не зупиняються на одному числі, а заносять у свій блокнотик декілька. Більше нікуди заносити. Оцінка в блокнотику - об'єктивніша формальної, вона має декілька компонентів, тому її можна назвати векторною, а формальну оцінку - скалярною.

9.2 Сепарабельність оцінки.

Шкала оцінок частіше буває 3-х, 4-х, 5-ти і 10-ти бальною. Для більш точного оцінювання кількості знань викладач часто придумує свій особистий алгоритм оцінювання, самостійно звужує або розширює формально встановлену шкалу, врізаючи її знизу або додаючи знаки - плюси і мінуси. Часто така розширена шкала, при близькому розгляді, просто є подвійною (дуальною) - число стає мірою кількості «заморожуваних» знань, а плюси і мінуси - мірою умінь і активності учня, тобто якості.

Оцінка, тобто одне єдине число, часто вважається мірою гарантованою викладачем або екзаменаційною комісією кількістю та якістю ЗУН претендента. Мабуть, це не цілком прийнятно.

Не можна оцінити кількість і якість однією мірою. Швидше, і багато викладачів з цим згодні, оцінка - не міра, а пропуск (мандат) учня в одну з неформальних (формальних) груп:

- Учні, які засвоїли матеріал і вміють вчитися. Таким слід відкрити шлях для подальшого навчання і яких можна залучати до відповідальних заходів (виконання робіт, замовлень тощо);

- Учні, яких можна вважати першими претендентами на продовження навчання (надання робочого місця), якщо воно не зайняте представниками першої групи;

- Учні, які навчаються слабо і володіють невеликим обсягом знань. Але є ще надія, що становище виправиться;

- Учні, не виявляють зацікавленості до навчання, яких краще довше не вчити (або яким краще не надавати робочі місця).

Такий підхід до оцінки гарний у рамках одного навчального закладу. Однак, коли мова йде про переходи учнів з одного навчального закладу в інший, випускних та вступних іспитів, в такому разі не застосуємо, тому що в різних навчальних закладах присутні різні критерії отримання цього «пропуску».

Оцінка, яка подається як міра кількості та якості ЗУН, повинна бути:

- об'єктивною;
- уніфікованою;
- зрозумілою;
- сепарабельною.

Під сепарабельною в різних науках розуміють здатність об'єкта до поділу на складові компоненти без втрати властивостей. Ця здатність дуже важлива, наприклад, для біології, де часто не вдається розділити клітину без знищення її властивостей.

Тут під сепарабельністю ми розуміємо здатність екзаменаційної оцінки описувати різні властивості претендента не тільки в сукупності, але і окремо.

9.3 Що вимірюють у процесі тестування?

В даний час використовують три методи обробки та аналізу результатів тестування:

- класичний;
- IRT (Item Response Theory);
- наочний.

Два перших підходи базуються на статистичній обробці так званого «сирого» бала, тобто бали, набрані випробуванним в результаті тестування. Обробка включає:

- визначення труднощів або "ваги" кожного завдання, тобто числа, яке відображає кількість учасників, що з ним впоралися;
- перерахунок результатів - «сушіння сирих балів».

Наочні методи, як правило, не включають тимчасової роботи і дозволяють учаснику побачити результат відразу після сеансу.

Люди, що застосовують наочні методи, економлять свій і чужий час, і опираються при цьому на власний досвід, який говорить:

- якщо завдання складені і «зважені» досвідченим викладачем, то ніяка статистична обробка не вплине суттєво на результати тестування;
- які б системи перерахунку балів не застосовувалися, все одно вгорі рейтингової таблиці фігурують, злегка міняючись місцями, одні й ті ж учні.

Не можна з ними не погодитися. «Сушіння сирих балів» після завершення тестування з добротного створеним тестом дійсно вимагає великих трудовитрат і рідко приводить до істотних змін у рейтингових таблицях. Приклад тому - російський досвід проведення тестувань «Телетестінг», ЦКТ, ЄДІ. Досвідчені члени журі традиційних олімпіад, яким доводиться мати справу з різними системами перерахунку балів, дотримуються часто такої ж думки.

Математичний апарат класичної теорії простіший, ніж апарат IRT, там можливий підрахунок навіть за допомогою калькулятора, хоча зручніше скористатися спеціальними програмами.

Вважається, що класична теорія має ряд недоліків, головний з яких - більша, ніж у IRT, залежність результатів вимірювання від інструменту вимірювання (тесту).

Якісний аналіз добре представлених результатів наочно демонструє успішність навчального процесу, дає можливість оптимізувати, коригувати зміст та методики навчання. Використовуючи результати кожного випробуваного, можна індивідуалізувати навчання.

Від процедури тестування очікують, що після її завершення всі учасники будуть достатньо надійно розподілені (диференційовані) за певними групами. Кількість груп визначається роздільною здатністю застосовуваної шкали оцінювання - 5, 10, 20, 100, 1000.

Претенденти, що потрапили в одну групу, повинні мати однакові характеристики - знання, уміння, навички. Всім претендентам, що потрапили в одну групу, присвоюється однаковий рейтинг (число) або виставляються однакові оцінки в рамках вибраної шкали.

Давайте розглянемо найбільш ймовірну реальну ситуацію, що виникає при застосуванні тестового контролю ЗУН до групи учнів. Поставимо експеримент над групою учнів, зовсім незнайомих з матеріалом дисципліни, з якої складено тестові завдання. Проведемо з ними тестування і розглянемо результат.

Кількість завдань, включених в тестовий комплекс, зазвичай не перевищує сотні. Серед них є прості та складні, що вимагають від учня кмітливості, пам'яті, знання визначень та уміння розв'язувати задачі.

Якщо творці тестових завдань не перевантажували себе залученням сучасних тестологічних технологій і обмежилися (як завжди) лише складанням завдань закритого типу «з вибором однієї правильної відповіді з чотирьох», відбудеться наступне.

Відповідно до теорії ймовірностей, методом «випадкового тикуну» середньостатистичний претендент правильно виконає близько 25% завдань. Претенденти «розумніші», тобто такі, що можуть логічно мислити і знають слабкі місця творців тестових завдань, знайдуть правильні відповіді в 50% -80% випадках. Учасники, які не знають слабких місць технології і не знають дисципліни, але спробували залучити свій інтелект до пошуку правильних відповідей, покажуть результат близько 30% -40%.

Таким чином, навіть за допомогою непрофесійного тестового контролю можна досягти певного результату, тобто розподілити учнів на три групи - «інтелектуалів-хитрунів», «старанних - шукаючих» і «лінивих - хвостистів», результати яких потрапили в інтервал 0 % -35%.

Правда, розділити учнів на такі групи можна буде тільки приблизно, бо їх результати розподіляться по інтервалу можливих значень, як правило, безперервно. Хоча, може бути, ви й знайдете на кривій розподілу три піки, що відповідають середнім даним груп. Тоді завдання розмежування груп, звичайно ж, спроститься.

Тепер розглянемо випадок, коли:

- всі учні з навчальною дисципліною знайомі;
- знання, вміння та навички в усіх учнів не мають різких відмінностей;
- всі учні вчилися в однакових умовах;
- всі учні мають стимул до отримання високого результату при тестуванні.

Тоді крива розподілу результатів учнів буде гладкою, схожою на відому в теорії ймовірності розподіли типу Гауса, Пуассона або Максвелла - з одною вершиною, що піднімається з нуля на початку координат і монотонно спускається до нуля на межі 100%.

Якщо ж група учнів неоднорідна, тобто сформована, наприклад, з учнів звичайної та спеціалізованої шкіл, ситуація може змінитися і крива розподілу результатів перетвориться на двогорбу. Вершини її будуть орієнтовно відповідати середньому рівню знань учнів тієї та іншої підгруп.

А може статись і так, що вершин у кривій розподілу виявиться так багато, що вони стануть непомітними на фоні однієї, що їх об'єднує, підсумкової вершини.

Виховані на непорушних законах класичної теорії ймовірностей тестологи, зазвичай розглядають тільки криві розподілів результатів з однією вершиною. Більше того, вони вважають, що тільки такі криві і можуть виходити, якщо тестові завдання та процедура тестування правильні. Все інше - помилка вимірів.

Більш того, тестологи кажуть, що тестові завдання і процедура тестування правильні, якщо в результаті тестувань виходить саме одnogорба, плавна, впізнавана математиками крива.

На практиці така ситуація, коли крива виходить плавною і одnogорбою, зустрічається досить рідко, та методику її виправлення не можна завжди визнати однозначно об'єктивними.

Вважається, що гарним методом виправлення ситуації, боротьби з «багатогорбністю» є дописування завданням ваг (відносної складності тощо), які виражаються числами. Результат кожного претендента зважують, тобто підсумовують не бали (плюс бал за кожне правильно виконане завдання), а ваги.

Існують різні підходи до обчислення ваг завдань, націлені на «виправлення» кривій розподілу. Завдання такого виправлення математично складне, і не кожен з розроблених способів виправлення призводить до бажаного результату.

Так і повинно бути, оскільки вихідні положення теорії засновані на припущенні про однорідність груп претендентів, тобто статистично «гладкому» розподілі серед них ЗУН. А це трапляється не завжди. Наприклад, візьміть групу, де чверть претендентів вміє досконало вирішувати завдання лише по одній темі з фізики, три чверті - тільки по двох. Тестування такої групи за допомогою усередненого за темами тесту призведе до дивних результатів, з яких важко зробити монотонну одnogорбу криву.

Обчислені ваги можуть сильно залежати від того, як сформована група претендентів. Це означає, що кожне нове тестування, з новою групою претендентів, має супроводжуватись вирішенням задачі по обчисленню ваг знову, проте результати цих двох тестувань не можна буде однозначно зіставити.

Який зміст мають ваги, чи можна використовувати їх значення для вирішення інших завдань, більш значущих, ніж вирівнювання кривої розподілу під встановлену волюнтаристськи форму?

Повернемося до традиційної задачі об'єктивного оцінювання підготовки учнів і розподілу їх за трьома-чотирма групами відповідно до їх рівнів. Крива розподілу результатів з однією вершиною, якою би вона не була математично і естетично красивою, не дає підстав для об'єктивної відповіді на питання про те, де ж об'єктивно проводити межі груп.

Досвідчений тестолог поділить учнів за допомогою простого алгоритму, давно придуманого і дієвого. Зробіть наступне:

- протестуйте учнів і отримайте криву розподілу результатів;
- надайте всім учням рейтинги і складіть список учнів відповідно до набраних ними балів - від вищого до нижчого;
- визначте, скільки учнів може навчатися в сильнішій групі кращих (припустимо - 20);
- перших 20 учнів у рейтинговому списку визначте в першу групу, другу двадцятку в другу, третю - в третю і т.д.;
- якщо необхідно поставити оцінки, поставте першій групі «п'ятірки», другій - «четвірки», третій - «трійки», іншим - що залишилося.

Задача розподілу, на перший погляд, успішно вирішена.

Правда, тут все ж допущена необ'єктивність: число 20 було включено Вами в розрахунок волюнтаристським шляхом, без надійного обґрунтування. Встановлений Вами кордон «20» став бар'єром просто тому, що стільки стільців стоїть у Вас в навчальному класі. Може бути так, що цим поділом Ви поламаєте

чіюсь долю, чи може включите в групу небажаних учнів, які зруйнують дисципліну і навчальний процес.

Невідчутна, з точки зору теорії ймовірностей, різниця в 1-2% може призвести до того, що хтось отримає «четвірку» і назавжди втратить можливість вступити до престижного вузу, а хтось отримає «п'ятірку» не цілком не обґрунтовано, і в групу «кращих» потраплять люди, на навчання яких буде даремно витрачено багато часу і коштів.

Можна якимось чином обчислити інтервал довіри, оточити ним граничну точку і провести співбесіду з усіма учнями, результати яких туди потрапили. Після того вже провести остаточний розподіл, який стане більш точним за рахунок використання людського фактора - досвіду викладача. І це вже - не автоматична процедура.

Як добитися об'єктивної автоматичної диференціації учнів, тобто об'єктивності в оцінюванні їх навчальних досягнень? Очевидно, одного знання величин їх рейтингів для цього недостатньо. Для цього необхідно залучати методи математичного моделювання і статистики. Наприклад, методи факторного аналізу.

9.4 Як виміряти кількість ЗУН?

Хороший тест складається із безлічі систематизованих завдань, наповнення яких рівномірно «покриває» матеріал дисципліни, та для виконання яких потрібно вибрати приблизно однакові масиви інформації і приблизно однакові вміння. У тих випадках, коли «приблизної однаковості» авторам досягти не вдалося, будуть завдання «зважені», тобто такими, числові коефіцієнти яких будуть характеризуватися відносною наповненістю роботи претендента інформацією та вміннями.

Таким чином, правильне виконання кожного завдання тесту свідчить про наявність у претендента конкретного знання і вміння. Якщо різні завдання дозволяють виявити наявність різних ЗУН, то результат тестування - оцінка - буде арифметичною сумою отриманих балів або просто кількістю виконаних завдань.

Якщо сеанс досить тривалий, щоб середньостатистичний претендент зміг впоратися зі всіма завданнями, а завдання тесту повністю покривали матеріал, то обчислена, як вказано вище, оцінка цілком може служити мірилом кількості знань.

А як же виміряти якість?

9.5 Як виміряти якість ЗУН?

Кількісна оцінка ЗУН, безсумнівно, повинна бути відображена в оцінці якості. Проте вона зачіпає значно більший пласт характеристик претендента.

Якість ЗУН виявляється в конкретній виробничій обстановці. Ці прояви виражаються в:

- швидкості реакції на ситуацію;
- точності віддання вказівок;
- об'єктивності створюваних звітів;
- якості створюваної продукції;
- швидкості виконання робіт;
- здатності швидко перевчатися;
- здатності та бажанні підвищувати свою кваліфікацію;
- широті кругозору серед виробничої та наукової інформації;

тощо. Не можна об'єктивно оцінити якість освіти шляхом прямого тестового опитування. Проте таку оцінку можна зробити шляхом непрямого опитування. У непрямому опитуванні замість самого претендента беруть участь його керівники, учні, споживачі продукції.

Проведенням таких опитувань займаються (повинні займатися) територіальні Центри моніторингу якості освіти.

Якість в освіті важливіша за кількість. Фахівець, що вміє вчитися, в стратегічному плані більш цінний для фірми, ніж той, який просто володіє вузьким комплексом ЗУН і чудово впорався з тестуванням. Енциклопедисти досягають у науці та винахідництві великих успіхів, ніж вузькі спеціалісти.

Оцінка якості не може бути виражена одним числом. Скільки параметрів - стільки й показників. Оцінка якості - вектор. Одним з його компонентів є

оцінка кількості знань. Кількість ЗУН (тобто параметри особистості, які можна виміряти тестуванням) є невід'ємною компонентою оцінки якості.

Оцінка якості освіти за допомогою комп'ютерного тестування - складне завдання, яке не можливо сьогодні вважати вирішеним. Можливі шляхи вирішення цього завдання можуть бути пов'язані з наступним:

- тест на вміння пояснити прийняте рішення;
- тест на вміння пояснити хід розв'язання запропонованої задачі і вибір шляху її вирішення;
- тест на навичок асоціативного мислення, вміння знайти зв'язок між явищами і пояснити їх;
- тест на вміння відокремити стандартне від нестандартного, потрібного від непотрібне (у вирішенні даної проблеми), відоме від невідомого;
- тест на вміння пояснити хід розвитку конкретної навчальної дисципліни в історичному, логічному, філософському аспектах;
- ситуаційне тестування, коли потрібно виявити вміння знаходити нестандартне вирішення за досить малий час, а потім пояснити його і провести порівняння з іншими розв'язками, виявити переваги та недоліки.

Це цілком досяжне завдання, котре можна реалізувати з сьогоднішнім рівнем розвитку тестології.

9.6 Як оцінити якість оцінювання якості?

Оцінка якості навчання, яка поставлена досвідченим і визнаним викладачем, вважається об'єктивною. Хоч, самі учні часто з цим не згодні, оскільки знають слабкості викладача і вміють ними користуватися. Викладач інколи проявляє менше зацікавленості до деяких розділів своєї дисципліни і не ставить запитань по тій чи іншій темі. Або навпаки, викладач може мати свою провідну тему і студент, обмовившись про зацікавлення нею гарантовано підвищує свою оцінку.

Є три способи оцінити об'єктивність оцінки викладача:

- опитування випускників. Необ'єктивні екзаменатори не користуються повагою в студентів, навіть якщо такі викладачі і оцінювали їхні відповіді на високі;

- опитуванням роботодавців, у яких працюють випускники;

- опитуванням викладачів навчальних закладів вищої ланки або при опитуванні викладачів, що працюють на курсах підвищення кваліфікації, куди переходять ці випускники для продовження навчання.

Іншими словами, якість і кількість ЗУН оцінюється викладачем вірно, якщо:

- він користується повагою своїх випускників;

- його рекомендаційний лист є гарантією отримання випускником престижного робочого місця;

- результати його роботи користуються заслуженою повагою серед колег.

Ясно, що така «оцінка об'єктивності оцінки» часто неприйнятна з причин довготривалості, трудомісткості тощо.

У нас відсутня і культура опитувань про якість навчання: психологи придумують анкети, служби працевлаштування їх поширюють, але більшість анкет залишаються незаповненими, а заповнені анкети складаються в мішки і зникають. Опубліковані службами працевлаштування результати опитувань часто виявляються необ'єктивними, заснованими на казкових мріях – «Як би вони могли виглядати?»

Хоча, подекуди, робляться спроби побудови систем безперервного навчання, де рішення про прийом абітурієнтів на чергову сходи́нку навчання або про прийом фахівця на роботу ґрунтуються на авторитетній рекомендації викладача і договорах про спільну освітню діяльність. У масштабі країни це - поки дивина.

9.7 Тестування, як засіб об'єктивізації оцінки якості та кількості.

Якщо іспит - це лише захід з видачі мандатів і ми до цього звикли, то чи потрібно змінювати ситуацію? Чи потрібно вводити нову шкалу оцінювання,

придумувати тонкі та складні методи розрізнення якості ЗУН? Чи потрібні спеціальні комп'ютерні інформаційні технології?

Можна припустити, що відповідь більшості педагогів, які мають справу з великими аудиторіями і для яких проблема оцінювання є нагальною, буде ствердною. Наведемо передумови такої відповіді.

1. Оцінка виставляється викладачами завжди на основі власного досвіду. Передісторія, зовнішній вигляд і поведінка претендента, відомості, отримані раніше від колег та інших людей грають при цьому певну роль.

2. Навіть маючи алгоритм оцінки, важко з його допомогою «розвести» претендентів, результати яких потрапили в область близькості оцінок. Не можна переконливо пояснити людині, чому 72 бали - це «п'ятірка», а 71 бал - це «четвірка». Тут допоможуть лише більш складні оцінюючі засоби підвладні комп'ютеру.

3. Неможливо за короткий проміжок часу гарантовано оцінити знання претендента в межах всього матеріалу дисципліни. Тим більше, якщо мова йде про групу. Питання викладач часто ставить в якомусь випадковому порядку (досвідчений студент може навіть «диригувати» цією «випадковістю»), і оцінює відповіді, орієнтуючись лише на наявність ключових слів і виразів у поясненні студента.

4. Претендентів іноді буває дуже багато, а кількість вакансій, на які вони претендують - мало. Така ситуація постійна при вступі абітурієнтів на престижні спеціальності вузу. Тоді «бійка» за кожний «бал» досить драматична, доказів правоти і неправоти членів екзаменаційних комісій можуть бути найрізноманітніші. Опорою в екзаменаційних розглядах можуть стати багаторівневі та багатовимірні шкали оцінювання, а кращим суддею - комп'ютер.

5. Традиційний іспит - завжди лотерея, міцні знання необов'язкові. Таке відношення до традиційної технології педагогічних вимірювань у переважній більшості претендентів. Тільки ОІТ зможуть допомогти змінити таке ставлення.

Проте, серед педагогів чимало прихильників збереження і традиційних способів оцінювання ЗУН. Традиційний спосіб дійсно часто є оптимальним у змісті малозатратності та об'єктивності. Наприклад, у випадках, коли:

- вища оцінка ЗУН не завжди є для учнів бажаною путівкою в життя; наприклад - в ПТУ, де більшість випускників, незалежно від оцінок, потрапляють після випуску в рівні виробничі умови;

- оцінка виставляється формально за двозначної шкалою «зараховано - незараховано», скажімо, при здачі абітурієнтами математичного факультету університету вступного іспиту з української мови;

- недбале ставлення до даної дисципліни в навчальному закладі, дисципліна вважається другорядною, заняття були непродуктивні або не проводилися взагалі.

Надалі виключимо такі випадки з розгляду.

У тестології створено методи, що дозволяють значно підвищити об'єктивність виміру якості та кількості ЗУН. Це так незвично для традиційної педагогіки, що багато викладачів та керівників цьому не вірять. І правильно роблять. Саме таке ставлення до нововведень відрізняє нашу систему освіти від закордонних систем.

Як правило, керівники установ савлять перед собою наступні питання:

- А раптом після тестування виявиться, що явні двієчники отримали позитивні оцінки?

- А раптом п'ятірок буде замало?

- А раптом п'ятірок буде занадто багато?

- А раптом п'ятірку отримає небажаний «абітурієнт», тобто такий, який не володіє потрібними якостями, адже знання дисципліни виявити тільки на усному іспиті?

- А раптом двійку отримає «бажаний» абітурієнт, який відвідував усі підготовчі курси та гуртки?

- А раптом більше половини учнів отримають двійки, а викладачі їм ставили в процесі навчання позитивні оцінки?

Судячи з досвіду російських дослідників, викладач визнає технологію тестування лише:

- коли він сам протестує свої знання і отримує трійку;

Проте керівник навчального закладу визнає технологію тестування лише за трьох умов:

- коли він дізнається, що викладачі закладу вже тестувалися і отримали «різні» оцінки;
- коли він отримує право керування кількістю підсумкових п'ятірок і двійок;
- коли йому «на цифрах» доведуть, що тестувати дешевше, ніж влаштовувати іспити за повною схемою.

Тестування є однією з найбільш технологічних форм проведення автоматизованого контролю з керованими параметрами якості.

Вітчизняна система освіти отримує нову якість. Сучасні комп'ютерні засоби, корпоративні мережі, телекомунікаційні системи дозволяють створити об'єднаний (на рівні організації, регіону, країни) ресурс для діагностики якості освіти та збирати статистичні дані результатів тестування дистанційними методами. Це потужний інструмент управління якістю, можливості якого поки що не досліджені.

9.8 Роздільна здатність і об'єктивність вимірювальних процедур

Під «роздільною здатністю» технології оцінки розуміють ступінь можливості адекватного розподілу (диференціювання, розбиття) всіх претендентів на критеріально-однорідні групи відповідно до рівня їх знань, особистісним якостям, практичною підготовкою тощо.

Зазвичай мається на увазі, що межі між групами учнів, проведені процедурою оцінки з високою «роздільною здатністю», легко пояснити, й вони у нікого не викликають суперечок. Однак багато хто з нас стикалися з випадками логічно нез'ясовного проведення кордонів між «п'ятіркою» і «четвіркою» або «четвіркою» і «трійкою» не лише на усних іспитах, але і при комп'ютерних тестуваннях.

Найчастіше, коли технології тестування розробляються неспеціалістами, межі визначаються волонтаристським шляхом, «зі стелі» (тут дуже проситься слово «завжди»).

Іноді роботодавці пред'являють специфічні претензії до вузівської оцінки якості, вважаючи, що прийнятий на роботу фахівець повинен, в першу чергу, грамотно проектувати вироби, а не знати напам'ять визначення та математичні теореми. Вимоги вузівських приймальних комісій також нерідко розходяться з тим, що потрібно від випускників у школах. Цю різницю в точках зору на якість вихідної та вхідної «продукції» цілком можна пояснити і вона повинна враховуватися при проведенні тестувань.

Роздільну здатність технології можна як підвищити, так і понизити. Вона може бути підвищена:

- по-перше, шляхом переходу до багатобальної шкали (наприклад, 100-бальні або 1000-бальні);
- по-друге, шляхом використання багатовимірних оцінок (наприклад, коли оцінювання кожної складової освіченості фахівця ведеться окремо і обчислення інтегральної оцінки здійснюється із застосуванням методик факторного аналізу).

Роздільна здатність залежить від властивостей не тільки кожного із застосовуваних тестових завдань, а й усієї сукупності ТЗ в цілому.

Кращим слід визнати той масив ТЗ, який:

- по-перше, «рівномірно покриває» весь навчальний матеріал (тобто в завданнях знайшла рівне відображення навчальна інформація подібних по інформаційній вазі та важливості розділів дисципліни);
- по-друге, питання і відповіді на самі завдання не викликають нарікань з точки зору грамотності, взаємної відповідності тощо;
- по-третє, ваги оцінок, які учні отримують у свою «скарбничку» в процесі тестування, об'єктивно відповідають співвідношенням складності та важливості ТЗ;

- по-четверте, завдання не можна виконати шляхом простого логічного висновку або простого впізнавання виходячи з яких-небудь прикмет.

Тестові завдання описують багатьма характеристиками, наприклад; величиною приписаних їм емпіричних параметрів:

- вагою;
- складністю;
- труднощами;
- трудомісткістю;

змістом:

- інформаційно-відображаючі;
- проблемно-евристичні;
- розвиваючі.

шкалою (логікою) оцінювання:

- ТЗ з двозначною шкалою оцінювання (правильно - не правильно);
- ТЗ з трьохзначною шкалою оцінювання (правильно - не зовсім правильно - не правильно);
- ТЗ з чотирьохзначною шкалою оцінювання (правильно - не повно - не точно - не правильно).

Більшою роздільною здатністю володіють тести з великим розкидом ваг завдань, з багатозначною шкалою оцінювання, побудовані на основі розвиваючого методу.

Принципового підвищення рівня об'єктивності оцінки знань (саме - знань, а не умінь) можна досягти лише при повному усуненні людського фактору з процедури оцінювання і переході до автоматизованого контролю. Все частіше і частіше педагоги висловлюють позитивне ставлення до автоматизації рутинних опитувань, бо зрозуміло, що тільки комп'ютеру під силу поставити кілька десятків запитань на кшталт «Назвіть дату битви під Ватерлоо» кожному із сотні учнів та оцінити їх відповіді за 15 хвилин.

Вважається, що процедуру вимірювання та оцінювання вмінь учнів, за винятком простих випадків, і тих, що піддаються формалізації, поки

автоматизувати неможливо. Ця сфера діяльності, наразі, підвладна лише людині і педагоги тут мають рацію, відстоюючи свій людський пріоритет перед машиною.

9.9 Систематичні помилки в тестуванні

Систематична помилка називається систематичною тому, що вона присутня в результаті завжди, систематично, при настанні певних умов. Вона не є випадковою, викликаною неконтрольованими змінами зовнішніх параметрів. Це - характеристика тесту, яку можна контролювати і враховувати.

Іноді буває так, що простіше визначити і врахувати величину систематичної помилки, що виникає в конкретних ситуаціях, ніж виправити тест. Адже виправлений тест теж може характеризуватися систематичними помилками, які ще не визначені.

Систематична помилка тесту - це важлива його характеристика, пов'язана, наприклад, з:

- відмінністю валідності тесту, вимірної для різних груп претендентів (помилка нахилу);
- співвідношенням між груповими значеннями вимірних характеристик по тесту і за критерієм (помилка інтерцепта).

Ці поняття використовуються для акцентування того, що закладено в тесті, - на протипагу можливої випадкової помилки. Це значною мірою пов'язано із специфікою вибірки, на якій адаптується тест. Розглянемо ці поняття детальніше.

- Систематична помилка нахилу. Різниця коефіцієнта валідності тесту для двох груп обстежуваних може бути пов'язана з використанням у ролі критерію суб'єктивних оцінок. Систематична помилка нахилу тесту може виникати і при порівнянні висновків про ефективність його застосування на представниках різних етнокультурних, соціальних та професійних груп. Залежить вона і від величини їх вибірки. Необхідно встановлювати відмінності між коефіцієнтами валідності різних вибірок. Якщо такі відмінності істотні (наприклад, у тесті з хімії в україномовних та національних групах), то краще за все перевірити тест

ще раз на двох незалежних вибірках, для того щоб з'ясувати, чи дійсно ці відмінності настільки істотні. Якщо відхилення викликані специфікою даного контингенту обстежуваних, то слід, прорахувавши цю систематичну помилку нахилу, враховувати її як при аналізі отриманих тестових даних, так і при прийнятті педагогами рішень про відсів та розподіл учнів.

- Систематична помилка інтерцепта означає, що тест систематично завищує або занижує значення критерію для частини претендентів. Виявити її можна, коли результати аналізуються по підгрупах обстеженого контингенту, що особливо необхідно на стадії вивірки нового тесту. Можливо, що вона буде виявлена навіть при однаковій валідності тесту для двох груп. Наслідком бувають різні прогностичні значення тесту для цих груп (наприклад, етнічної більшості та меншості). Якщо до цих груп ми будемо застосовувати єдиний нормальний показник, одна з них може опинитися в більш сприятливих, а інша в менш сприятливих умовах. Потрібно провести розрахунок систематичної помилки інтерцепта для тестів, орієнтованих на обстеження груп з суттєвими відмінностями за освітніми та соціоекономічними рівнями.

Причиною систематичної помилки інтерцепта може бути не тільки зміст знань і умінь, закладених в завданнях тесту, але і семантика тесту - ступінь звичності або складності формулювання його завдань. Для відмінності у національних групах можуть бути істотні етнопсихологічні чинники, особливо в тестах, пов'язаних з гуманітарними цінностями і знаннями.

Можуть позначитися демографічні і біографічні відмінності учнів. Так, для випускників міських шкіл, в яких тестування застосовувалося в старших класах, сама форма подання тесту при вступних іспитах до ВНЗ звична. Це робить тестування для них операцією знайомою, рутинною. У той же час для абітурієнтів з віддаленого села, котрі ніколи не стикалися з тестами, вже сама ситуація тестування пов'язана із додатковими стресовими факторами.

Можлива й інша, семантична ситуація: невірно сформульоване завдання у вступному тесті. Не виключено, що в гіршому становищі опиняться якраз кращі випускники престижних шкіл, звиклі до більш суворих і точних формулювань

завдань. Подумавши, що за «простотою» в завданні ховаються пастки, вони можуть почати шукати в ньому «другий зміст», особливо в ситуації відбору при високому конкурсі. Витративши на виконання завдання багато часу і сил, не знайшовши явного підступу, ці добре підготовлені випускники в ситуації, коли завдання побудовані з вибіркоким варіантом відповіді (а так будується більшість стандартизованих тестів), можуть зупинитися не на більш очевидному та правильному варіанті, а на тому, що «хитріший», і помилитися. А трієчники думати не будуть - виберуть більш очевидний варіант і отримають свій бал.

9.10 Рейтинг і оцінка. Їх визначення та призначення

Обчислений за підсумками проведення тестування рейтинг учня це число, за величиною якого намагаються судити про цілий спектр достоїнств і недоліків.

Рейтинг, виражений всього одним числом (будемо називати його, так як математики, скалярним рейтингом) можна назвати об'єктивним, але не всеосяжним показником. Подібно до того, як переможний результат багатоборця не говорить про те, що він вміє краще за інших - плавати чи бігати.

Кардинально змінити ситуацію можна лише перейшовши до рейтингової моделі, де рейтинг складається з декількох показників, і є векторним.

Спробуємо це зробити. Всі завдання тесту розіб'ємо на групи однотипних, з точки зору викладача, завдань. Наприклад, це можуть бути завдання на застосування стандартної формули або запитання, що стосуються обмеженої літературної теми. Такі групи назвемо субтестам.

Весь тест тепер буде складатися з кількох рівноправних субтестів. Кількість завдань субтеста Z_s , що пропонують учневі, має бути настільки великою, щоб закони випадковості накладені на субтест мали мізерний вплив на результат.

Нехай результат виконання завдань субтеста буде однозначно характеризувати лише одну з якостей учня, наприклад, уміння вирішувати

арифметичні завдання на відсотки. Або, наприклад, вміння швидко орієнтуватися в масивах історичних або хімічних даних.

Значення результату виконання субтеста S можна виразити, наприклад, у вигляді відношення кількості правильно виконаних завдань до кількості пред'явлених:

$$r_s = \frac{Z_+}{Z_S} \quad (1)$$

Виконання завдань часто оцінюють по-різному. Наприклад, виконання одного завдання або групи завдань можна оцінювати за такою шкалою:

- Завдання виконано вірно (+1 бал);
- Завдання виконано невірно (-1 бал);
- Відповідь містить помилки (від -1 до +1 балів).

Не обговорюючи переваги різних шкал, приймемо за інтервал можливих значень r_s відрізок від -1 до +1.

Повний результат проходження тесту виразиться рядком чисел або вектором рейтингу:

$$R = \{r_1, r_2, \dots, r_n\}, \quad (2)$$

де n - кількість субтестів в тесті, $1 \leq S \leq n$. Оскільки абсолютні значення r_s , згідно (1), можуть приймати значення від нуля до одиниці, вектор R вказує лише на точки всередині куба або частини куба в n -вимірному просторі результатів субтестів. Назвемо цей простір простором результатів. Точку, яку вказав вектор рейтингу будемо називати результатом, а куб - область знаходження всіх можливих результатів - кубом результатів.

Можна спробувати уявити, як щільно заповнений куб результатами. Нехай субтестів буде лише два. Тоді куб результатів буде двовимірним і перетвориться у внутрішню частину одиничного квадрата.

Залежно від обраного тесту, обраної групи учасників та їх мотивації, різні області квадрата будуть заповнені з різною щільністю. Наприклад, зовсім мало результатів виявиться поблизу «ідеального» результату; дуже багато

результатів може сконцентруватися в області «найбільш ймовірного» результату, який можна отримати методом «випадкового тикуну» тощо.

Тепер, після розгляду цього наочного прикладу, можна по-іншому подивитись і на завдання процедури тестування. Вона полягає в тому, щоб після проведення тестування, точки-результати учнів виявилися згрупованими, розподіленими всередині декількох (небагатьох) малих областей всередині куба результатів. Учні, результати яких потрапили в одну таку область, будуть вважатися близькими за рівнем (якістю) знань і отримують однакові оцінки.

Осі координат в просторі результатів краще уявити перпендикулярними один одному, а відповідні їм одиниці вимірювань - нерівними. Хоча, деякі автори вважають, що більш наочно буде, якщо уявити осі координат похилими, а нерівність одиниць вимірювань виразити через величини синусів і косинусів кутів взаємного нахилу осей координат. Прийняття одного з цих уявлень – справа особистісного вибору.

Результати проходження різних субтестів r_s можуть (і повинні) мати різні відносні ваги p_s , що відображають як відносну важливість даного субтеста серед інших, так і складність поданих там завдань. Ваги субтестів застосовують, наприклад, коли обчислюють скалярний рейтинг шляхом простого підсумовування результатів проходження субтестів за формулою:

$$R = \sum_{s=1}^n p_s r_s \quad (3)$$

Якщо автор тесту не вважає за потрібне розрізняти важливість і складність завдань, він приймає всі ваги рівними одиниці. Найчастіше, автори приписують завданням ваги не об'єктивно, спираючись лише на власний досвід і здоровий глузд.

У просторі результатів можна визначити так звану метрику або формалізовану відстань між результатами різних учасників тестування:

$$M = |R_2 - R_1| \quad (4)$$

На відміну від випадку роботи зі скалярним рейтингом (3), де відстань між результатами учасників є простою різницею рейтингів, у n -вимірному випадку відстані можна визначати по-різному, у відповідності з цілями тестування.

Неважко зрозуміти, що при підстановці в (4) замість одного з векторів нульового вектора, отримаємо величину (довжину) іншого вектора. Зміст цієї довжини може бути різним.

Довжина вектора рейтингу не має такої ж ролі, як величина скалярного рейтингу. У векторному випадку значно важливіше знати не довжину вектора рейтингу учня, а відстань від нього до вектора «ідеального» рейтингу Ідеальний рейтинг - це рейтинг «ідеального» учня, який виконав абсолютно правильно всі завдання тесту.

Чим менша відстань від вектора рейтингу учня до вектора «ідеального» рейтингу, і чим більша відстань від вектора рейтингу учня до області «найбільш ймовірного» результату, тим вище потрібно оцінити знання учня.

Тому для того, щоб розподілити учнів на групи відповідно до виявлених якостей, необхідно, в першу чергу, обчислити відстані між усіма обчисленими векторними рейтингами.

Після обчислення відстаней між рейтингами всіх учасників, можна спробувати згрупувати результати учасників так, щоб відстані між результатами кожної групи були меншими, ніж відстані до результатів інших груп. Це робиться за допомогою методів факторного аналізу. За допомогою факторного аналізу можна підібрати метрику (4) і величини ваги p_s , так, щоб групи результатів були більш виразні, більш розділені між собою. Оскільки такий поділ можна зробити строго математично, без залучення додаткової емпіричної інформації, результати його можна визнати об'єктивними.

Методам факторного аналізу присвячено чимало книг. Довгий час вони не отримували практичного розвитку у зв'язку з низьким рівнем застосовування обчислювальної техніки. Зараз, на початку XXI століття, ситуація змінилася і застосування методів факторного аналізу в тестології перетворюється на буденність.

9.11 Системи кумулятивного індексування

Системи кумулятивного індексування якості та кількості отриманих ЗУН (системи ІКІ - Індивідуального Кумулятивного Індекссування) тісно пов'язані з рейтинговими системами тестування. Вони є дієвим інструментом:

- підвищення рівня якості ЗУН;
- активізації навчального процесу;
- контролю якості.

Як правило, різноманітні системи оцінки ЗУН носять констатуючий характер, не впливають прямо на якість і рідко приносять задоволення учням і педагогам.

Кумулятивні системи здатні позитивно впливати як на кількість, так і на рівень якості освіти. Ідея кумулятивних систем проста:

- кожна дія учня має визначену ціну в балах;
- одержані кожним учнем за свої дії бали підсумовуються (акумуляуються) протягом певного навчального періоду;
- невдачі учнів ніяк не позначаються (при неправильних відповідях, пропусках занять тощо), суми балів не знижуються, залишаються тими ж;
- суми балів залежать від дати здачі контрольних робіт (запізнився зі здачею контрольної роботи - отримай меншу кількість балів, що додаються);
- підсумкові семестрові оцінки виставляються педагогом на підставі порівняння сум балів (кумулянтах) у навчальній групі. Скажімо, перші 10% учнів, що мають найвищі кумулянти, отримують «п'ятірки», потім ті, хто набрав від 70% до 90%, отримують «четвірки» і т.д.;
- навчальні «борги», а також самостійно вивчений додатковий навчальний матеріал також можна «здавати», збільшуючи свої кумулянти.

У 70-80-х роках минулого століття на низці освітніх конференцій обговорювалося питання про використання у вузах СРСР систем ІКІ. Такі системи інтенсивно розвивалися в 50-60-х роках за кордоном. У СРСР серед першопрохідців були викладачі ФЕН НГУ - факультету природничих наук Новосибірського державного університету.

У 1982-1992 рр., за підтримки колег з Ленінграда і Новосибірська аналогічні експерименти проводилися викладачами ДСДУ на хімічному, математичному, фізичному, історичному факультетах, а також у спеціалізованих класах декількох державних та недержавних середніх шкіл Владивостока. При цьому система ІКІ в ДВГУ відрізнялася від своїх попередниць.

Комп'ютерні технології тоді тільки з'явилися в СРСР, були важкодоступними, і про те, щоб застосувати комп'ютер на заняттях, доводилося тільки мріяти.

Системи ІКІ мають ряд позитивних і негативних рис. Співвідношення їх кількості залежне динаміці в часі на користь позитивних.

Перерахуємо негативні риси системи ІКІ, які проявляються на етапі впровадження:

- у групах, де введена система ІКІ на навчальних заняттях з будь-якої дисципліни, учні починають присвячувати їй значну кількість сил і часу, а на інші дисципліни не звертають належну увагу. Це породжує негативні емоції серед колег-викладачів;

- система ІКІ суперечить традиційній п'ятибальній системі (порушується графік звітності тощо), що викликає негативне ставлення до введення кумулятивної системи з боку керівників навчального процесу;

- система ІКІ різко диференціює успіхи учнів, суми їх балів часто відрізняються в десятки разів. Це породжує нервозність батьків учнів, тому що вони звикли до традиційної п'ятибальної шкали, «нівелює» розходження в активності учнів;

- після впровадження системи ІКІ виникає необхідність додаткової роботи з учнями, що не сприйняли систему на початковому етапі та значно відстали (з ними можна працювати і традиційно, проте через деякий час вони все одно захочуть бути «як усі» і будуть сприймають систему як належне) ;

- система ІКІ може викликати «неспортивні» явища серед учнів (зникнення у «передовиків» зошитів, взуття), тому що не всім бажаним мати п'ятірку в чверті вигідна присутність на заняттях лідерів.

Перерахуємо позитивні якості кумулятивних систем:

- в учнів значно зростає інтерес до додаткової літератури і навчання, що істотно випереджає програму;
- велика активність учнів на консультаціях і колоквиумах;
- публікація кумулянтів призводить до виникнення духу змагання, появи прагнення учнів до оволодіння додатковими знаннями;
- зникають поточні «борги», а якщо з'являються, то здаються оперативно;
- на заняттях зникає «негатив», пов'язаний з острахом учнів неправильно розв'язати задачу, не виконати завдання, отримати двійку - учні завжди мають право і можливість поліпшити свої результати.

Вцілому, позитивні якості систем ІКІ з лишком перекривають негативні. Останні ж зникають після періоду адаптації.

Мабуть, найкращим «ареалом» для впровадження кумулятивних систем є середовище дистанційного навчання. Саме тут знання оцінюється комп'ютером і, відповідно, є всі передумови для відпрацювання кількісних критеріїв.

Поточні значення кумулянтів учнів різних навчальних закладів, будучи обчисленими за однакової системи і зібраними в комп'ютерній БД, цілком можуть слугувати основою для проведення робіт з порівняльного моніторингу якості освіти.

9.12 Визначення скалярного рейтингу

Результати тестувань використовуються для обчислення рейтингу учнів. В основному використовують два види рейтингу: поточний і підсумковий.

1. Поточний рейтинг - метод обліку досягнень студентів у процесі виконання заздалегідь визначеної сукупності завдань, з кожної навчальної дисципліни. Поточний рейтинг покликаний забезпечити ритмічну навчальну роботу. Завдання різної складності заздалегідь оцінені заліковими балами. Виконання кожного завдання збільшує загальну суму балів учнів. Завжди

відомо - скільки балів набрано кожним учнем. Використовується два варіанти поточного рейтингу:

- визначається мінімально необхідна сума залікових балів для отримання заліку без здачі іспиту або якої-небудь додаткової співбесіди;
- визначається кількість балів, необхідних для оцінки за п'ятибальною шкалою.

2. Підсумковий рейтинг - метод визначення рангу (місця) учня в групі за результатами навчання по одній або декількох навчальних дисциплінах.

Якби всі учні мали різні тестові бали, то ранжирування не становило б праці: чим вище бал, тим вище й займане місце. Але зазвичай один і той же тестовий бал отримують кілька людей, і тоді виникає питання - як розподілити всіх випробовуваних на всі наявні місця?

Якщо рейтинг обчислюється в малій навчальній групі, то кількість місць дорівнює кількості учнів у групі. Після тестування всі отримують відповідні місця. Тим хто має однакові бали, дають біль-менш однакові ранги, з наступним пропуском стільки місць, скільки людей є з однаковими балами. Це дозволяє рівномірно і справедливо розподілити досліджуваних на всі місця рейтингу. Наприклад, якщо у групі 13 учнів, то результати ранжуються так:

РАНЖУВАННЯ

Бали	Частоти	Сум. F	Рейтин
1	1	13	13
2	1	12	12
3	1	11	11
4	2	10	9-10
5	3	8	6-8
6	2	5	4-5
7	1	3	3
8	1	2	2
9	1	1	1

Іноді проводиться рейтинг у групі з дуже великою кількістю учнів. Тоді немає особливого змісту розставляти їх на всі наявні місця. Якщо кількість учнів більше ста, то краще застосовувати так званий відсотковий рейтинг, де в якості норми береться 100 процентних місць. Суть процентного рейтингу полягає в розподілі всіх учнів саме на ці сто процентних місць. Це зручно при порівнянні масивів даних, отриманих, наприклад, у різних регіонах у процесі атестації випускників.

Технологія визначення відсоткового рейтингу відображена в наведеній нижче таблиці.

Там використані позначення:

- 1 колонка - бали досліджуваних;
- 2 колонка - частоти;
- 3 колонка - кумулятивні частоти (cum.f);
- 4 колонка - усереднені частоти, отримують складанням значення cum.f в рядку зі значенням cum.f нижче цього рядка і діленням усього виразу на 2. Наприклад; $1 + 0 = 1$, $1 / 2 = 0,5$; $1 + 1 = 2$, $2 / 2 = 1$; $3 + 1 = 4$, $4 / 2 = 2$;
- 5 стовбець - процентні частки, отримують діленням усереднених f (4 колонка) на N.

У даному прикладі $N = 120$

- 6 колонка - відсоток успішності піддослідних, в залежності від отриманого тестового балу (відсотковий рейтинг);
- 7 колонка - відсотковий рейтинг, який отримують вирахуванням $100 - R$, з наступним округленням до цілого рейтингу; даний рейтинг присвоюється випробуваним, що мають відповідні тестові бали.

ВИЗНАЧЕННЯ ВІДСОТКОВОГО РЕЙТИНГУ

Бал X	Наср. f	Cum. F	Уср. F	Частки	% рейт.	Місце
19	2	119	118	.983	98.3	2
18	3	117	115,5	.962	96.2	4
17	6	114	111	.925	92.5	8
16	7	108	104.5	.871	87.1	13
15	6	101	98	.817	81.7	18
14	8	95	91	.758	75.8	24

13	10	87	82	.683	68.3	32
12	12	77	71	.592	59,2	41
11	13	65	58,5	.487	48.7	51
10	10	52	47	.392	39.2	61
9	9	42	37.5	.312	31.2	69
8	9	33	28.5	.237	23.7	76
7	8	24	20	.167	16.7	83
6	4	16	14	.117	11.7	88
5	3	12	10.5	.087	8.7	91
4	6	9	6	.050	5.0	95
3	2	3	2	.017	1.7	98
2	0	1	1	.008	.8	99
1	1	1	.5	.004	.4	100

Наповнення таблиць та опис технології в даному параграфі відповідають (в скор. та адапт. вигляді) тексту четвертого розділу книги: *Аванесов В.С «Композиция тестовых заданий». Учебная книга. 3 изд., доп. М.. Центр тестирования, 2002 г. -240 с.*

РОЗДІЛ X. ОІТ ДЛЯ ЗАХИСТУ ЯКОСТІ ТА УПРАВЛІННЯ ЯКІСТЮ ОСВІТИ

Захист якості освіти та управління якістю освіти - поки незвичні поняття для українського педагога. Вони прийшли в освіту з виробничої сфери. Поява їх тут обумовлена масовим впровадженням ОІТ і появою нових ефективних поглядів на управління розвитком суспільних відносин і виробництва.

10.1 ОІТ і керування якістю.

Ефективне управління системою освіти країни неможливе без наявності зворотнього зв'язку - надійної інформації про стан системи, яка включає показники, які характеризують:

- кадрове та матеріально-технічне забезпечення освіти;
- навчальний процес;
- результати навчання.

Отримання такої інформації в повному обсязі неможливо в даний час в Україні з ряду причин, наприклад, відсутність:

- надійної та науково обґрунтованої системи показників, що характеризують стан системи освіти;
- загальнодержавної системи моніторингу освіти.

Управлінню якістю присвячено чимало наукової та методичної літератури. Є стандарти якості установ, наприклад популярний стандарт - ISO-9000. Нажаль, для освітньої галузі багато з описаних там положень не підходять.

Більшість авторів цих книг та теорій розглядають, в основному, якість управління виробництвом, вважаючи якість продукції вторинним поняттям. Дійсно, вони по-своєму праві. Там, де виробляють матеріальний товар, завжди є багато написаних на папері інструкцій, технологій. Немає проблем з виконавцями, хто інструкцій не виконує - просто звільняють, благо, що безробіття ще існує. Проблема тут з керівниками, які вміють підійматися на верхні поверхи керівної піраміди і не вміють керувати. Для них і створюють десятки томів - стандартів якості, де розписані алгоритми управлінської діяльності на всі випадки життя. Приголомшливе полегшення керівного життя!

Трудовий колектив працює строго за книгами, сам себе підбирає, виховує і контролює. Все за стандартом. Чим краще керівництвом виконуються інструкції - тим якісніше виробництво.

В освіті - не так. Особливо - в українській. Наявність докладної обов'язкової інструкції, як вчити, - це «загибель» якості освіти. Занадто непростий товар - освіта. Ніхто не знає точно не тільки, що це таке, але і як його виробляти. Зате є старі добрі педагогічні рецепти - ділова гра, публічність успіху, змагання тощо.

Є багато педагогів, що чудово знають, як забезпечити методичність і дидактичність, і як потім розповісти про це з трибуни, але мало таких, які б дійсно застосовували в навчальному процесі ці рекомендації. Чому мало, зрозуміло - це важко і часто неможливо. Ось саме тут - в справі нового впровадження тисячоліттями перевірених технологій навчання - керівникові і можуть допомогти комп'ютерні технології.

Для підвищення рівня якості освіти недостатньо тільки методично і дидактично вести навчальний процес. Грамотно організоване багатоетапне змагання між освітніми установами у вигляді, наприклад, освітньої комп'ютерної олімпіади може призвести до значного зростання успішності навчання без особливого збільшення навантаження на навчальний процес. У такій олімпіаді можуть помірятися знаннями не лише учні, а й вчителі, які претендують на перехід у нову категорію.

Результати масових комп'ютерних олімпіад можуть служити:

- підставою для атестації фахівців та установ;
- підставою для виставлення підсумкових оцінок учнів;
- невичерпним джерелом даних для моніторингу якості освіти в установі, регіоні, країні.

Саме такі олімпіади - не що інше, як чудовий інструмент управління якістю освіти. Наведемо головні характеристики олімпіад:

- Масовість (всеосяжність);
- Безперервність (багатостадійність);

- Привабливість;
- Дух боротьби (змагання);
- Значимість (результати впливають на оцінки в журналі, розміри премій для учасників навчального процесу, оцінки атестаційних комісій, підвищення категорій і зарплат тощо);

- Об'єктивність;
- Публічність («країна знає своїх героїв»).

10.2 ОІТ і організація управління якістю

Є багато методичних посібників про нові освітні інформаційні технології, наповнених закликами і таблицями, які повинні допомагати підвищувати якість освітньої продукції. Користі від більшості з них не багато. З іншого боку, кожен діючий педагог може назвати рецепти успіху в цій галузі.

Опишемо один з таких рецептів застосування ОІТ, що здається цілком достойним для освоєння та масового застосування. Він відомий, але ОІТ грають тут не основну, а допоміжну, хоча і дуже важливу роль. Мова піде про впровадження в навчальний процес старого доброго змагання.

Відомі комп'ютерні комплекси, за допомогою яких можна організувати масове комп'ютерне тестування.

Краще, якщо інтерфейс комплексу буде задовольняти викладеним вимогам.

Не менше ніж інтерфейс комп'ютерного комплексу, важливий антураж тестування. Наведемо опис антуражу, який вважається найкращим:

1. Турнір є багатостадійним, що ведеться протягом навчального року;
2. У турнірі беруть участь всі учні освітніх установ, що змагаються, без винятку. Результат кожного повинен увійти в скарбничку команди;
3. Результати кожного турнірного бою негайно публікується із зазначенням вкладу кожного учасника;
4. Вклади учасників завжди позитивні. Вклади тільки накопичуються;

5. Результати перемог негайно позначаються в житті: оплачені екскурсії, походи, особисті та командні медалі, публічні привітання та винагороди для педагогів, виставлення оцінок у журнал за підсумками битв.

Все це важко було б організувати, не маючи комп'ютерних класів. З ними все інакше. Маючи в розпорядженні двох інженерів і ресурс шкільного класу з 12 ПК, можна протягом дня провести тестування 300-400 чоловік. Публікація результатів, з урахуванням автоматизації не займе більше ніж півгодини.

Можна запропонувати й інші шляхи використання ОІТ в благородній справі розвитку системи освіти та в управлінні якістю її продукції.

10.3 ОІТ і кадрове забезпечення

Підвищення кваліфікації та перепідготовка кадрів у сфері освіти, а саме в області ОІТ є невід'ємною частиною заходів, спрямованих на підтримку і підвищення якості та кількості ЗУН випускників освітніх установ. Нові шляхи підвищення ефективності перепідготовки та підвищення кваліфікації пов'язані з появою і впровадженням нових комп'ютерних і телекомунікаційних технологій навчання та оцінки якості ЗУН.

Не кожна районна освітня структура укомплектована штатом методистів, спроможних вести цю роботу в повному обсязі. Проведення ж заходів щодо перепідготовки через крайову структуру настільки ж значуща, як і витратна: відривати від навчального процесу на тривалий час великі групи вчителів, забезпечувати їх транспортом, житлом, обладнанням - часто стає недозвальною розкішшю.

Комп'ютерні класи є вже у всіх районних структурах. Організувати на їх базі короткочасні курси перепідготовки - не проблема. За наявності програмного забезпечення впоратися із завданням дистанційної групової перепідготовки в області професійних знань зможе і один підготовлений методист.

Саме таким шляхом пішли, в рамках договору про спільну діяльність, Департамент освіти і науки адміністрації Приморського краю і ДВГУ.

Важливою особливістю є орієнтованість Програми на перепідготовку кадрів в області ОІТ, моніторингових технологій (масових тестувань знань, збору і обробки даних, прогнозу) і технологій менеджменту якості.

10.4 ОІТ і захист якості освіти

Створення національної системи відкритої освіти (СВО) є пріоритетним напрямком діяльності Міністерства освіти. Вирішення цього завдання нерозривно пов'язано з процесом загального реформування системи освіти. Формування СВО спирається на нові, незвичні для України принципи - технологія і структура створюється з ініціативи «знизу».

Освіта вже сприймається нами не тільки як якийсь поняття, але і як товар. В Україні будується ринок освітніх послуг. А якщо це так, то сам товар, виробники і споживачі товару, організатори виробництва та споживання товару повинні бути відповідним чином захищені правовим механізмом.

Чинний Закон «Про освіту» не включає чіткого, доведеного до чисельних характеристик, визначення якості та рівня якості освіти; немає в Законі поняття конкуренції та монополізму в сфері освіти, так само, як нема і багатьох важливих понять, пов'язаних з ринком, відсутній несуперечливий опис механізму захисту якості.

Система освіти, як виробничий механізм, складається з (залежить від) багатьох компонентів і факторів, що впливають на якість та залежних від нього: викладачів і вчителів; виробничого персоналу навчальних закладів і науково-методичних центрів; авторів навчальних посібників та освітніх технологій; абітурієнтів та учнів ; адміністраторів навчальних закладів та управлінь освітою; роботодавців; меценатів, що вкладають кошти в розвиток навчальних закладів; закордонних організаторів, що впроваджують українські освітні технології; експертів і гарантів якості; держава.

Всі вони повинні бути захищені не тільки від впливів ззовні, але і одне від одного. Наприклад, цілком логічні такі заходи захисту: рішення про реформування, отримання посад тощо, не повинні прийматися адміністратором без урахування об'єктивних показників якості; випускник вузу з більш високим

рейтингом якості повинен мати пріоритет при працевлаштуванні; роботодавець повинен мати чіткі державні або інші гарантії якості прийнятого на роботу випускника вузу.

Виробничі принципи вимагають, щоб якість освітніх послуг виражалась кількісно у взаємному порівнянні. Причому ці вирази - наприклад, рейтингові таблиці - повинні бути захищені достатніми гарантіями.

Моніторинг якості освіти не менш важливий, ніж медичний, хімічний, радіаційний, екологічний та соціальний. Це основа для побудови у державному масштабі механізму захисту, гарантування і підвищення якості освіти. Цей механізм повинен спиратися не тільки на матеріальну та методичну базу мережі ЦМКО, не лише на результати її діяльності, але й на деякі нетрадиційні для нашої системи освіти аспекти. Обговоримо деякі з них.

Роботодавець, як правило, вибирає майбутнього співробітника з декількох претендентів. Його вибір буде більш об'єктивний, якщо він стане спиратися на результати діяльності регіонального ЦМКО - таблицю рейтингів навчальних закладів та оцінки незалежного тестування залишкових знань. Так реалізується один з аспектів механізму захисту якості. Якщо при цьому претенденти законодавчо захищені один перед одним і перед роботодавцем, на основі тих же рейтингових таблиць, то реалізується другий аспект захисту. Одночасно з першими двома, реалізується третій аспект - взаємний захист освітніх установ.

Сучасний роботодавець зацікавлений в швидшому залученні майбутніх співробітників, до закінчення ними навчання. Його вибір претендентів буде об'єктивним, якщо він отримає доступ до даних незалежного тестування ЗУН студентів, наприклад, в мережі Інтернет. Студенти, знаючи, що за їх успіхами стежать майбутні роботодавці, безсумніву, більш серйозно поставляться до своєї освіти. Таким чином, підвищується індивідуальна освітня мотивація і реалізується четвертий аспект - підтримка підвищення рівня якості.

Вузи зацікавлені у завчасному формуванні контингенту абітурієнтів.

Дані ЦМКО є основою для:

- побудови міжвузівської політики такого формування;

- проведення профорієнтаційних заходів.

Публікація результатів діяльності ЦМКО - рейтингових таблиць - неминуче приводить до появи конкуренції серед учнів та освітніх установ. Заклади прагнуть залучити кращих викладачів, впровадити кращі освітні програми, підвищити оснащеність тощо, що призводить до загального підйому рівня освіти.

Комп'ютеризований моніторинг дозволяє аналізувати діяльність не тільки вузів в цілому, а й конкретних викладачів і керівників. Ці дані дозволяють об'єктивно оцінити діяльність і стати основою захисту прав:

- викладачів та вчителів як виробників освіти;
- абітурієнтів та учнів, як покупців і носіїв освіти;
- навчальних закладів, як організаторів виробництва освіти;
- роботодавців, як споживачів освіти;
- держави, як організатора ринку освітніх послуг.

Ведення порівняльного моніторингу якості освіти на рівні регіонів та держав не тільки стане основним стимулом конкуренції в цій сфері, але і є абсолютно необхідним кроком на шляху до інтеграції освітніх систем і побудови глобальної системи відкритої освіти.

Слід уважно поставитися до рішення про конкретний вид механізму захисту та гарантування якості. Ми повинні бути впевнені, як це не парадоксально, що якість освіти буде захищено і від самого механізму свого захисту.

Кожне нововведення призводить не тільки до позитивних, але й до негативних наслідків. Зокрема, одним з можливих негативних результатів впровадження нового механізму, може виявитися поява загального адміністративного прагнення до уніфікації освіти. Це може привести, як це трапилося в деяких країнах, де такі механізми давно застосовувалися, не до розвитку освіченості, а до падіння рівня якості освіти. Про глибину зв'язку цих, на перший погляд, незв'язаних явищ - уніфікації та стійкості - свідчить наука.

Властивості стабільності та самовідновлення систем прямо пов'язані зі ступенем їх неоднорідності. Встановлено, що найменш стійкими є уніфіковані, однорідні системи. Цей факт, крім того, дає ключ до розуміння одного з «чудес»: саме в неуніфікованості та традиційній різноманітності освітньої системи закладена її магічна здатність до самостійного відновлення і підтримання рівня якості.

10.5 ОІТ і проблеми фінансування

Виникаюча в Україні система ДО спочатку була поставлена в жорсткі умови. Вузи виявляли фінансову самостійність, яка полягала часто в тому, що дистанційним просто називали традиційне заочне навчання для залучення абітурієнтів.

Є думка, що вітчизняна система освіти сама «породила» своє недофінансування. Тут є частка істини, бо система давно вже страждає відомою вадою - неефективним менеджментом. «Чоботар без чобіт», - як каже приказка. Вузи виробляють тисячі цілком успішних менеджерів і маркетологів, і забувають про них, беручи на відповідні посади в штат не фахівців, а співробітників з науковими званнями і викладачів. Співробітники вузів завжди працюють на два фронти - кандидат наук «не людина», поки не став керівником (менеджером), а господарник – «не людина», поки не став кандидатом наук. Ефективність виробничої та наукової діяльності при цьому, природно, невисока.

Є й інша відома сторона «медалі» - недофінансування. Пояснимо. Вузи «тягнуть» з собою з недавнього минулого традицію підвищення оцінок у студентській заліковці не за знання, а за «користь» - організацію заходів, ремонту тощо. Здійснити організаційний «подвиг» для деяких студентів значно простіше, ніж зрозуміти матеріал складної дисципліни. Організація забирає досить багато сил у студентів, що пішли по такій стежці, на навчання часу не залишається. Знань у них немає, але вуз «закриває на це очі». У підсумку вони гарантовано захищають диплом і стають молодими спеціалістами. Далі вони, обійшовши колег, що старанно вчилися, традиційно швидко роблять кар'єру,

отримують ступені, не будучи фахівцями, і, врешті-решт, вирішальний голос при розподілі державних коштів.

Це добре, що у нас в суспільстві є такі активні самобутні менеджери. Наша біда в іншому - ці самобутні менеджери не володіють достатнім обсягом знань ні в своїй спеціальності, ні в менеджменті. Зі студентства вони несуть в собі неприязнь не тільки до викладачів, кривлячи душею при виставленні оцінок, а й до системи освіти вцілому. Вони побоюються професіоналів і неохоче допускають їх до своїх лав. Вони по-своєму чесні, і в них просто рука не підніметься проголосувати за направлення досить великих коштів на освітні проекти, тобто, по суті, до сфери ОІТ. Покопайтеся в додатках до дипломів депутатів і міністрів, і Ви знайдете серед них безліч трієчників.

Ми впевнені, що система відкритої освіти, яка будується на основі ОІТ, здатна позбавити нашу освітню структуру від цих та інших недоліків.

РОЗДІЛ XI. ЗАСТОСУВАННЯ КОМП'ЮТЕРНИХ ТЕХНОЛОГІЙ МАСОВОГО ТЕСТУВАННЯ

11.1 Історичні передумови масового тестування

Автоматизовані технології масового тестування ЗУН в ряді країн використовуються вже кілька десятиріч. Є відомості, що професійне тестування держ. службовців проводилося в Китаї ще до нашої ери. Згідно історичних джерел, певні технології тестування на профпридатність (наприклад, визначення якостей воїнів) розроблялися і використовувалися в державах стародавнього Середземномор'я.

У Росії, США, Великобританії, Франції, Німеччині, Австралії, Канаді, Японії, Голландії, Австрії, Китаї та інших країнах існують державні та приватні організації, що займаються розробкою тестових комплексів, їх розповсюдженням та організацією масового тестування. Є організації, які займаються регулярним збором інформації тестувань з метою ведення моніторингу якості освіти. Результати такої діяльності дозволяють планувати розвиток систем управління якістю.

Розвинена індустрія тестового контролю знань у США. У кожному штаті США сьогодні функціонують десятки спеціалізованих центрів тестування. Багато таких центрів мають відділення за кордоном, займаються «рекрутуванням» абітурієнтів та дешевої робочої сили. У США є загальнонаціональні тестові системи (SAT, GRE та ін) для:

- оцінки знань учнів;
- атестації освітніх установ;
- відбору перспективних абітурієнтів до університетів;
- оцінювання характеристик людей при прийомі на роботу і на службу в армію;
- моніторингу освітнього середовища;
- глобальних соціологічних досліджень.

Масові тестування, результати яких напряму впливають на рівень державних дотацій або «визначають долю» абітурієнтів університетів, як

правило, проводяться незалежними організаціями. Особистості експертів і авторів тестових завдань при цьому тримають в секреті. Це робиться для підвищення рівня об'єктивності. У результаті проведення таких масових тестувань:

- учні - отримують атестаційні оцінки за періоди навчання;
- абітурієнти - отримують інформацію про те, в які університети вони можуть бути зараховані автоматично з пільгами оплати за навчання, або - наскільки успішно вони можуть брати участь у абітурієнтському конкурсі;
- педагоги - отримують об'єктивну характеристику якості своєї роботи;
- навчальні заклади - одержують офіційне місце (рейтинг) у рейтингових таблицях штату, країни, співдружності (це найважливіша підстава для підвищення чи зниження рівня оплати навчання, залучення спонсорів, отримання пільг у оподаткуванні тощо);
- батьки - одержують інформацію про освітні установи;
- випускники навчальних закладів - отримують можливість претендувати на більш престижні робочі місця при прийомі на роботу (у багатьох штатах США це право випускників захищено на законодавчому рівні);
- державні органи управління - отримують підстави для розподілу дотаційних коштів між освітніми установами на наступний рік (це найважливіший момент управління якістю освіти в США).

Існують договірні відносини між багатьма європейськими та скандинавськими країнами в області проведення міжнародних порівняльних досліджень якості освіти на основі тестових технологій. Такі відносини мають хорошу основу - історію розвитку міжнародних фізико-математичних олімпіад, що зародилися в середині ХХ століття в СРСР і країнах соціалістичного табору.

Тестування ЗУН могли б бути більш популярними не тільки серед абітурієнтів і безробітних, але і серед фахівців і пенсіонерів, якщо б творці їх сценаріїв не нехтували (ставилися як до чогось несерйозного) важливими рисами людської натури - спортивним азартом, прагненням до лідерства тощо. Люди, які прагнуть продемонструвати кращий результат у тестуванні,

обов'язково будуть вивчати підручники, енциклопедії. Велика частина російського населення вважає, що перемога в порівнянні знань важливіша від спортивної. Азартних людей багато, і це означає, що хороший сценарій тестування може привести до помітного зростання освіченості населення.

11.2 Масові тестування знань в Росії

У Росії роботи в області масового комп'ютерного тестування знань стали реально проводитися тільки в останні два-три десятиліття. До безперечних «старших» лідерів у цій галузі слід віднести Новосибірський державний університет НДІ МІОО НГУ (розробка і реалізація технологій освітніх тестувань студентів та школярів) і Московський державний університет (Центр «гуманітарних цілей технології», розробка та реалізація проекту Всеросійської комп'ютерної олімпіади «Телетестінг», масове комп'ютерне тестування на професійну придатність).

Багато років організовує масові всеросійські тестування на основі бланкової, та частково комп'ютеризованої технології Центр тестування (ЦТ) Морф. В останні три роки ЦТ Морф розгорнув масові щорічні тестування і на основі повністю комп'ютерної технології.

Детальний опис цих проектів можна знайти, наприклад, на сайтах:

1. Проект «Єдиний Державний Іспит»

<http://www.ege.ru/>

<http://ege.ru/>

2. Телетестінг: всеросійське комп'ютерне тестування для старшокласників

<http://www.teletesting.ru/>

3. Центр тестування МО РФ, проект «Централізоване тестування»

<http://www.rustest.ru/>

<http://rustest.ru/>

4. Центр "Гуманітарні технології", проект «Телетестінг»

<http://www.ht.ru/>

У Росії з'явилося чимало фахівців - тестологів (далі - в алфавітному порядку): В. С. Аванесов, А. М. Майоров, Л. В. Макарова, В. І. Нардюжев, А. І.

Субетто, А. О. Татур, В. А. Хлебніков, М. Б. Челишкова, А. Г. Шмельов та ін захищено багато дисертацій, присвячених тестовими технологіями. Можна говорити про створених в різних регіонах країни наукових шкіл тестологів.

З 1995 року масові комп'ютерні тестування школярів ведуться в Приморському краї (цнот ДВГУ, Приморський крайовий ЦНІТ, Приморський крайовий ЦМКО). Зусиллями ДВГУ, районних і міських управлінь освітою, на території Приморського краю створені та створюються близько двох десятків територіальних Центрів моніторингу якості освіти (ЦМЯО). Серед них найбільші - Владивостоцький ЦМЯО і Дальногорський ЦМЯО.

До безперечних успіхів створеної ДСДУ структури віднесемо шестиразове всеросійське лідерство за кількістю проведених сеансів у період централізованих тестувань.

У 2002 році ДСДУ першим у Росії зробив і здійснив проект масового круглорічного комп'ютерного тестування знань усіх студентів молодших курсів очних відділень університету (більше 8 тис. чоловік) з 10 загальних дисциплін (роботи ведуться спеціалізованим підрозділом ДВГУ - Відкритим університетом).

11.3 Засоби профілактики негативних ситуацій при комп'ютерному тестуванні

У процесі тестування не всі претенденти прагнуть сумлінно виконувати завдання. Можливі негативні ситуації, пов'язані з широким спектром мотивів: від принципового несприйняття тестування як методу контролю знань, до суто дитячих витівок навіть з боку дорослих.

Це буває наслідком змови групи людей, не зацікавленої з яких-небудь причин проходити тестування і яка хоче його зірвати. Відмова спотворює результати, особливо якщо кількість суб'єктів тестування, що відмовилися перевищує 5%. Декларовані відмови (незалежно від обґрунтованості) знижують у решти претендентів мотивацію роботи.

Особливо негативними є публічні відмови, що позначаються на репрезентативності вибірки тесту, і ініціаторами яких є лідери груп. Спроби

переконань і вмовлянь ні до чого, зазвичай, не приводять. Залякана група при дидактичному тестуванні дає завідомо занижені результати.

Способами профілактики ситуації відмов, як показує практика є:

- продуманий вибір часу тестування;
- визначення розумної тривалості тесту (з урахуванням того, що після 30 - 35хв роботи з тестом продуктивність рішень знижується);
- комфортність приміщення, температури, відсутність відволікаючих факторів (шум);
- попереднє роз'яснення мети і задач тестування, значення його результатів для об'єктивного оцінювання навчальних досягнень учнів;
- зовнішній вигляд і стиль спілкування персоналу повинні не відволікати претендентів, а стимулювати сумлінну роботу, створюючи доброзичливу атмосферу.

Другий вид перекручувань - прагнення претендента, по можливості, ховатися за нейтральними відповідями типу «важко відповісти». Вважається, що якщо в тесті у кого-небудь з обстежених до 30% відповідей такого типу, то його слід виключити з вибірки (при оцінці репрезентативності, надійності, валідності тесту).

Досвід показує, що якщо більше 50% претендентів не можуть виконати завдання тесту - це завдання невдале, має низьку дискримінантну здатність.

Третім видом є навмисні спотворення відповідей з боку тих, хто знає правильну відповідь, але з якихось причин вирішує пожартувати. Серед них зустрічаються любителі «системного» підходу. Вони відзначають як правильний варіант кожен перший (третій, четвертий, по черзі) варіант виконання, або на першу половину тесту відповідають позитивно, а на другу - негативно, тощо.

Перевірку щирості того, хто відповідає, краще проводити методами, які не викликають виникнення відчуття підозри. Психологічні опитувальники включають групи запитань, які контролюють сумлінність претендента - «шкали брехні», а в професійно складені анкети - спеціальні контрольні питання.

У дидактичні тести включають такі питання також. Цей прийом доцільно використовувати не в первинному варіанті тестів, а в тому, де після відбору завдань проведена стандартизація. Якщо при тестуванні діагностом були помічені відволікаючі чинники - він зобов'язаний їх зафіксувати в звіті і спробувати дати прогноз можливого впливу цих факторів на результати. До таких факторів належать: поява в аудиторії постороннього або представника адміністрації, шум біля аудиторії, яскраві зовнішні прояви - хвороба будь-кого з претендентів (нудота, голосний кашель).

Для того щоб попередити спроби тестованих знайти відповідь за допомогою підказки або інших тестованих, необхідно заздалегідь в інструкції з проведення тестування чітко обумовити порядок дій педагога-діагноста:

- зауваження з фіксацією порушника і завдання, за яким запитувалася допомога;
- виключення для порушників цього завдання з числа оцінюваних;
- повторне тестування порушника в індивідуальному порядку;
- припинення тестування всієї групи тощо.

11.4 Дидактичний принцип індуктивної мотивації в організації масових комп'ютерних тестувань

Мотивація учасників - найважливіша умова не тільки масовості, а й дидактичної доцільності масових тестувань. Неправильно мотивовані учасники зберігають у своїй пам'яті мимоволі виникненні у них під час проходження тестування відчуття втоми і «примусу». Це може негативно позначитися на подальшій освітній траєкторії особистості. Позитивна ж мотивація може безпосередньо «підштовхнути» зростання якості освіти в територіальних школах.

У кінцевому підсумку провідниками заходів мотивації є керівники та персонал територіальних центрів тестування (моніторингу якості освіти).

Можна створити і розвісити плакати, можна провести роз'яснювальні передачі для школярів і їх батьків по телебаченню. Але весь створений при

цьому настрої учасників тестувань та їх батьків випарується, коли хто-небудь з персоналу в комп'ютерному класі висловить сумнів:

- в якості тесту;
- в доцільності участі в тестуваннях;
- в об'єктивності підведення підсумків тестуванні;
- у щирості заяв керівників системи освіти, щодо долі кращих учасників;
- в чесності керівництва територіального центру та відсутності фальсифікації;
- в істинності цільового витрачання коштів (грошей), отриманих від суб'єктів тестування за участь у тестуваннях, тощо.

Людей, які не здатні до самоконтролю, не здатних тримати при собі свої почуття і думки, не повинно бути в штаті центру.

Позитивна мотивація учасників, створювана централізованими заходами, підкріплена загальним позитивним настроєм співробітників центру (опосередковано переданою мотивацією), може творити дива.

Організовуючи тестування в масштабах регіону, слід подбати про мотивацію співробітників територіальних центрів. Свою, створену зовнішніми організаторами, мотивацію вони передають (індукують) учасникам тестувань. У ДСДУ це робиться наступним чином:

- створена прозора система виплат винагород співробітникам центрів, ставлячи суми в пряму залежність від кількості проведених сеансів;
- щодня співробітники центрів отримують неформальні листи, де висловлюється виключно позитивне ставлення організаторів до щоденних успіхів, називаються імена кращих, проводяться порівняння та обговорення;
- організовано змагання між центрами за отримання призів за підсумками заходів.

Дієвість цих заходів очевидна: ДСДУ шість разів став російським лідером за масовістю участі школярів у централізованих тестуваннях.

11.5 Моніторинг якості освіти в Росії. Міжнародне співробітництво

Якість освіти школярів, роботи вчителів і шкіл Росії в недалекому майбутньому буде контролюватися незалежними експертними центрами, координуючими набір абітурієнтів до вузів, що будують таблиці рейтингів освітніх установ, що проводять моніторинг освітніх систем. Нові для Росії ринкові терміни - менеджмент, маркетинг - міцно вкорінятимуться в лексиконі працівників освіти.

Роботи з впровадження моніторингових технологій в даний час координуються Міністерством освіти РФ (<http://www.informika.ru/>) та Інститутом загальної середньої освіти Російської Академії освіти (ИОСО РАН) (<http://www.ioso.ru/>)

Росія здійснює ряд міжнародних проектів у галузі тестування. Інформацію про це можна знайти на офіційному сайті www.informika.ru. Результати цих заходів використовуються для побудови міжнародних рейтингів освітніх систем. Щорічно публікуються рейтингові таблиці, де можна почерпнути відомості про те, школярі якої країни краще знають, наприклад, математику чи хімію. Об'єктивність оцінок забезпечується спеціальними стохастичними процедурами, на підставі яких вибираються конкретні школи і класи на території країн, де проводяться тестування.

В останні роки відзначається підвищений інтерес до проблем якості освіти в усьому світі. Країни об'єднують зусилля в розробці підходів до оцінки та управління якістю освіти. Створюється система моніторингу якості освіти у світі. У ній беруть участь більше 50 країн. Її організатори:

- Міжнародна асоціація з оцінки навчальних досягнень IEA (International Association for the Evaluation of Educational Achievement);
- Організація економічного співробітництва і розвитку - OECD (Organisation for Economic Co-operation and Development).

Прикладами міжнародного співробітництва в цій області є міжнародні порівняльні дослідження, серед яких слід виділити: IAEP-II-1991, TIMSS- 1995, TIMSS-R 1999, CIVIC -1999, CiViC-2000, PISA-2000.

Повний цикл подібного дослідження в даний час, у зв'язку з швидким розвитком теорії і практики педагогічних вимірювань, а також технології проведення тестування та обробки результатів, становить 3-4 роки. Це дозволяє організувати ефективний моніторинг освітніх систем за ключовими показниками якості та ефективності освіти. Одного зрізового дослідження, проведеного один раз на десять років, недостатньо для виявлення тенденцій розвитку або наслідків реформ.

Росія брала участь у міжнародних дослідженнях навчальних досягнень в області:

- математики;
- природничих предметів;
- громадянства;
- читання.

(IAEP-II, TIMSS-1995, TIMSS-1999, SITES, CIVIC-1999, CIVIC-2000, PISA-2000, PIRLS-2000).

Найбільш важливим міжнародним дослідженням з оцінки якості освіти у світі в 90-х роках, в якому брала участь Росія, було III Міжнародне дослідження з оцінки якості математичної та природничої освіти - TIMSS. Мета дослідження - порівняльна оцінка природничо-математичної підготовки учнів середньої школи в країнах з різними системами освіти та виявлення факторів, що впливають на рівень цієї підготовки.

У дослідженні TIMSS беруть участь багато країн: Австралія, Австрія, Англія, Аргентина, Бельгія (мови навчання: французька і фламандська), Болгарія, Угорщина, Німеччина, Греція, Гонконг, Данія, Індонезія, Іран, Ірландія, Ісландія, Ізраїль, Італія, Іспанія, Канада, Кіпр, Колумбія, Корея, Кувейт, Латвія, Литва, Мексика, Нідерланди, Нова Зеландія, Норвегія, Португалія, Росія, Румунія, Сінгапур, Словацька Республіка, Словенія, США, Таїланд, Філіппіни, Франція, Чеська Республіка, Швеція, Швейцарія, Шотландія, ПАР, Японія.

Завдання дослідження TIMSS:

- Збір, аналіз та узагальнення інформації про системи освіти.

- Аналіз та порівняння систем природничо-математичної освіти на рівні визначення цілей, планування змісту освіти та вимог до підготовки учнів з конкретних предметів.

- Аналіз навчального процесу навчання математики і природничих предметів.

Результати проведених в Росії міжнародних досліджень з оцінювання якості освіти публікуються у збірниках матеріалів. Їх готує Центр оцінки якості освіти (ЦОЯО). Наприклад:

- ВИПУСК 1. Ковальова Г. С. Основні підходи до порівняльної оцінки якості математичної та природничої освіти в країнах світу (за матеріалами міжнародного дослідження - TIMSS)

- ВИПУСК 2. Порівняльний аналіз математичної та природничої підготовки учнів основної школи Росії. Деніщева Л. О., Дюкова С. Є., Ковальова Г. С., Корощенко А. С., Краснянська К. А., Мягкова А. Н., Найдьонова Н. М., Резнікова В. З., Суравегина І. Е. / Під ред. Г. С. Ковальнової.

- ВИПУСК 3. Що відповідають учні, вчителі та директори шкіл Росії на питання міжнародних анкет? (Альманах міжнародного дослідження TIMSS / Частина I /) Укладачі: Гуссак Н. А., Ковальова Г. С.

- ВИПУСК 4. Порівняльна оцінка природно-математичної підготовки випускників середніх шкіл Росії. Деніщева Л. О., Ковальова Г. С., Краснянська К. А., Лошаков А. А., Найдьонова Н. М., Нурмінській І. І., Тарасов Ю. П. / Під ред. Ковальнової Г. С.

- ВИПУСК 5. Що відповідають випускники та директори середніх шкіл Росії на питання міжнародних анкет? (Альманах міжнародного дослідження TIMSS / Частина II /) Укладачі: Гуссак Н. А., Ковальова Г. С.

- ВИПУСК 6. Порівняльний аналіз програм і підручників з математики та природничих предметів. Івченкова Г. Г., Краснянська К. А., Ковальова Г. С. та інші / Під ред. Ковальнової Г. С.

- ВИПУСК 7. Порівняльна оцінка математичної та природничої підготовки учнів початкової школи Росії. Івченкова Г. Г., Ковальова Г. С., Краснянська К. А. та інші / Під ред. Ковальової Г. С.

- ВИПУСК 8. Моделі взаємодії факторів, що впливають на рівень математичної та природничої підготовки школярів Росії. Ковальова Г. С., Краснянська К. А., Найдьонова Н. Н. і др. / Під ред. Найдьонова Н. М.

- ВИПУСК 9. Технічний звіт з проведення в Росії міжнародного дослідження - TIMSS. Каменщикова Є. К., Ковальова Г. С., Крайнова Н. А., Найдьонова Н. М. / Під ред. Ковальової Г. С.

Проект TIMSS є одним з найбільш великомасштабних освітніх досліджень за охопленням досліджуваних проблем і за кількістю країн-учасниць.

TIMSS об'єднав зусилля фахівців більше 50 країн світу. У процесі дослідження вивчаються тисячі програм і підручників різних країн з математики та природничих предметів, проводиться тестування та анкетування сотень тисяч школярів, тисяч вчителів і директорів шкіл.

Для проведення дослідження TIMSS в Росії залучається велика кількість спеціалістів: представники РАО та МО РФ, працівники освіти десятків регіонів Росії, директори та вчителі сотень шкіл, десятки тисяч учнів. Формування досліджуваної вибірки здійснюється відповідно до міжнародних вимог, аналізується і затверджується міжнародними експертами.

Участь Росії в TIMSS дозволило також отримати інформацію про систему освіти Росії, якої вона раніше не мала.

Особливістю TIMSS є його значна орієнтація не на ранжування країн за рівнем навчальних досягнень учнів, а на пояснення отриманих даних.

11.6 Організація та проведення дослідження TIMSS-R (1999) у Росії

У TIMSS-R (1999) брало участь 38 країн: Австралія, Англія, Бельгія, Болгарія, Гонконг, Йорданія, Іран, Ізраїль, Індонезія, Італія, Канада, Кіпр, Південна Корея, Латвія, Литва, Македонія, Малайзія, Молдова, Марокко, Нідерланди, Нова Зеландія, Росія, Румунія, Сінгапур, Словацька Республіка, Словенія, США, Таїланд, Тайбей (китайський), Туніс, Туреччина, Угорщина,

Філіппіни, Фінляндія, Чеська Республіка, Чилі, ПАР, Японія. З них 26 країн брали участь у дослідженнях 1995 р. і 1999 р., 12 країн приєдналися до дослідження на другому етапі.

Відмінною особливістю дослідження TIMSS є наукове співробітництво спеціалістів всіх країн, що беруть участь у дослідженні. Це дозволяє включити в роботу провідні колективи світу і таким чином підтримувати високу якість дослідження. Так, у проведенні дослідження та розробці його інструментарію приймали участь багато науково-дослідних центрів або професійних організацій світу: Служба педагогічного тестування (ETS, США), Канадський Центр Статистики (Statistics Canada, Канада), Секретаріат ІЕА (Нідерланди), Центр обробки даних (DPC, Німеччина) та інші. Для координації зусиль фахівців різних країн були створені дорадчі комітети, які склалися з провідних фахівців світу. Координація всього дослідження здійснювалася Міжнародним координаційним центром, який розташовується в Бостонському коледжі в США (Бостон). Експериментальна перевірка інструментарію TIMSS здійснювалася в усіх країнах.

Дослідження TIMSS в Росії проводилося фахівцями ЦОКО ІОСО РАВ при участі Департаменту загальної середньої освіти Міністерства освіти РФ. У завдання російських фахівців входили організація і проведення дослідження в країні, участь в розробці та адаптація інструментарію до умов Росії, формування представницької вибірки учнів Росії, підготовка матеріалів дослідження (їх переклад, адаптація, підготовка до друку, тиражування), організація і проведення тестування та анкетування по школах країни, забезпечення контролю якості проведення дослідження в Росії, введення та обробка даних, аналіз результатів дослідження, підготовка наукових і технічних звітів, створення банку даних дослідження. Організація та проведення дослідження TIMSS в регіонах Росії здійснювалися за участю Морф через міністерства освіти, управління чи департаменти освіти відібраних регіонів Росії, педагогічні інститути, інститути підвищення кваліфікації робітників

освіти та інші установи. Близько 1500 фахівців взяли участь у проведенні дослідження TIMSS в Росії в 1999 р.

Вибірка учнів 8 класу початкової школи Росії будувалася з 89 регіонів, об'єднаних у 10 економіко-географічних зон на основі ймовірнісно-стратифікованого підходу. При виборі регіонів Росії розрахунок проводився на основі федеральної статистики Міністерства освіти. На федеральному рівні відсутня відповідна міжнародним вимогам статистика, необхідна для формування представницької вибірки учнів. Тому процедура вибірки учнів в Росії містить додаткову стадію - вибір регіонів. Це дозволяє власними силами збирати необхідну регіональну статистику відповідно до міжнародних вимог.

Всього в дослідженні було відібрано 47 регіонів. Відбір шкіл проводився для кожного обраного регіону окремо. Всього в дослідженні брало участь 189 шкіл. З них 75% міських, 25% селищних і сільських шкіл. У кожній школі вибирався один клас, всі учні якого брали участь у тестуванні. Всього в дослідженні брало участь 4332 учнів 8 класу. Вчителі, що ведуть у відібраних класах математику і всі природничі предмети, брали участь в анкетному опитуванні. Всього було охоплено 189 вчителів математики і 756 вчителів природничих предметів.

При створенні інструментарію дослідження TIMSS-R (1999) використовувалися ті ж підходи, що і в дослідженні TIMSS (1995). Повний інструментарій включав тести досягнень (8 варіантів по 75-80 завдань закритого та відкритого типу з математики і природничих предметів, виконання одного варіанту було розраховане на 90 хв.); анкети для учнів, вчителів математики та природничих предметів, адміністрації школи, спостерігачів за проведенням тестування в школі, експертів-предметників з математики і природничих предметів; керівництва по формуванню вибірки, проведення апробації інструментарію та основного дослідження, перевірі завдань з відкритими відповідями, вводу та обробки результатів, забезпечення контролю якості проведення тестування в школах.

При розробці тестів додатково враховувалося, що на основі їх виконання будуть порівнюватися результати двох зрізових досліджень 1995 р. і 1999 р. У зв'язку з цим у них частково використовувалися завдання 1995 р., які не були опубліковані у відкритій пресі. Опубліковані завдання були замінені на інші з тих самих розділів. Особливу увагу приділяли відповідності статистичних параметрів. У всіх анкетах за ключовими показниками використовувалися ті ж питання.

При інтерпретації результатів міжнародних досліджень математичної підготовки учнів 8 класу і випускників середньої школи, в яких брала участь Росія в 1995 р. і 1999 р., слід мати на увазі притаманні їм особливості. У подібних дослідженнях для перевірки, як правило, виділяються питання змісту навчання, які є загальними для більшості країн-учасниць. Тим не менше, і при такому підході зміст перевірочних завдань у тій чи іншій мірі не відповідає змісту програми навчання математики відповідної паралелі класів у кожній з країн.

Для Росії невідповідність завдань, що перевіряють зміст програми 8 класу значно більша, ніж для багатьох країн. Це пов'язано з тим, що вона сильно відрізняється від програм інших країн за номенклатурою питань і вимогами до підготовки учнів. По відношенню до особливостей змісту курсів математики в Росії завдання розподілилися наступним чином. Більше половини (60%) завдань складені на матеріалі, який традиційно вивчається в курсі математики 5-6 класів, а в більшості країн продовжує вивчатися в 7-8 класах. Тільки 23% завдань відповідали вимогам курсу, що вивчається в 7-8 класах. При цьому кількість основних тем курсу алгебри та геометрії у цих класах досить незначна або ці теми зовсім не знайшли відображення у змісті перевірки. Тому наші учні не отримали можливості продемонструвати рівень засвоєння значного за обсягом матеріалу, який був ними вивчений до моменту проведення тестування.

Форма перевірки міжнародних досліджень виявилася незвичною для росіян. За 90 хвилин учень повинен був відповісти на 75-80 тестових завдань

різного типу, складених на основі матеріалу 6-ти навчальних предметів природничо-математичного циклу. У цих умовах явно проявилися як позитивні якості, так і недоліки математичної освіти.

Аналіз результатів досліджень у 1995 р. і 1999 р. показав, що стан математичної підготовки восьмикласників Росії не зазнало істотних змін в порівнянні з власними результатами 1995 р., а також у порівнянні з іншими країнами.

По відношенню до результату Росії (середній бал по тесту) країн-учасниць можна розподілити на три групи: результати істотно вищі за російські, результати статистично не відрізняються від російських, результати істотно нижчі від російських. Склад країн-учасниць дослідження в 1999р. змінився в порівнянні з 1995р., змінився і склад цих груп.

Цікаво відзначити, що в 1995 і в 1999р.р. перша група включала 7 країн. Це країни Азії (Сінгапур, Корея, Тайбей, Гонконг, Японія) і дві європейські країни - Чеська Республіка (тільки 1995р.) і Фламандська частина Бельгії. Вони показали більш високі результати в порівнянні не тільки з Росією, але й з усіма країнами-учасницями.

У 1995р. серед країн, результати яких не відрізняються від російських, входять країни, що мають гарний рівень математичної освіти: Франція, Угорщина, Нідерланди, Австралія, Канада. Результати цих же країн (крім Франції, яка не брала участі у дослідженні) не відрізняються від російських і в 1999р.

У 1995р. серед країн з результатами нижче від російських - США, Англія, Німеччина, Італія. Результати цих же країн (окрім Німеччини, не брала участі у дослідженні) залишились нижчими від російських і в 1999р.

11.7 Єдиний державний іспит у Росії

Серед нових інформаційних технологій, впроваджуваних Морф в системі освіти, виділяється технологія Єдиного державного іспиту (ЄДІ). Масштабне експериментальне впровадження ЄДІ в регіонах стало можливо тільки завдяки всебічній підтримці територіальних органів влади.

Спочатку в ЄДІ застосовувалася тільки бланкова технологія, але, останнім часом позначився «крен» у бік впровадження комп'ютерного тестування.

Основа ЄДІ - об'єднання шкільного і вузівського іспиту - поки не знаходить загальної підтримки в багатьох великих державних вузах Росії. Причин цьому багато, а головна - це недовіра керівництва вузів до положень технології ЄДІ запропонованої Морф:

- фактично відповідальність за якість контингенту прийнятих за експериментом абітурієнтів перекладається з вузівських приймальних комісій на плечі невідомих вузам контролерів;
- немає ясності з технологією планування набору студентів;
- немає ясності з ДФЗ - державними іменними фінансовими зобов'язаннями - які держава передбачає закріплювати за випускниками шкіл.

Особливості технології ЄДІ дозволяють автоматизувати виявлення недоліків і недобросовісності в пунктах проведення ЗНО. Статистичний аналіз таких явищ показав, що більш за все недоліки виявляються в «переселенських» територіях, де населення проживає тимчасово або прибуло раніше з інших регіонів і країн.

До «делікатних» моментів технології слід віднести, наприклад, невідпрацьованість форми бланків і пов'язану з цим багаторазову перевірку сканованих даних. Наприклад:

- На бланку реєстрації в назвах полів використовуються різні друкарські шрифти, які відволікають увагу школярів від зразків і призводять до багатоваріантного написання цифр і букв.
- У переліку зразків відсутні написання латинських букв, вони можуть бути потрібними.
- Написання інструкцій «червоним по червоному» і «рожевим по рожевому» приводить до складнощів для дітей зі зниженим рівнем зору.
- Написи - пояснення до полів незрозумілі школярам, викликають суперечки і вимагають додаткових усних пояснень.

- Страх школярів написати неправильну літеру або невірно намалювати хрестик, підтримувана педагогами, відіграє негативну роль і може негативно позначатися на підсумковій оцінці знань.

- Низька швидкість обробки результатів і доставки сертифікатів до місць проживання.

Багатьох із зазначених тут, а також інших «делікатних» моментів використовуваної технології можна було б уникнути, якщо видозмінити технологічний план, більш широко включити в нього використання можливостей комп'ютерної техніки.

Перешкодами до впровадження комп'ютерних технологій ЄДІ можуть стати, наприклад:

- слабе володіння учнями комп'ютерною клавіатурою;
- мала забезпеченість комп'ютерною технікою та відсутність відповідних центрів тестувань;
- наявність тестових завдань виконуваних у вільній формі, контроль яких поки неможливо автоматизувати за допомогою комп'ютера.

Перші дві з цих перешкод тимчасові і, швидше за все, будуть усунуті на більшості територій протягом найближчого року - двох. Третє - більш складне, і пов'язане воно з передбачуваною необ'єктивністю оцінки з використанням тестової системи. Запропонована організаторами ЄДІ система оцінки «вільних» завдань не позбавлена недоліків, оскільки перевіряючі і перепереверючі виконані завдання фахівці неминуче привносять в оцінку «людський» чинник.

При розробці технологій ЄДІ і ЦКТ, фахівцями міністерства несправедливо забуті позитивні моменти, відпрацьовані і впроваджені Центром "Гуманітарні технології" під керівництвом професора А. Г. Шмельова при проведенні всеросійських заходів «Телетестінг». Це, зокрема, - змагальність. Навряд чи є особлива потреба в забороні повторної здачі ЄДІ, тим більше, при переході на комп'ютерну технологію, коли варіанти комплектів завдань не повторюються. Якщо технічні середовища дозволяють зробити можливим багаторазове складання тесту, система освіти територій і регіону тільки виграє,

коли школярі у вільному режимі, в порядку живої черги стануть повторно (багаторазово) проходити тренувальне тестування з підсумковим заліком кращого результату. Показники якості знань від цього покращаться.

ЄДІ не може залишатися одноразовим заходом, що проходить лише раз на рік. Така дорога налагоджена машина не може стояти і чекати свого часу протягом року. Технологія ЄДІ повинна оселитися в кожній школі на постійній основі, перерости з інструмента для вимірювань в інструмент для підтримки, підвищення, гарантування і захисту якості освіти.

Моніторингові дослідження дозволяють діагностувати освітнє середовище і прогнозувати його розвиток. Такі роботи, що мають широке поширення у ряді країн, не є повсякденністю в Росії. Благим побажанням залишається застосування результатів моніторингу для впорядкування формування контингенту абітурієнтів. І вузам, і абітурієнтам було б непогано визначитися заздалегідь з показниками конкурсу і величинами прохідних балів на спеціальності. Поки отримані на ЄДІ бали не є гарантією вступу абітурієнта на обрану спеціальність в обраному вузі. Це - одна з проблем, що утримують ВНЗ від участі в експерименті ЄДІ.

Впровадження ЄДІ вимагає багатьох попередніх заходів:

- навчання фахівців;
- створення та обладнання спеціалізованих центрів;
- відпрацювання шляхів фінансування;
- роз'яснювальна робота з населенням.

11.8 Моніторинг, атестація, сертифікація. Досвід США

Вищою можливою мірою об'єктивності оцінки якості освіти та навчання претендента, який закінчив курс навчання у навчальному закладі, можна досягти, якщо сумувати (обчислювати) її, виходячи з трьох компонентів:

- результат тестування ЗУН претендента;
- результат опитування споживачів випускників про якість випускників конкретного підрозділу навчального закладу;
- рекомендації викладачів випускника.

У багатьох країнах розвинена культура рекомендаційних листів, а згадані опитування проводяться незалежними моніторинговими центрами. Діяльність територіальних моніторингових центрів зазвичай фінансується або державою, або великими освітніми асоціаціями. Статистичні дані користуються попитом:

- роботодавці - визначають, випускників яких навчальних закладів слід брати на роботу в першу чергу.
- керівники навчальних закладів - визначають, яких викладачів слід запросити на роботу.
- абітурієнти - вибирають навчальний заклад для вступу.
- батьки учнів - планують подальше навчання своїх дітей.
- студенти - визначають, до якого викладача слід піти для виконання дипломної роботи і виробничих практик.
- органи управління освітою - регулюють цінову політику і визначають плани розподілу субсидій.

Абітурієнти, учні, претенденти на заняття вакантних посад і діючі фахівці регулярно проходять атестацію. Як правило, всі особи, що успішно пройшли атестацію отримують документ, що засвідчує вимірний рівень їх ЗУН.

Претенденти, які успішно пройшли атестацію отримують сертифікат. Фахівців або випускників освітніх установ, що мають сертифікат про якість і кількість своїх ЗУН, називають сертифікованими.

Атестації або сертифікації піддаються навчальні заклади. На відміну від Росії, де ця процедура спрощена і складається з одноразового заповнення і перевірки величезної кількості красиво заповнених таблиць, за кордоном до атестації часто підходять як до перманентного, тривалого заходу. Гасло тут було б таким: «Атестація через все життя!». Моніторинговими центрами ведуться опитування, заповнюються бази даних про випускників тощо. Увага приділяється і газетним скандалам, і журналістським розслідуванням про екзаменаційні підробки і нецільовому витрачанні коштів в системі освіти.

Спеціальні установи, де проводиться атестація установ, ЗУН фахівців, зазвичай називають сертифікаційними центрами.

Сертифікати можуть мати різні ступені (рейтинги), відповідно до рівня якості який вони засвідчують. Рівні сертифікатів (дипломів, атестатів тощо) можна порівняти по рейтинговим таблицям, що щорічно публікуються.

Служба працевлаштування не має права віддати робоче місце фахівця з сертифікатом нижчого рейтингу, якщо в черзі є фахівець з сертифікатом більш високого рейтингу. У разі порушень прав фахівця, за нього успішно вступають судові органи.

Аналогічно, сертифікат охороняє право освітньої установи на отримання субсидій певного рейтингового рівня.

Як правило, процедура атестації включає обов'язкове бланкове або комп'ютерне тестування. Сертифікаційний центр веде свою діяльність від імені

- держави;
- великого освітнього закладу;
- недержавної установи.

У залежності від свого статусу, сертифікаційний центр видає претендентам (випробуваним), що успішно пройшли атестацію

- державні сертифікати;
- сертифікати державного зразка;
- посвідчення;
- інші документи.

Вважається, що якість діяльності сертифікаційного центру буде вищою, якщо він незалежний, тобто досліджувані ним структури ніяк не впливають ні на оснащеність, ні на рівень винагород співробітникам центру.

Важко забезпечити незалежність освітнього сертифікаційного (або моніторингового) центру в Росії. Співробітники центру - такі ж люди, як всі навколишні, з такими ж мало захищеними і забезпеченими правами. У них є діти, яким потрібно вступати до вузу чи переводитися з школи в школу. Отже, не дивлячись на всезагальне прагнення до принциповості, тут є всі передумови до традиційного російського бартеру «ти мені - я тобі».

11.9 До розвитку технологій тестування

Існуючі теорії тестової оцінки якості та кількості поки що є далекими до завершення. Тим не менш, побудовані на їх основі технології отримали широке поширення в Росії та світі загалом.

В Росії спостерігається бурхливе впровадження існуючих (часто критикованих професіоналами) технологій тестування, а стримане ставлення до розвитку теорії пояснюється:

- традиційною здоровою недовірою педагогів до будь-якої автоматизації навчального процесу;
- прагматичністю цілей, що ставляться керівниками освітніх установ на перший план.

Прагматичний підхід до оцінки ЗУН, що прийшов до Росії разом з реформами останніх років на зміну існуючому зовсім нещодавно бережному ставленню, знизив рейтинг російської освіти в світі. Про це говорять результати міжнародних досліджень.

Якість середньостатистичного випускника російських освітніх установ тепер нижче, ніж два-три десятиліття тому. Серед роботодавців зміцнилася думка, що на вкладиш диплома спеціаліста можна тепер не звертати уваги. Роботодавці стали самі екзаменувати, по-своєму розумінню, прийнятих на роботу фахівців. Відповідальність освітньої установи перед державою за якість освіти постійно підміняється особистою відповідальністю роботодавця за правильний вибір робітника.

Але не все втрачено. Відповідальне державне ставлення до розвитку теорії оцінювання і впровадження нових комп'ютерних технологій атестації ЗУН допоможуть повернути російській системі освіти колишню славу.

Навряд чи можливо досягти помітного підвищення об'єктивності вимірювань в рамках існуючих технологій тестування. У цьому переконують багато професійних досліджень, проведених самими творцями технологій та їх колегами. Майбутні успіхи у справі об'єктивізації пов'язують зі створенням:

- нових підходів до конструювання тестових завдань і тестів;

- нових форм представлення тестових завдань;
- нових технологій оцінювання якості і кількості ЗУН.

Серед перспективних напрямів розвитку форм тестових завдань можна назвати так званий грейдінг (від англ. - grading - сортування, класифікація) - методику, в якій оцінка виконання тестових завдань є диференційованою (ступеневою). При застосуванні методики grading, спектр оцінки ширший, ніж двозначний (вірно - невірно).

Застосовується grading частіше при тестуванні знань з математики. Визнаного україномовного (російськомовного) аналога поки що немає (перша згадка, яка відома нам про цю технологію належить В. С. Аванесову).

Іншим варіантом розвитку технології тестування є перехід від найпоширенішої, вже звичної і зрозумілої лінійної схеми, що використовує завдання в закритій формі, до іншої - схеми «м'якого тестування» (перевести цей термін можна як «Soft testing» або «Soft grading»).

Сам термін «м'яке тестування» означає, що з технології усунені надто різкі межі оцінки варіантів виконання тестових завдань. Замість двоступеневої шкали «вірно - невірно», тут використовується чотириступінчаста: «вірно - неповно - неточно - невірно». Така шкала більше відповідає людській логіці, ніж двозначна.

Тут можна використовувати вже наявні бази завдань закритої і відкритої форми. Однак, часто вони вимагають «косметичних» коректив. Корективи внести нескладно, провівши в кожному завданні сортування невірних варіантів відповідей на неправильні, неточні і неповні. Нижче наведено приклади, як це можна робити для різних навчальних дисциплін.

Змінено тут і форму подання тестових завдань. Випробуваному пред'являється вибірка тестових завдань закритої форми, причому, виконуючи завдання, він бачить на екрані - не тільки один варіант відповіді на питання. Саме цей, єдиний, варіант і повинен оцінити випробуваний. Решта варіантів приховані, і можуть відкриватися випробуваним по черзі, не даючи можливості

скористатися простим логічним порівнянням. Такий підхід дозволяє значно підвищити достовірність вимірювання.

Процедура пред'явлення завдань випробуваному оформлена тут у вигляді ділової рольової гри. Суб'єкт тестування грає роль вчителя, який проводить усне опитування у віртуальному класі. Інтерфейс оформлений відповідно - учні, дошка, парти. Учні можна перепитувати і виправляти вже виставлені оцінки.

Описаний варіант сценарію реалізований колективом Центру нових освітніх технологій в 1995-1996 рр.. (Цнот ДСДУ) в електронному тренажері «STeacher».

Модернізований тренажер, який реалізує технологію «м'якого тестування», створений колективом Тихоокеанського інституту дистанційного навчання та технологій (ТІДНТ ДСДУ) і в даний час застосовується для тестування знань студентів і школярів, модернізований тренажер отримав назву «Дидактор».

У 2004 році створена версія тренажера «Дидактор-6», де реалізовані мультимедійні можливості, необхідні, наприклад, при вивченні мов. «Дидактор-6» дозволяє вести тестування з використанням відкритих тестових завдань і в нетрадиційній формі, наприклад, як диктант.

11.10 Адаптивне тестування

Адаптивне навчання і адаптивний тестовий контроль, у своїй єдності, являють собою сучасний комп'ютерний варіант реалізації принципу індивідуалізації навчання. Цей принцип не міг бути реалізований ефективно в умовах традиційної класно-урочної форми занять. Початок адаптивного навчання можна віднести до часу виникнення педагогічних праць Коменського, Песталоцці та Дістервега.

Адаптивне навчання:

- засноване на використанні оцінок, одержуваних при вхідному адаптивному тестовому контролі. За підсумками контролю вибирається перший навчальний матеріал, з якого й починається адаптивне навчання. Наступний за цим контроль виявляє міру засвоєння запропонованого матеріалу, і т.д.;

- допомагає відійти від класно-урочної форми навчання, відкриває можливості свободи вибору траєкторії навчання;
- вимагатиме визначення переліку навчальних і тестових завдань для перевірки підготовленості не тільки за мінімальними, а й більш високим вимогами;
- дозволяє забезпечити видачу навчального матеріалу оптимального рівня складності. Відомо, що легкі матеріали не володіють помітним розвиваючим потенціалом, а важкі завдання знижують навчальну мотивацію у більшості учнів;
- забезпечує знання без звичних, для масової освіти, численних прогалин в індивідуальній підготовленості учнів;
- досягає необхідної структури і бажаного рівня знань;
- дозволяє регулювати трудомісткість і кількість пропонованих завдань в залежності від відповіді учня на поточне завдання. У разі правильної відповіді наступне завдання він отримає важче, у разі неправильного - легше.

Доцільність адаптивного контролю впливає з міркувань раціоналізації традиційного процесу тестування, де всім претендентам дається однаковий набір завдань. Підготовленому суб'єкту тестування немає необхідності давати легкі завдання через високу ймовірність їх правильного рішення. Через високу ймовірність неправильного рішення немає змісту давати важкі завдання слабкому претенденту. Використання завдань, що відповідають рівню підготовленості, істотно підвищує точність вимірювань і мінімізує час індивідуального тестування.

Три варіанти адаптивного тестування:

1. Пірамідальне тестування - всім претендентам дається завдання середньої складності і вже потім, в залежності від відповідей, кожному дається завдання легше чи важче. На кожному кроці застосовується принцип поділу шкали складності завдань навпіл. Наприклад, якщо в тесті є двадцять одне впорядковане за складністю завдання, тестування починається з одинадцятого. Якщо претендент діяв правильно, йому пред'являється шістнадцяте завдання і,

в разі успіху, дев'ятнадцяте. Аналогічно, слабо підготовлений претендент після невдачі на одинадцятому завданні пробує сили на шостому, а далі, за тим же принципом, до моменту стабілізації близьких за складністю завдань;

2. Flexilevel - коли контроль починається з будь-якого рівня складності, а потім відбувається поступове наближення до реального рівня підготовленості;

3. Stradaptive - коли тестування проводиться з БТЗ, де завдання розділені за рівнями складності. Після вірного виконання, наступне завдання береться з більш високого рівня складності, після невірного - навпаки. Це вимагає попереднього визначення складності всіх завдань.

У адаптивному тесті заздалегідь відомі параметри складності і диференціюючої спроможності кожного ТЗ. Він створюється у вигляді БТЗ, де завдання упорядковані у відповідності з їх характеристиками. Головна характеристика завдань адаптивного тесту - це їхня складність, отримана емпірично. Це означає: кожне завдання адаптивного тесту пройшло емпіричну апробацію на заданій вибірці, яка відповідає передбачуваним характеристикам претендента.

Адаптивний контроль дозволяє раціоналізувати традиційне тестування. У рамках адаптивного контролю ефективно вибираються і використовуються ті завдання БТЗ, які відповідають рівню підготовленості претендента. Це дозволяє істотно підвищити точність вимірювань і мінімізувати час сеансів до, приблизно, 5-10 хвилин.

11.11 Незвичайні форми тестування

Цей текст сповнений реклами, тому ми вважаємо, що ні автор, ні компанія Google, ні адміністрація CNET News.com на нас не образяться за цей передрук. Для тих, хто не знає: Силіконова долина - район США з найвищою концентрацією програмістів на одиницю площі.

В самому серці Силіконової долини (Маунтін-В'ю, штат Каліфорнія) на виїзді з Ралстону, що веде до Санта-Клари, з'явився рекламний щит з математичною головоломкою. Щоб вирішити її, більшості водіїв, що проїжджали по хайвею №101, довелося вдатися до допомоги Google.

Пошуковий гігант, що стояв за анонімною зухвалою порожньою рекламою з коротким чорним написом, прагнув лише привернути увагу людей з математичним складом розуму і запропонувати їм роботу. Цей рекрутський хід однозначно замішаний на відомому афоризмі математика Пола Ардос: «Математик - це пристрій для переробки кави в теорему».

Напис на рекламному щиті говорить: {first 10-digit prime found in consecutive digits e}.com. Відповідь, 7427466391.com, приводить математичних геніїв на веб-сторінку, де без всякого натяку на те, що за цією грою стоїть Google, пропонує вирішити ще одне рівняння.

Перемога над цим рівнянням веде на сторінку Google Labs, відділення досліджень і розробок компанії, де говориться: «Одна з речей, які ми засвоїли, створюючи Google: те, що шукаєш, легше знайти, якщо воно саме шукає тебе. Ми шукаємо кращих у світі інженерів. І ось ви тут. До нас щодня надходить безліч резюме, і ми придумали цей нехитрий процес, щоб поліпшити відношення сигнал/шум».

Google методично підходить до найму людей. Компанія, в якій станом на 31 березня 2004 працювало 1907 співробітників, з першого дня існування дуже серйозно відноситься до своїх методів. Один з перших співробітників фірми, що працює в ній ще з тих часів, коли та розміщувалася в гаражі в Пало-Альто, розповідає, що під час його інтерв'ю зі співзасновниками Google Сергієм Бріном і Ларрі Пейджем на сусідньому столі лежало не менше восьми довідників по найму людей. Рекрутери компанії теж кажуть, що перш ніж прийняти когось, вони проводять безліч інтерв'ю, іноді із залученням тестів з математики або бізнес-стратегії.

Ще один дотепний спосіб рекрутингу Google полягає в тому, що для виявлення потенційних працівників компанія розміщує спонсоровані посилання на сторінках результатів пошуку. Наприклад, у свій час таке посилання було пов'язане з ключовим словом/ім'ям Udi Manber - це шеф А9, нового відділення пошукової технології Amazon.com.

Цей приклад повчальний, його неважко адаптувати в повсякденній освітній практиці. Оголошення «загадки» можуть розміщатися на шкільному сервері, на сайті школи в Інтернет, на стінах шкільних коридорів.

РОЗДІЛ XII. Програми для створення тестів

Напевно будь-яка доросла людина хоча б кілька разів у житті проходила тестування. Іноді тестування було розвагою, іноді доводилося відповідати на запитання, щоб здати іспит або влаштуватися на роботу. Тести настільки популярні, що існує досить багато спеціалізованих додатків і онлайн-сервісів для їх створення.

Основна, але далеко не єдина категорія людей, яким може стати в нагоді програма для створення тестів, - це викладачі навчальних закладів. Різноманітні опитувальники постійно використовуються також практикуючими психологами, працівниками відділів кадрів, які займаються відбором кандидатів на вакантні посади в компанії тощо. Тестування часто буває необхідне і в тих випадках, коли в організації виконується перевірка рівня професійної підготовки співробітників у тій чи іншій галузі. І, звичайно, тести активно використовуються веб-майстрами, адже немає кращого способу утримати відвідувача на вашому сайті, ніж запропонувати йому з'ясувати, до якого психотипу він належить, або дозволити йому перевірити знання в тій області, де він вважає себе експертом.

Розглянемо чотири різні програми для створення тестів та опитувань.

12.1 Quiz Press

Quiz Press 2.5.8 - прості тести для публікації в Інтернеті

Розробник: Sol Robots

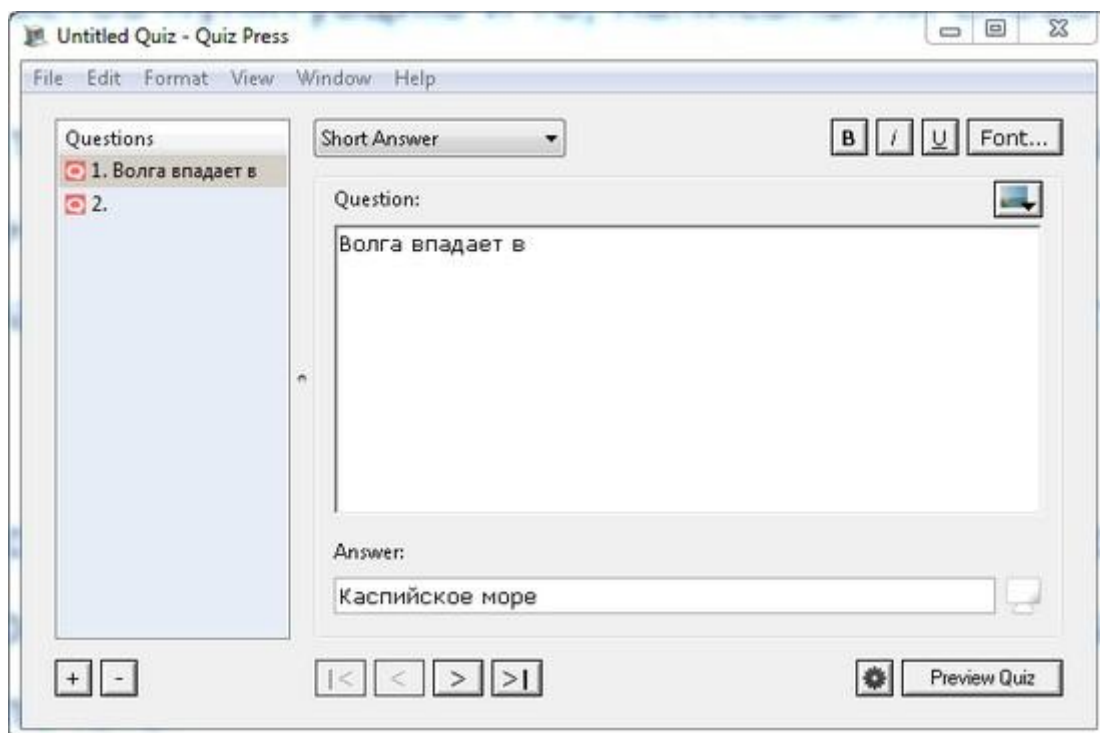
Розмір дистрибутива: 2,7 Мбайт

Поширення: shareware

Російський інтерфейс: немає

За допомогою Quiz Press викладачі можуть створювати тести для перевірки знань учнів. Програма дає можливість генерувати тести з питаннями різних типів: такі, в яких потрібно вибрати правильні відповіді із запропонованих варіантів, де потрібно вписати відповідь самостійно, де потрібно вказати, чи є твердження помилковим або вірним, та інші. Причому один тест може містити питання різних типів. При складанні питань у

викладача є можливість використовувати мультимедійні файли (графіку, флеш, звук). За замовчуванням при оцінюванні за кожну правильну відповідь студенту нараховується один бал, проте в налаштуваннях тесту можна встановити різну кількість балів, в залежності від типу питання. Крім цього є можливість обмежити час, який студент може витратити на вирішення завдань, додати нагадування про швидке завершення відведеного часу, вказати, чи потрібно враховувати при оцінці відповідей пунктуацію і те, написані слова з великої літери чи ні.



При створенні тестів можна швидко роздрукувати листок з правильними відповідями. Сам тест, створений за допомогою Quiz Press, представлений у вигляді набору файлів, готових для розміщення на веб-сайті. Крім цього, можна експортувати те, що отрималось в інші популярні формати тестів.

Результати повідомляються студенту в кінці тесту, а також пересилаються викладачеві в зашифрованому вигляді на адресу електронної пошти, вказану в налаштуваннях програми, або ж виводяться на друк.

12.2 Wondershare QuizCreator

Wondershare QuizCreator 4.01 - flash-тести і опитувальники

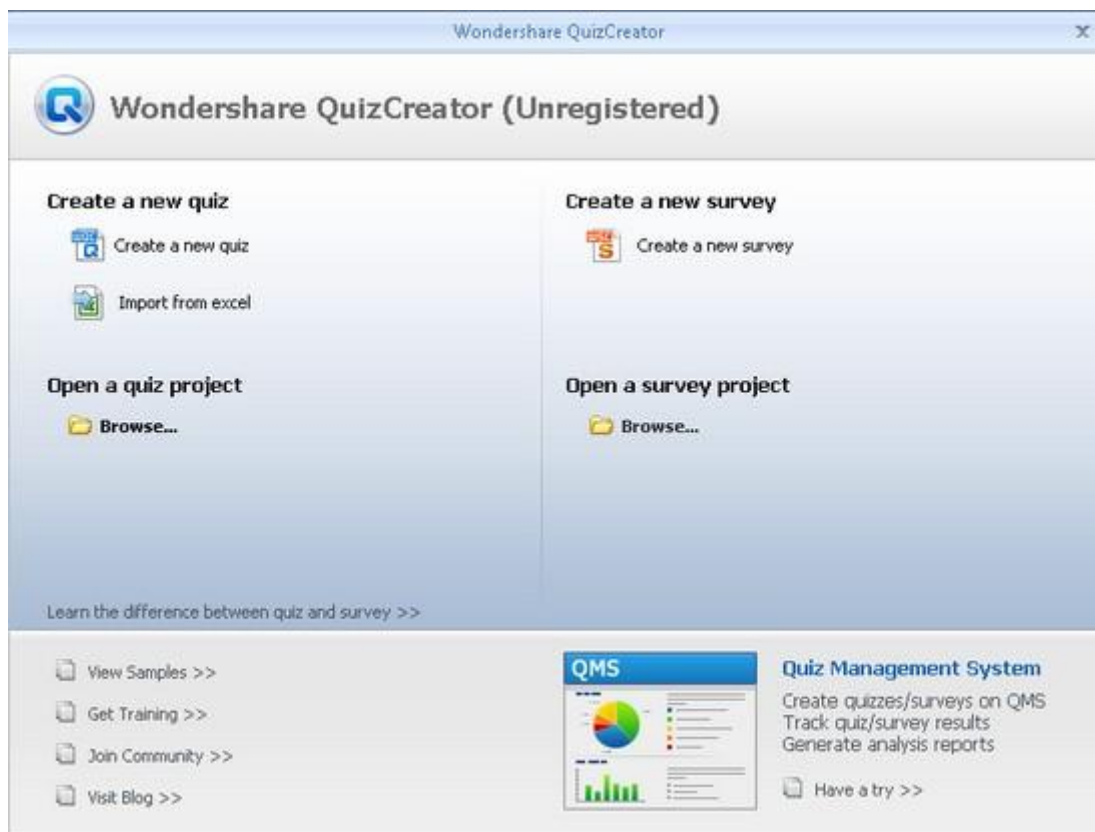
Розробник: Wondershare

Розмір дистрибутива: 11,5 Мбайт

Поширення: shareware

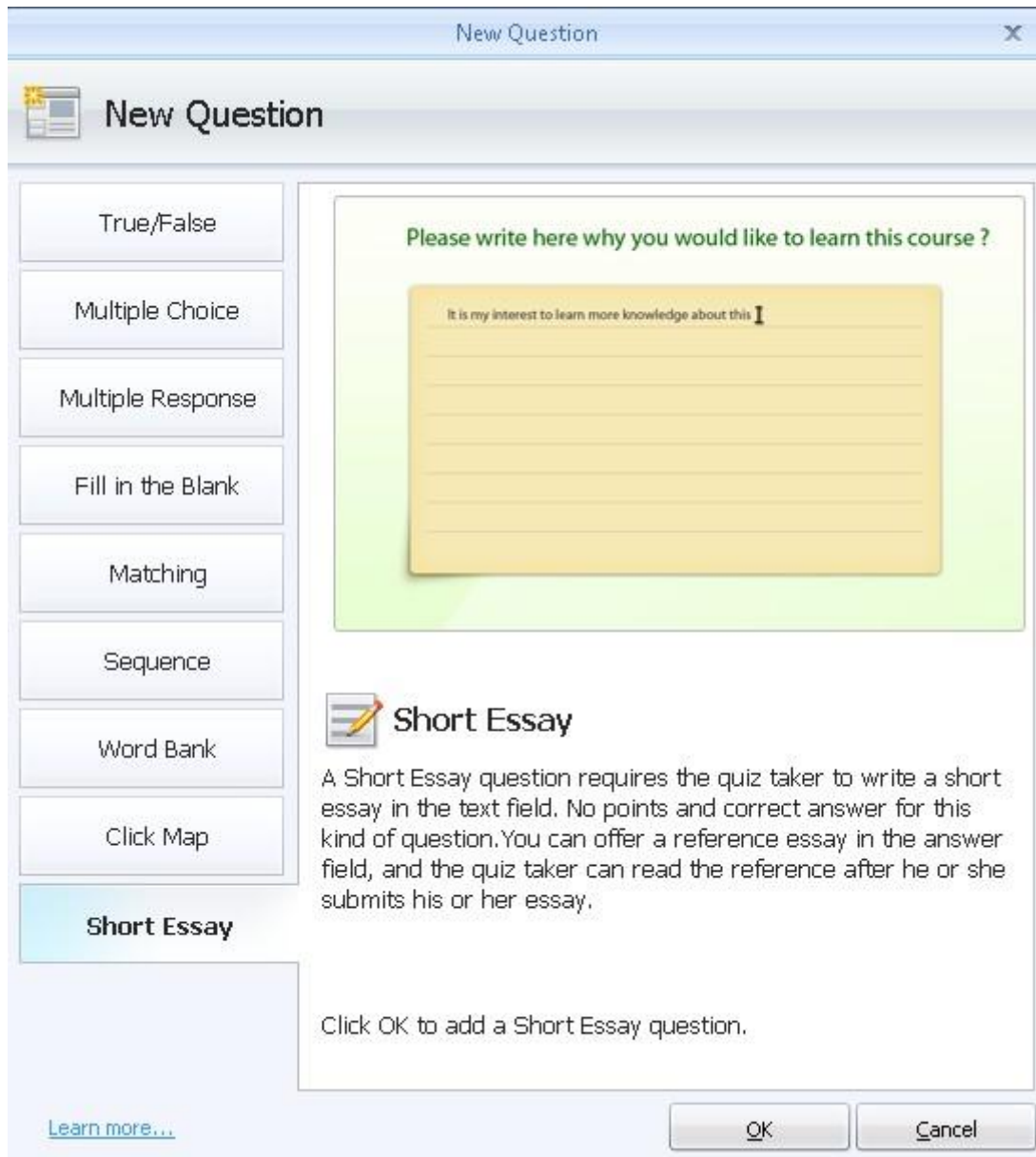
Російський інтерфейс: немає

Якщо потрібно створювати не тільки тести, а й звичайні опитування, в яких не потрібно враховувати правильність відповідей, можна звернутися до можливостей програми Wondershare QuizCreator. З її допомогою можна згенерувати опитувальник з найрізноманітнішими типами питань.



Піддослідні можуть відповідати «так» чи «ні», вибирати один або декілька варіантів відповідей із запропонованих, розташовувати варіанти відповідей у певному порядку, вписувати коротку відповідь в потрібне поле, співвідносити зображення з текстовими написами, підбираючи для них пару, і так далі. Ті ж варіанти питань доступні і при складанні тестів. Крім цього, тести можуть включати питання, відповіді на які заховані на зображенні. Цей варіант зручно використовувати, наприклад, при перевірці знань з географії. Викладач може додати в тест карту і завданням студента буде показати на ній певні області. Також можна створювати питання, для яких студенту не пропонуються варіанти відповідей. У цьому випадку відповідальний повинен вписати

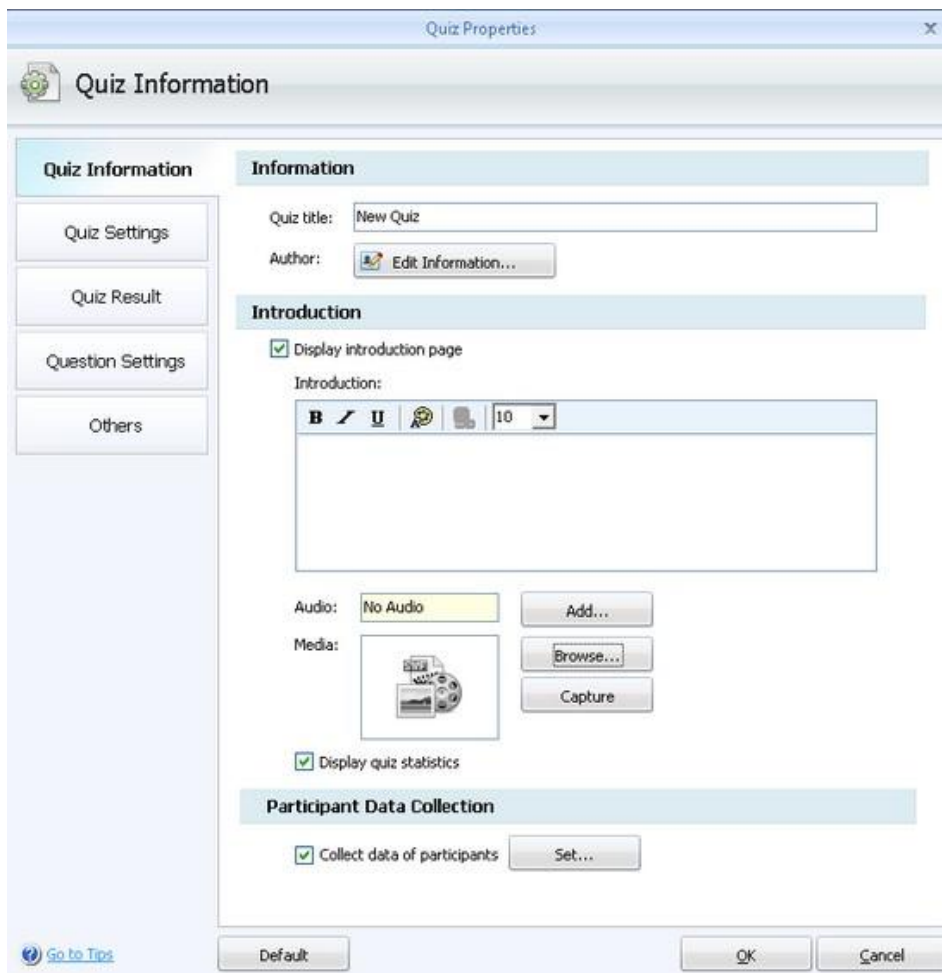
відповідь вручну, і він буде порівнюватися з тими варіантами, які викладач вказав при складанні тесту. У Wondershare QuizCreator є докладний опис всіх типів питань, тому заплутатися в них програма не дасть. Відзначимо також, що в програмі є можливість імпорту списку з питаннями з Microsoft Excel.



Кожне питання може містити не тільки текстову інформацію, а й різні мультимедійні дані. Причому вони можуть виступати як частиною питання, так і просто ілюструвати його. Наприклад, додаючи питання від теорії відносності у тесті з фізики, можна розмістити на сторінці фотографію Ейнштейна. Правильна відповідь на кожне питання за замовчуванням оцінюється в десять балів. При бажанні можна встановити певну кількість балів для кожного питання. Крім цього, є можливість забирати певне число балів від загального

результату при неправильній відповіді на питання, а також встановлювати кількість спроб, які даються студентові для пошуку правильної відповіді.

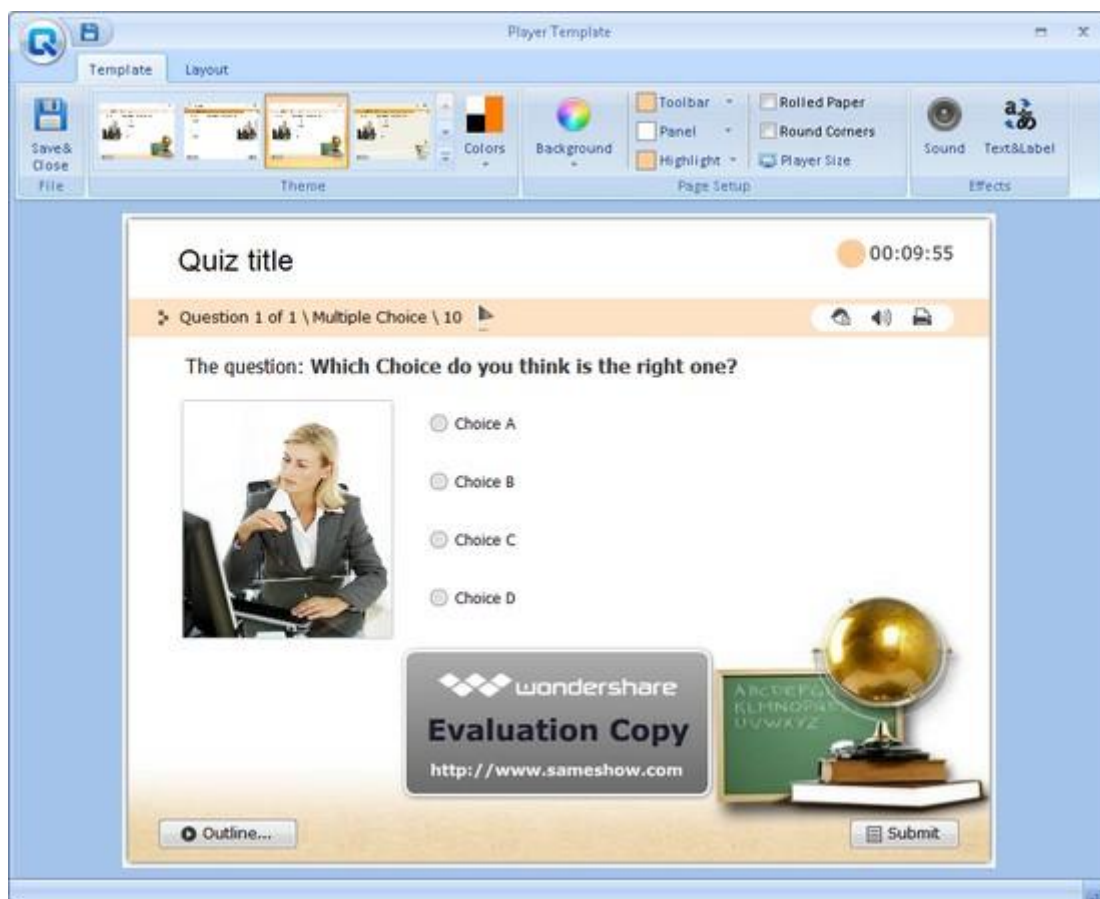
Після того, як питання будуть додані, програма пропонує визначити численні параметри тесту. Можна вказати автора та контактні дані, додати вступну сторінку з інформацією про тест (це може бути не тільки текст, але і графіка, звук, відео), визначити, як тест буде реагувати на дії користувача у випадку правильних і неправильних відповідей, встановити відсоток правильних відповідей, при досягненні якого тест буде вважатися пройденим, вказати, які повідомлення користувач повинен отримати у разі успішного проходження або провалу тесту. Крім цього, є можливість налаштувати перехід на задану веб-сторінку після завершення тесту.



Не забути і засоби для захисту. Так, для запуску тесту може знадобитися вказувати пароль або навіть дані облікового запису. Крім цього, є можливість дозволити запуск тесту тільки з певного домену, що дозволить захистити його

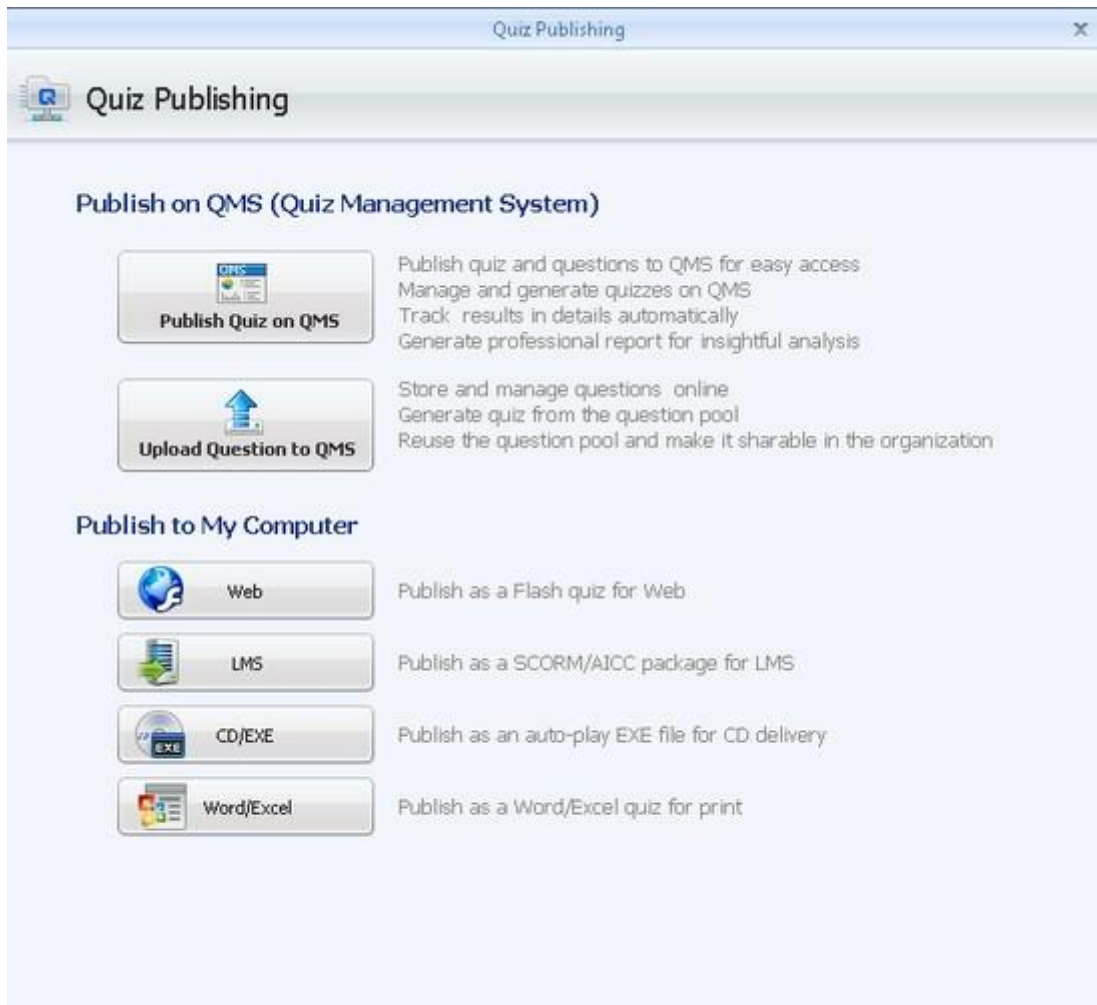
від копіювання при розміщенні в Інтернеті. Статистика замовчує дані про тих, хто проходив тест, чи не збирається його проходити, проте автор може зробити обов'язковим введення імені, адреси електронної пошти та інших даних. Також можна встановити ліміт часу для відповідей на певні питання або для тесту цілком, дозволити або заборонити перегляд своїх відповідей, повторне проходження тесту і перехід до наступного питання, коли ще не дано відповіді на попереднє.

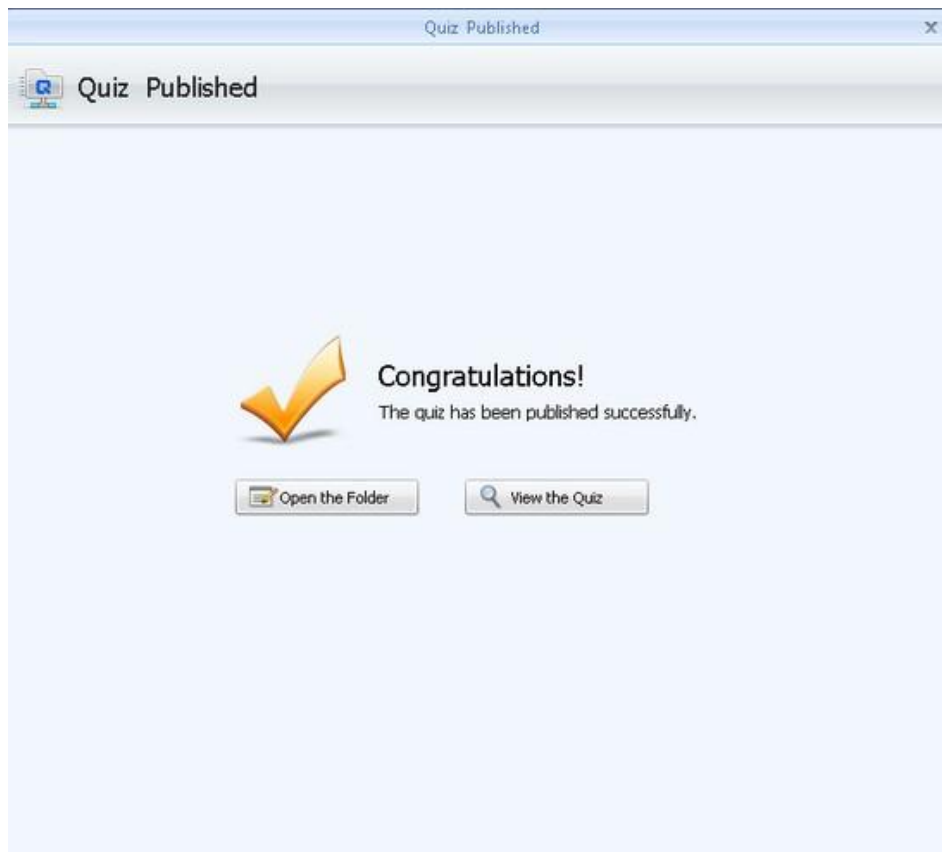
Багато уваги в Wondershare QuizCreator приділено зовнішньому вигляду тесту. У користувача є можливість вибрати одну з безлічі запропонованих тем оформлення, відредагувати її (змінити кольори і написи, прибрати елементи, визначити положення графічних даних тощо), додати аудіофайл, який буде програватися у фоні.



Згенерований у програмі тест може бути збережений в різних форматах в залежності від того, яким чином планується його поширювати. Наприклад, можна зберегти його у форматі Word / Excel для відправки на друк, записати диск з меню автозапуску. Однак найбільш сучасним і цікавим варіантом, поза

всяким сумнівом, є створення flash-версії тесту для розміщення на сайті. При виборі цього варіанту з'являється можливість вказати електронну адресу, на яку будуть відправлятися результати, а також дозволити відправлення результатів студентам.





Що стосується опитувань, то при їх створенні можна використовувати більшість описаних налаштувань і можливостей експорту. Виняток становлять ті, які актуальні тільки при оцінці рівня знань (різні оповіщення для правильних і неправильних відповідей та інше), що у випадку з опитуваннями неістотно.

І тести, і опитування можна зберігати в Мережі на сервісі Quiz Management System, який також є розробкою компанії Wondershare. Зручність цієї системи в тому, що з її допомогою можна поширювати тести тільки серед зареєстрованих користувачів, переглядати результати виконаних завдань, використовуючи різноманітні фільтри, генерувати звіти. А ще Quiz Management System підтримує імпорт питань, що створюються в Wondershare QuizCreator, і складання на їх основі нових тестів безпосередньо у веб-інтерфейсі.

12.3 Schoolhouse Test

Schoolhouse Test 3.1.6 - помічник для викладачів

Розробник: Schoolhouse Technologies

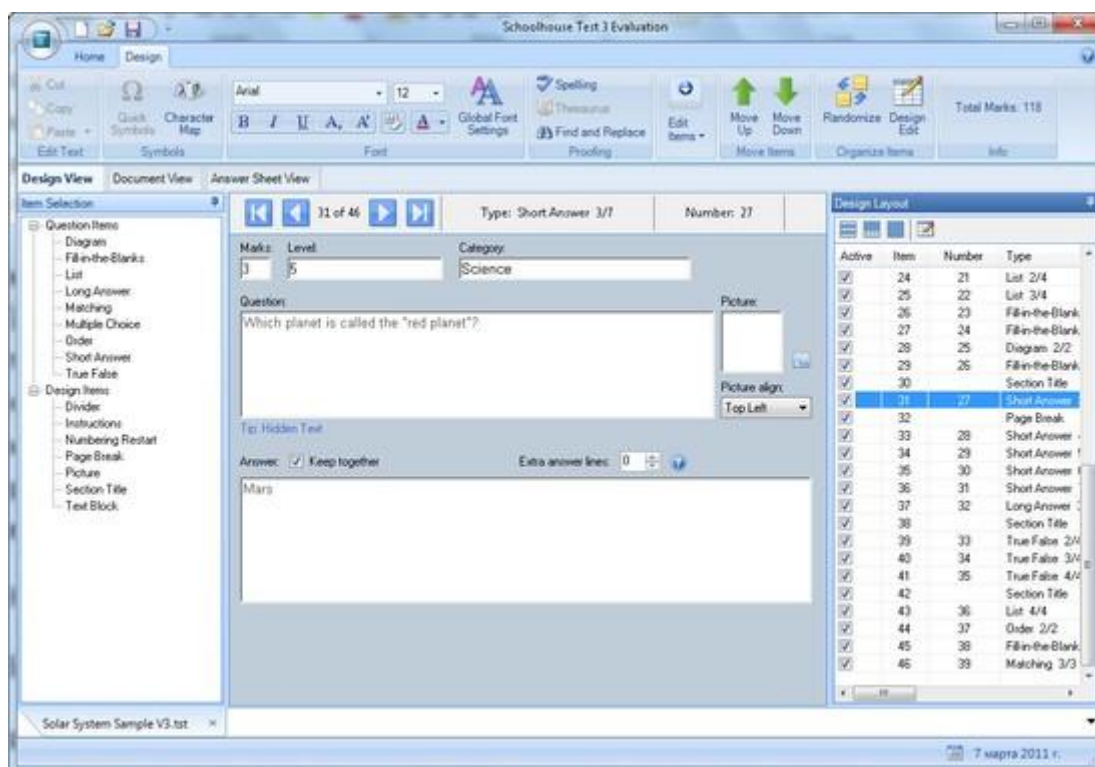
Розмір дистрибутива: 16,5 Мбайт

Поширення: shareware

Російський інтерфейс: немає

На відміну від попередніх програм, Schoolhouse Test орієнтований, насамперед, на викладачів навчальних закладів. Мета програми - спростити роботу з базою даних питань, а також максимально автоматизувати процес створення тестів і перевірки результатів. При цьому ніяких можливостей для створення онлайн-ових завдань не передбачено - програма дозволяє лише роздрукувати результати роботи, а також експортувати тести в PDF.

Робота зі Schoolhouse Test починається з формування бази питань. Додаючи питання, викладач може вказати галузь знань, а також рівень підготовки студента і число балів, які той отримає за правильну відповідь.

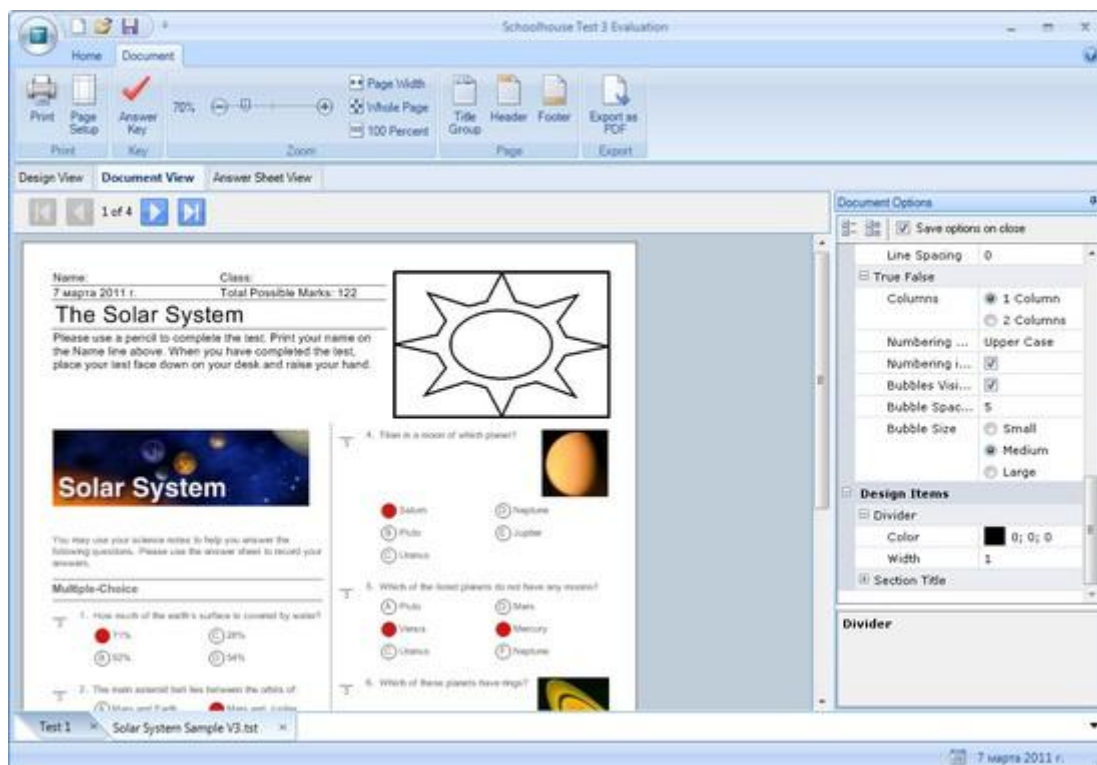


Питання можуть супроводжуватися графічними файлами, а крім цього, картинку можна прикріпити до будь-якого з варіантів відповіді. Schoolhouse Test підтримує самі різні типи відповідей, серед яких і такий незвичайний, як малювання студентом діаграми. Також є можливість створення тестів, у яких учню потрібно вставити пропущене слово. Для створення питань такого типу потрібно ввести фразу цілком, після чого клацнути по слову, на місці якого в завданні повинен бути пропуск. Schoolhouse Test відрізняється гнучкими засобами для управління всіма питаннями. Наприклад, при створенні завдань з

декількома варіантами відповідей кожен з них можна тимчасово відключити або виключити з перемішування випадковим чином.

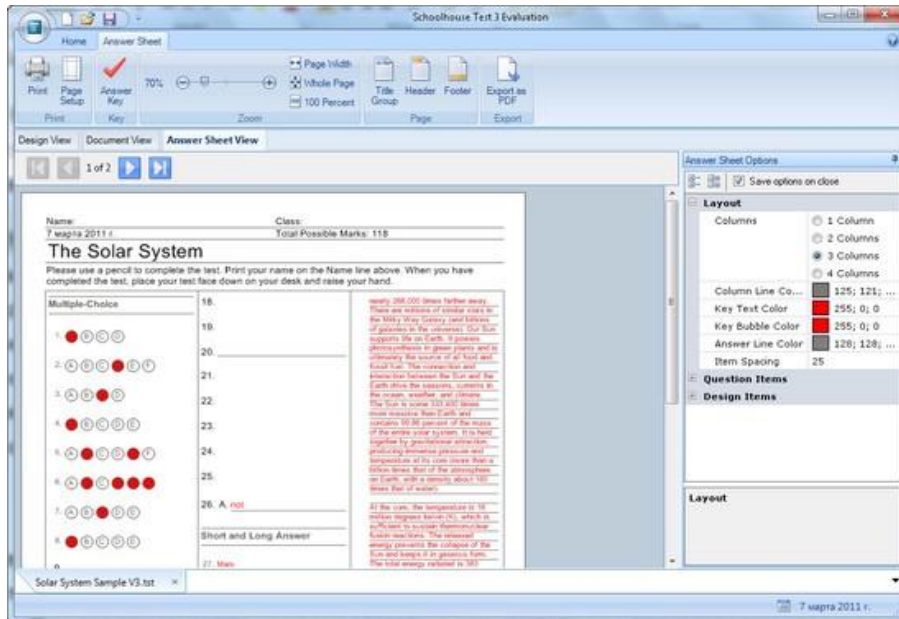
Коли база даних питань буде створена, викладач може швидко вказати, які з них необхідно включити в тест, а також створити декілька різних варіантів тестів з заданою кількістю питань у кожному.

Schoolhouse Test має три режими представлення даних. У першому - Design View - ведеться робота з базою питань. Тут викладач може відібрати питання для тесту, встановити їх порядок, додати супроводжуючі графічні файли, внести інші правки в зміст завдання. У режимі Document View можна побачити, як буде виглядати тест у друкованому вигляді, змінити його оформлення, колонтитули, параметри друку сторінки. Для зручності викладача в цьому режимі доступна можливість відображення правильних відповідей. Таким чином, можна швидко переглянути сторінки з завданнями і перевірити, чи не закралася чи помилка.



При переході в третій режим представлення даних - Answer View - програма генерує лист відповідей. Його зручно використовувати для швидкої перевірки тестів, а також для економії паперу. При його використанні

передбачається, що студенти отримають питання окремо і будуть заповнювати не листок з питаннями, а свій лист відповідей.



12.4 Adit Testdesk

Adit Testdesk 2.4 - повний комплекс рішень для тестування

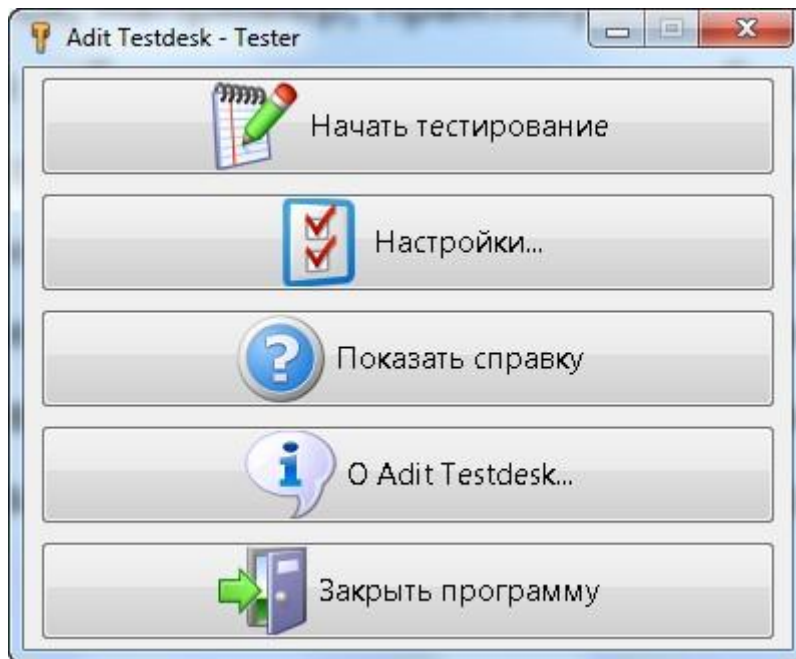
Розробник: Adit Software

Розмір дистрибутива: 41 Мбайт

Поширення: shareware (є декілька редакцій)

Російський інтерфейс: є

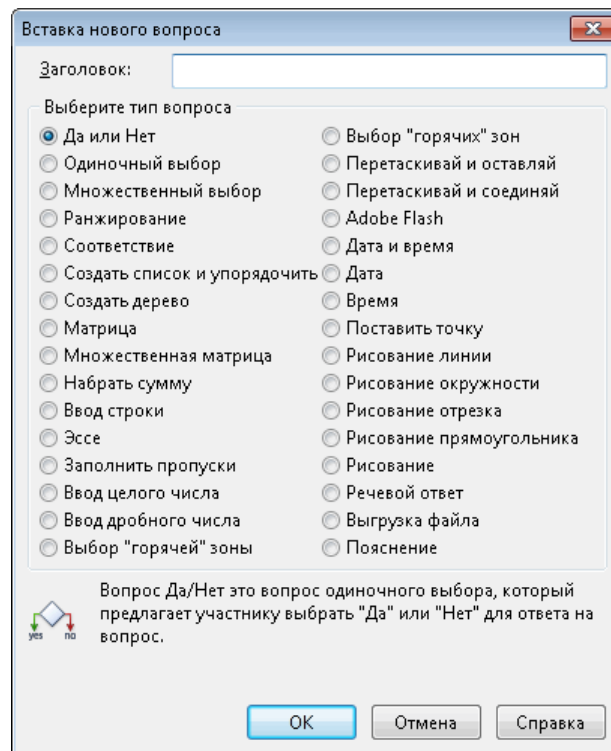
Програма Adit Testdesk - це універсальний засіб для створення тестів будь-якої спрямованості. Особливість цього рішення в тому, що за допомогою утиліти можна не тільки згенерувати тести, а й організувати тестування. До складу Adit Testdesk входить кілька утиліт: редактор тестів Editor, засіб для створення складних скриптових тестів з використанням мови програмування ScriptMaker, програма Reporter для оцінки результатів та їх експорту в формат XML, а також клієнт для проходження тестів Tester. При використанні мережевої редакції програми доступна також утиліта Testserver, за допомогою якої можна організувати тестування в локальній мережі або через Інтернет. Також є можливість підготувати друковану версію будь-якого тесту, використовуючи для цього вбудований текстовий редактор.



Adit Testdesk пропонує кілька основних типів проектів, які розрізняються за способом оцінювання. Тип "з правильними відповідями" підійде для створення тестів, в яких потрібно оцінити знання учнів або співробітників. Тип "психологічний" передбачає можливість встановлення певного числа балів для кожного з варіантів відповідей. Для складання більш складного алгоритму оцінювання можна використовувати тип "скриптова", а варіант "опитування", який підійде в тих випадках, коли правильність відповідей значення не має і є бажання дізнатися думку опитуваних.



Програма підтримує просто неймовірну кількість типів питань, багато з яких в інших додатках недоступні. Наприклад, можна створити такі питання, де в якості відповіді людина, яка проходить тест, повинна поставити крапку, побудувати граф, ввести час або дату, зіставити число кожному твердженню, вибрати "гарячі" зони на зображенні, намалювати коло, прямокутник, відрізок або лінію і так далі.

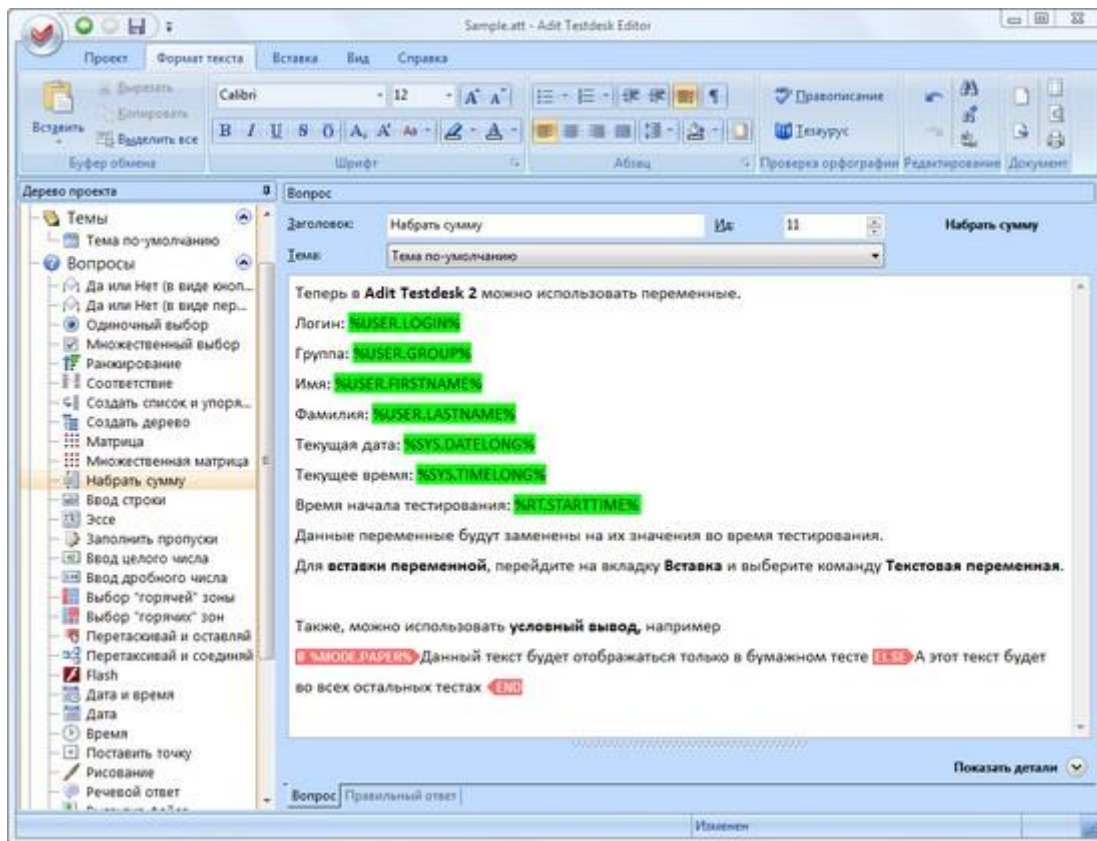


Складаючи питання, можна не просто вказати правильну відповідь або кілька варіантів відповідей, але і задати спеціальні умови для оцінювання його правильності. До кожного питання можна прикріпити графічний, аудіо- або відеофайл, а також створити навчальний приклад, представлений у вигляді дерева невеликих текстів.

Кожен проект, що створюється в програмі, містить не тільки перелік питань, але й також теми і профілі. Теми використовуються для більш зручної організації питань: кожне питання можна віднести до певної теми, після чого їх стає значно простіше шукати. Крім цього, якщо обмежується час на проходження тесту, то можна окремо встановити ліміт на питання, які відносяться до різних тем.

Профілі - це дуже зручний засіб управління великою базою питань. Завдяки профілям можна створювати схожі тести з різним оформленням і з різними налаштуваннями, а також з різними питаннями. Це може стати в нагоді, наприклад, якщо викладач створює кілька варіантів одного і того ж тесту для групи студентів. Оформленням тестів можна гнучко управляти за допомогою модуля "Дизайнер тексту". З його допомогою можна змінити колір фону, шрифт, розташування різних елементів, додати власні написи або логотипи.

Окремо варто згадати про можливості використання в тестах змінних. Вони дозволяють змінювати питання або інші параметри тесту саме в процесі проведення тестування, в залежності від того, які відповіді дає користувач. Наприклад, якщо опитуваний вказує, що він - чоловік, йому показується одна група питань, а якщо повідомляє, що він - жінка, питання виникають зовсім інші. Крім цього ці змінні зручні при оцінюванні психологічних тестів, де при виборі тої чи іншої відповіді можуть вилучатися бали від вже набраної суми або додаватися бонус. Змінні можна використовувати і для надання допоміжної текстової інформації, яка виводиться в процесі проходження тесту. Наприклад, якщо при проведенні оцінки знань отримані незадовільні результати, за допомогою змінних можна вивести текст з посиланнями на навчальні матеріали, які повинні допомогти студентові в подальшому навчанні.



Багато уваги в Adit Testdesk приділено безпеці. Всі тести зберігаються в зашифрованому вигляді і автор тесту може встановити права на перегляд і зміну даних. Скажімо, одним користувачам можна дозволити додавати нові профілі тестування, інші можуть вносити зміни в питання, а ще одна група має право лише переглядати готові тести. Доступ до тестів може бути обмежений IP-адресою користувача, ідентифікатором групи і логіном. Також можна встановити квоти тестування, наприклад обмежити число пройдених тестів для користувача двома на добу. Для тих, хто планує активно використовувати Adit Testdesk для тестування через Інтернет, існує спеціальна безпечна мережева редакція. У ній реалізована захищена передача даних з використанням протоколу SSL, а також безпечне віддалене адміністрування з підтримкою протоколу SSH.

Незважаючи на те, що всі чотири утиліти, розглянуті в цьому огляді, використовуються для однієї і тієї ж мети - створення тестів, у кожній з них своє призначення, яке визначається, перш за все, тим, які варіанти використання готових тестів передбачили розробники. Так, результати роботи в

Quiz Press можна зберегти лише у вигляді набору файлів для завантаження на FTP-сервер, тому очевидно, що програма може використовуватися для створення тестів, які плануються розміщувати в Інтернеті.

Wondershare QuizCreator може зберігати тести та опитування в форматі Flash, що буде зручно для веб-майстрів, які бажають урізноманітнити інформацію на своєму сайті. За допомогою цієї програми можна також записувати тести на диск. Schoolhouse Test призначений виключно для створення паперових тестів і є зручним інструментом для організації питань в зручному для викладача вигляді. Нарешті, Adit Testdesk пропонує для проведення тестувань власний клієнт, а значить, орієнтований, насамперед, на використання в організаціях та навчальних закладах, де потрібно реалізувати повний цикл тестування.

РОЗДІЛ XIII. ПРОБЛЕМИ АДАПТИВНОГО ТЕСТУВАННЯ

При формування світової освіти (освіта перестає бути внутрішньою справою окремої країни) необхідно виробити єдині стратегії в оцінці якості освітніх послуг. Система оцінювання знань і умінь, наявна в кожній країні, вже не є статичною, замкнутою в собі самій. Однак реформування системи освіти неможливе без реформування системи оцінювання, контролю якості освіти.

Важливе значення в даний час набуває культура оцінки та оцінювання якості навчання. Оцінювання - це процес збору і обробки дієвої та надійної інформації, яка дозволяє учасникам освітнього процесу приймати рішення, необхідні для вдосконалення роботи та досягнення кращих навчальних результатів. Своєрідною базою даних для оцінювання є державні стандарти, програми для викладача і для студента. Механізмами ж контролю та оцінки ефективності освітнього процесу стають тести.

Всім відомо, що зараз в системі освіти Росії паралельно з традиційною системою оцінки і контролю результатів навчання склалася нова парадигма - тестування. Це викликано потребою суспільства в отриманні незалежної, об'єктивної інформації про навчальні досягнення учнів.

Великі за останні десятиліття масштаби використання тестів у багатьох країнах світу призвели до створення компаній та служб, які займаються розробкою тестів, їх розповсюдженням, організацією масового тестування, а також регулярним збором інформації з метою моніторингу якості освіти. Західні країни зараз значно випереджають російську систему тестування з підготовки наукових кадрів, за рівнем розвитку теорії тестових методів, за технічною та програмно-обчислювальною оснащеністю процесів розробки, застосування, обробки та зберігання тестових завдань. Така ситуація відставання суперечить тенденції досягнення високої якості вітчизняної освіти, хоча в останнє десятиріччя намітилися й позитивні зміни в цій області. Приміром, майже 10 років функціонує Російська державна система тестування з РСІ.

Стратегія та тактика вдосконалення систем контролю. Збільшення масштабів тестування в освіті та розробка державних освітніх стандартів породжують спектр проблем, умов і вимог, які націлені на вдосконалення систем контролю та оцінки якості підготовки учнів, оцінки самого процесу тестування, підвищення його ефективності шляхом адаптації в умовах негайного реагування на індивідуальні особливості підготовки піддослідних.

Спроби привнести елементи адаптивності в процеси навчання та контролю налічують багаторічну історію. До певної міри можна вважати, що основні встановлені ідеї адаптивності беруть свій початок у вченні великого чеського педагога-мислителя Яна Амоса Коменського. Розробляючи нові форми і методи навчання учень прийшов до висновку про пріоритет тих навчальних завдань, які відповідають природним можливостям дітей, стверджуючи тим самим необхідність посилення завдань, що стимулюють успіхи в навчанні. У наступні десятиліття ідеї Я. А. Коменського підтримувалися в різних формах навчання великою кількістю педагогів, в числі яких були і росіяни: К. Д. Ушинський, Л. М. Толстой і багато інших, які бачили можливість оптимізації навчання та контролю кожного з учнів у введенні в навчання різних ігрових форм, в активізації самонавчання та самоконтролю, у вивченні психологічних особливостей особистості кожного учня. Таким чином, вже в кінці XIX - початку XX століть у педагогіці простежується установка на облік індивідуальних особливостей учнів, на розвиток їх самостійного і творчого ставлення до навчання, надання їм можливості прояву своїх схильностей. Говорячи сучасною мовою, можна стверджувати, що у вказаний період були закладені перші фундаментальні ідеї адаптивного навчання і контролю. Щоправда, реалізувати ці ідеї у той час так ніхто і не спромігся.

Пізніше, в 20-і роки XX століття, коли в освітній системі нашої країни спостерігався період критики формалізму і схоластики, що залишилися в спадок від дореволюційної школи Росії, ідеї розвитку творчої самостійності учнів, самоконтролю результатів навчальної діяльності отримали широке поширення.

До числа перших радянських педагогів, які бачили перспективу розвитку освітнього процесу не у відмові від контролю, не в штучному перетворенні особистості учня, а в підвищеній увазі до індивідуально-психологічних особливостей учнів, можна віднести С. Т. Шацького та П. П. Блонського. Вони вважали, що важливіше за все створити ділові стосунки між вчителем і учнем, які виникають при виконанні завдань важких, але посильних для учнів. Отже, завдання педагога полягала в тому, щоб відібрати ці важкі, але все таки посильні завдання, які сприяють встановленню атмосфери співпраці в процесі навчання і контролю.

У наші дні вона оформилася у вигляді основних ідей адаптивного навчання і контролю, де адаптивність досягається шляхом відбору оптимальних за важкістю завдань для навчання і контролю результатів навчальної діяльності учнів чи студентів.

У зв'язку з цим виникає серйозна проблема методологічного характеру. Вона пов'язана з теоретичним обґрунтуванням спеціальних методів для виділення оптимальних за важкістю завдань при організації контролю і навчання кожного учня чи студента. Рішення проблеми передбачає пошуки відповідей на ряд питань, серед яких головним є: встановлення зв'язку між рівнем підготовленості учнів і характеристиками завдань, що виділяються для ефективно організації контролю і навчання кожного учня чи студента.

Ситуація в області адаптивного тестування. М. Б. Челишкова, відомий вчений у галузі педагогічних вимірювань і оцінки знань учнів, провела аналіз численних праць вітчизняних і зарубіжних дослідників в області тестування і зробила висновок про те, що нині відсутні значимі результати по встановленню зв'язку між рівнем підготовленості учнів і характеристиками знань, що виділяються для ефективно організації адаптивного навчання. У своїх дослідженнях М. Б. Челишкова пише про те, що звертаючи увагу на практику можна простежити негативну тенденцію, яка полягає в домінуючій загальній спрямованості традиційних процесів навчання та контролю на якогось «середнього учня», відсутнього в реальному житті. Загальновідомо, що всі учні

різні. Тому тенденція на «посередність» призводить до того, що у частини більш сильних учнів не виникає стимулів для отримання нових знань та їх розвитку. Інша частина учнів, у яких більш низький рівень підготовки, виявляється не в змозі опанувати новими знаннями навіть за допомогою викладача (Челишкова М.Б., 2001, с. 10-15).

Можна відзначити, що в останні роки в практиці освіти складається ситуація, під впливом якої традиційне тестування, здійснюване за допомогою стандартизованих тестів, поступово переростає в сучасні, більш ефективні форми адаптивного тестування, які ґрунтуються на відмінних від традиційних теоретико-методологічних засадах і інших технологіях конструювання та пред'явлення тестів.

У нашій країні теоретичні дослідження з проблем адаптивного тестування представлені ще недостатньо. Назвемо імена деяких вчених, що працюють у цьому напрямку: Абрамян Г.В., Аванесов В.С., Нардюжев В.І., Челишкова М.Б., Шмельов А.Г. За кордоном з початку 90-х років адаптивне тестування отримало широке застосування у сфері практики. Однак можна відзначити, що у вітчизняних та зарубіжних наукових працях з даної проблематики ще недостатньо широко представлені питання класифікації адаптивного тестування, його функції, теоретичні питання адаптивного навчання та адаптивного тестового контролю.

Очевидно, що в даний час проблема адаптивних тестів надзвичайно актуальна. Біля витоків адаптивного тестування лежало прагнення до підвищення ефективності тестових вимірювань, що, як правило, пов'язувалося із зменшенням числа завдань, часу, вартості тестування, а також з підвищенням точності оцінок, отриманих суб'єктами тестування за результатами виконання тесту. Дослідники бачили можливість підвищення ефективності в адаптації тестів, труднощі які враховував діапазон підготовленості суб'єктів тестування.

Ми повною мірою згодні з нашою американською колегою Бетті Лівер (Лівер Б., 2001), яка вважає, що адаптивним повинно бути не тільки тестування, а й сам процес викладання. У своєму дослідженні на цю тему вона підкреслює,

що в процесі навчання треба постійно думати про те, що є учні-синтетичні, а є аналітики; зустрічаються в групі імпульсивні та рефлексивні студенти; необхідно пам'ятати про те, що індуктивних учнів, наприклад, втомлюють пояснення викладача і засмучує необхідність сидіти протягом тривалого часу, слухаючи розповідь вчителя і не маючи можливості спробувати зробити це самим. Дедуктивні учні живуть правилами, люблять слухати пояснення викладача. Не варто очікувати, пише Б. Лівер, від учнів з домінуванням правої півкулі, що вони будуть контролювати правильність вживання іноземної мови, якщо вони неправильно говорять рідною мовою. Види діяльності і тести, засновані на внутрішньому контролі, будуть виконуватися такими учнями дуже погано. Учні з домінуванням лівої півкулі, навпаки, занадто контролюють свою промову, і спроби залучити їх у бесіду, особливо на перших етапах вивчення іноземної мови, рідко мають успіх.

Тому часто результати стандартизованих класичних тестів дають неправильне уявлення про те, що учень знає або що він вміє робити.

Враховуючи навчальні стилі учнів, викладачі можуть створити власні адаптивні тести, що дають набагато більше інформації про здібності та знання учнів, ніж тести стандартизовані.

Так, тести з відкритими питаннями, стверджує Б. Лівер, дають учням синтетичного типу можливість проявити свої творчі здібності та здатності до синтезу. Тести на вирішення певних завдань дозволяють учням-аналітикам застосовувати свої здібності до аналізу. Письмові тести дають достовірні результати як для учнів візуального типу, так і для учнів кінетичного типу, а візуальні тести, безумовно, краще застосовувати для учнів візуального типу. Імпульсивні студенти добре справляються з традиційними тестами з обмеженим терміном виконання; рефлексуючим студентам потрібні тести, в яких не заданий ліміт часу. Для учнів синтетичного типу важкими є тести з можливістю множинного вибору варіантів відповіді. Студенти-аналітики, вважає Б. Лівер, виграють від тих видів діяльності, які дозволяють їм заглибитися в деталі, проаналізувати, зробити висновок. Таким чином, зі

сказаного вище впливає, що американські вчені активно використовують дані психології в навчанні та оцінки знань і вмінь учнів.

Назвемо чинники, що формують, на наш погляд, нинішню ситуацію у сфері адаптивного тестування в Росії і сприяють актуалізації цього напрямку:

- загалом і в вищій російській освіті відбувається різке розширення обсягу змісту освіти: інформатизація всіх рівнів освіти, збільшення ваги творчого аспекту в навчанні на всіх етапах освіти;

- тенденція до зростання масштабів тестування в освітній системі нашої країни, де за останні п'ять років кількість школярів і студентів, що пройшли через підсумкове тестування, зросла з 35 тисяч до 80 тисяч осіб;

- розробка нового покоління державних стандартів, зокрема з окремих дисциплін (маємо на увазі РСІ), які, в свою чергу, актуалізували проблему масових вимірювань щодо рівня і якості підготовки учнів. Створення ж державних освітніх стандартів пов'язується, як нам здається, з проблемою конкурентоспроможності майбутнього фахівця, який повинен володіти багатоплановою, багаторівневою і багатоаспектною компетенцією, у тому числі в галузі іноземних мов і російської мови як іноземної.

Тестування, що проводилося русистами-практиками протягом останніх років в Російському університеті дружби народів (щорічно при переході з однієї освітньої сходи на іншу тестується понад 200 осіб) показало, що, дотримуючись принципів російської методики викладання РКІ з її особистісно-орієнтованим підходом до навчання і для підвищення якості тестового контролю, для додання тестуванню динаміки, необхідно враховувати індивідуальні особливості як самих суб'єктів тестування, так і особливості їх підготовки.

У зв'язку з цим однією з актуальних завдань методики за останній час стало завдання репрезентативного охоплення змісту контролю в тестах з РСІ. Практика тестування показує, що вимоги, закладені в державних стандартах з російської мови як іноземної і у відповідних їм тестах нерідко неузгоджені з реальними можливостями учнів. Так, в Російському університеті дружби

народів навчається велика кількість студентів з країн Азії, Африки, Латинської Америки. Викладачам-русистам добре відомі особливості даних учнів та пов'язані з цим труднощі при навчанні їх російської мови. Часом ці люди не вміють не тільки писати, але й читати рідною мовою, погано володіють уміннями спілкування. Завдання викладачів російської мови на етапі довузівської підготовки, як правило, переростає з практичної в освітню. У силу цього досить складно підготувати студентів з 160 країн світу, які розмовляють на 400 мовах, до об'єктивного та ефективного контролю.

Безумовно, викладачі-практики шукають шляхи підвищення якості підготовки учнів, намагаються запобігти перетворенню навчання російській мові в процес підготовки до підсумкового тестування. З цією метою створюються серії адаптивних тестів, що дозволяє учням звикнути до нових для більшості з них форм контролю, до нових завдань відкритої форми, в яких ставиться завдання для студента-іноземця за певними параметрами організувати свою відповідь. Зауважимо, що іноді не всякий відповідь носія російської мови не відповідає повною мірою вимогам і параметрам, що задаються у тестах для іноземних громадян.

Про облік у контролі національно-орієнтованого змісту навчання. Все більш актуальною, на наш погляд, стає проблема обліку не стандартного, а національно-орієнтованого змісту освіти і навчання російській мові в зарубіжних країнах. У різних країнах склався різний досвід викладання російської мови, обумовлений особливостями освітніх систем зарубіжних країн.

Наш аналіз цієї проблеми показав, що в так званих східних методиках (використовуються в країнах Азії: Китаї, Японії, В'єтнамі та ряд інших) навчання граматиці, лексиці, читання, письма переважає над навчанням усного мовлення. Припустимо, що деякі національні «стереотипи» можна подолати за час інтенсивного навчального навантаження з мови на підготовчому факультеті, враховуючи також, що студент живе в мовному середовищі. А як бути зі студентами, які живуть і навчаються у себе на батьківщині, де практика

викладання визначається освітніми нормами не країни досліджуваної мови, а рідної країни? В якості прикладу наведемо досвід тестування, проведеного нами в Японії в університеті міста Саппоро. Наші контакти, нехай дуже короткі, показали, що японці - досить скромні, ввічливі, тактовні, стримані. У їхній системі виховання, яка складалася століттями, закладені такі традиції, як згода зі співрозмовником (обов'язкова згода), правило м'яко висловлювати власну думку, вислуховувати думку іншого (японець намагається висловлювати думку колективу). Очевидно, тому важко діалогу студентам складання тесту з завданнями такого типу: «заперечити співрозмовнику», «докорити співрозмовника», «висловити загрозу». Менталітет японців не дозволяє давати мовному партнеру такі жорсткі установки. Закордонна практика тестування показує низьку ефективність таких завдань, як: «Висловіть захоплення, захоплення чим-небудь». У Китаї, Японії та ряді інших країн, де не прийнято виражати «бурхливі» емоції, мова людини (особливо при сторонніх, особливо при іноземцях, особливо у присутності викладача) дуже стримана і тактовна. Домогтися реалізації закладеної в тесті інтенції такого роду досить нелегко.

Складно також залучити, наприклад, арабського або індійського студента в діалог про їхнє ставлення до холодної зброї (матеріали типових тестів), тобто про те, потрібно чи не потрібно застосовувати його в особливих, критичних випадках. Цього просто може і не бути в життєвому досвіді учнів. Більше того, ми добре знаємо, що в багатьох країнах світу продаж зброї заборонено. Як у такому випадку викладачу досягти результатів у контролі навичок і вмінь випробуваного?

Безумовно, для тестування усного мовлення потрібно ретельно підбирати теми, характерні для національного менталітету, а не суперечливих йому.

Адаптивне ставлення до тестового контролю виражається і в обліку вікових особливостей учнів. Важко об'єктивно оцінити вміння в усному мовленні в учня за наступним завданням: «Вас зацікавила реклама чи оголошення в газеті, що розповідає про те, як виховують дітей у московському

дитячому садку, так як ви хочете влаштувати свою дитину в певний дитячий садок. Зателефонуйте за вказаною в оголошенні телефону і організуйте свій діалог на задану тему ». Зауважимо, що, по-перше, не для всіх іноземних студентів ясно, що таке російський дитячий сад, по-друге, на їхній батьківщині влаштування дітей у дитячий садок передбачає часто абсолютно інший набір питань, наприклад, обов'язкове питання про вартість цієї послуги. Крім того, постає питання - звідки студенту взагалі знати про те, що запитують батьки в такому випадку. Найбільш спритні випробовувані своє невміння організувати передбачуваний діалог жартівливо пояснював так: «Якби я був батько, я б не по телефону дізнавався про те, де буде виховуватись мій син або моя дочка, а поїхав би подивитися на місці» (подібне «дозвіл»-питання було запропоновано тестованими з Японії та ін.) На жаль, в наших тестах зустрічаються такого роду «проблеми» для випробуваного й викладача, який створює завдання. Звичайно, багато недоліків завдань можна усунути, зокрема, через орієнтацію тестового матеріалу не тільки на реалії України, але і на реалії тих країн, студенти (учні) з яких проходять тестування. Адаптація в даному випадку мається на увазі як ґрунтовну підготовку учнів, так і підготовку навчальних і контрольних матеріалів у плані співвіднесення мовних і культурних традицій різних країн. Тестування, безумовно, повинно бути адаптивним за своєю національною орієнтацією. Розвиток же національно-орієнтованих тестів та їх впровадження в систему тестового контролю, на наш погляд, покликаний стати завданням недалекого майбутнього.

Таким чином, досвід тестування з російської мови як іноземної учнів різних національностей, які оволодівають мовою як в Росії, в мовному середовищі, так і за кордоном, поза тим середовищем, переконливо підтверджує нашу тезу про те, що тестування з іноземної мови неодмінно має враховувати індивідуальні особливості учнів, перш за все їх національність, особливості менталітету, культурні та історичні традиції і т.д. У силу зазначених причин тестові матеріали повинні бути адаптивними.

Ще раз повернемося до розмови про сукупність обсягу тесту. Він, як відомо, включає 5 субтестів: по лексиці і граматиці, по говорінню, письма, читання та аудіювання. Ми підраховали, що в деяких тестах число завдань, включаючи лінійні, тобто такі, які відносяться до одного тексту або до одного із завдань (наприклад, по говорінню: 1) перегляньте фрагмент фільму; 2) опишіть ситуацію, 3) дайте опис дійових осіб; 4) висловіть власну оцінку того, що відбувається і ін), об'ємне. Чи завжди ефективно досягнення якості перевірки через збільшення кількості завдань? Очевидно, що кількість далеко не завжди вимірює якість. Тому говорячи про проблему адаптивного тестування, ми маємо на увазі, крім іншого і сказаного вище, проведення теоретико-практичних досліджень з проблем оптимізації процедури тестування. Під оптимізацією нами розуміється така побудова рівневого контролю, коли меншим числом завдань вдається швидше і точніше оцінити рівень і якість підготовленості тестованих. Здається, що для вирішення цієї проблеми буде потрібно ще чимало зусиль, а найголовніше - тісна співпраця практиків з теоретиками. У нас поки що мало теоретиків для наукового обґрунтування РСІ форми контролю.

Підводячи підсумок сказаного, дозволимо сформулювати висновок-визначення адаптивного тестування. Отже, під адаптивним тестуванням нами розуміється процес підготовки до контролю у формі тестових методик через навчальні тести градуйованого (від простого до більш складного) характеру і власне контроль, який коректно враховує індивідуально-особистісні, національні та культурно-орієнтовані особливості іноземних учнів, які оволодівають українською мовою в межах і за межами своєї країни, зокрема в середовищі мову якого вивчають.

РОЗДІЛ XIV. ОСНОВНІ ВІДОМОСТІ З ОБЛАСТІ ТЕСТОЛОГІЇ

14.1 Загальні відомості з тестології

Тестологія - (від англ. - Test - проба, випробування) - наука про вимірювання психофізіологічних та особистісних характеристик, а також кількості та якості ЗУН.

Тестологи вивчають і створюють способи, методи, технології вимірювань психофізіологічних та особистісних характеристик, а також кількості та якості ЗУН. Вчені створюють тестові комплекси, де реалізовані досягнення тестології у вигляді сукупності технологій, рекомендацій, тестів, автоматизованих систем, пристроїв.

Тестові комплекси застосовують для атестації навчальних досягнень учнів та абітурієнтів, для вимірювання кількості та якості навичок і вмінь при прийомі людей на роботу тощо.

Теорію тестових педагогічних вимірювань називають коротко IRT - Item Response Theory.

У IRT не ставляться і не вирішуються фундаментальні проблеми валідності та надійності тесту: тест там заздалегідь вважається надійним і валідним. Обчислення IRT зводяться до отримання оцінок параметрів складності завдання і до вимірювання рівня навчальних досягнень піддослідних. До досягнень IRT відносять використання однієї шкали у вимірах значень параметрів досліджуваних та завдань тесту. Це дозволяє співвідносити рівні ЗУН випробовуваних з мірою складності тестових завдань. На цих можливостях спільних оцінок параметрів досліджуваних та завдань заснований застосовуваний більшістю тестологів математичний апарат.

У нас поки не склалося уявлення про педагогічну тестологію як виділену в науці або науковій області:

- нема посади тестолога в реєстрах відділів кадрів;
- нема вакансій для академіків - тестологів в Академії Наук тощо;
- педагоги не сприймають тестології, оскільки в ній дуже багато математики;

- математики, фізики, інженери не сприймають тестології із-за властивих їй традиційних для педагогіки «туманних сентенцій»;

- програмісти вважають, що тестологія - пройдений етап, і там немає чого робити особистості, котра мислить творчо.

Зроблено одну зі спроб виправлення цього становища шляхом популяризації: основні положення і досягнення тестології викладено мовою зрозумілою педагогу без страшних математичних викладок та термінології. Іноді Ви зустрінете іноземне слово, але лише там, де тестологи не знаходять аналога на рідній мові.

14.2 Принципи конструювання тесту, створення тестових завдань і вибору технології тестування

Педагогічне тестування має досить добре розроблену принципову базу. Вона дуже добре описана в науковій і науково-методичній літературі.

1. Валідність тесту. Результати тестування групи претендентів повинні відповідати об'єктивним характеристикам, котрі дали суб'єктам тестування їх керівники, колеги, викладачі. Це найважливіший принцип, якому повинні підпорядковуватися зусилля упорядника тесту. Відповідно до безлічі зазначених характеристик, прийнято розрізняти види валідності.

2. Надійність тесту і технології тестування. Результати тестувань подібних груп претендентів за допомогою одного тесту повинні бути однаковими (подібними) і не залежати від часу. Це другий за важливістю принцип, якому повинні бути підпорядковані зусилля упорядника тесту. Частіше за все, надійність можна охарактеризувати лише інтуїтивно, оскільки:

- повторне тестування однієї групи претендентів по одному тесту призводить до зростання результатів за рахунок ефективності пам'яті, самонавчання тощо;

- дуже важко знайти подібні групи претендентів і довести, що повторне тестування однієї групи претендентів по одному тесту призводить до зростання результатів за рахунок ефективності пам'яті, самонавчання тощо;

- на результати впливає передісторія груп претендентів, технічно складно виключити попереднє спілкування між ними тощо.

Надійність - принцип, що прийшов у педагогіку з техніки. Це принцип, який висувається до інструменту технічного виміру. Але педагогіка - інша область з іншими закономірностями, і тому принцип надійності часто обгрунтовано піддається критиці (див. релевантність тесту).

3. Релевантність тесту і технології тестування. Виберемо два однакових за валідністю тести (або, точніше, дві технології тестування). З точки зору дидактики, наука підкоряє всі цілі зростанню якості навчання, якщо повторне тестування з використанням першого тесту (першої технології) призводить до більш вираженого поліпшення результатів, ніж з використанням другого тесту (другої технології), то перший тест (перша технологія) - краща. Нехай результати в обох випадках поліпшуються по-різному за рахунок різної вираженості ефекту самонавчання претендентів, привабливості процедури та ін. З точки зору принципу надійності, обидва тести (обидві технології) - не надійні. Чим кращий тест (технологія тестування) з точки зору дидактики, тим він (вона) гірші, виходячи з принципу надійності. Тому, в педагогіці, на відміну від техніки, краще говорити не про надійність тесту (технології), а про релевантність тесту (технології), тобто ступеня відповідності принципам дидактики. Релевантність тесту (технології тестування) - міра його здатності до ініціації процесів самоосвіти. Релевантність можна виразити чисельно, взявши відносну приросту усереднених результатів при повторному тестуванні з використанням різних тестів (технологій).

4. Наукова достовірність. У тест включаються тільки ті елементи знання та зв'язку між ними, які є об'єктивно істинними. Відповідно, спірні точки зору, цілком допустимі в науці, не включають у тест.

5. Узагальненість. Теми тестування не повинні бути надмірно узагальненими.

6. Значущість. Істотні ТЗ відображають структурні інформаційні елементи дисципліни або зв'язок між структурними елементами, без яких знання стають неповними, з пропусками.

7. Відповідність сучасним знанням. У ТЗ включаються тільки сучасні елементи знання. Винятком можуть бути ТЗ, складені спеціально для ведення навчальних занять процесу в рамках концепції «діалогу культур».

8. Відповідність джерел знань. Всі ТЗ повинні включати адресу джерела, яке може бути включено і продемонстровано частину тексту ТЗ у випадках, коли є різночитання в рекомендованій літературі.

9. Репрезентативність. Репрезентативний тест містить сукупність ТЗ, що відображають всі структурні елементи змісту дисципліни та їх зв'язки. Репрезентативна (представницька) БТЗ містить сукупність ТЗ, достатніх для ведення варіативного контролю по кожній із наявних у структурі дисципліни навчальних тем. Репрезентативна БТЗ може включати ТЗ з додаткових навчальних матеріалів для проведення олімпіад тощо.

10. Варіативність. Зміст тесту може і повинен змінюватись по мірі:

- зміни змісту дисципліни;
- зміни і появи нових наукових знань;
- зміни мети тестування;
- зміни спрямованості на особливі вибірки випробовуваних.

Для досягнення достатньої варіативності тесту в рамках навчальної дисципліни необхідно, щоб БТЗ містила тестові завдання, що відповідають різним навчальним посібникам і різним способам викладу навчального матеріалу.

11. Складність. Для кожного ТЗ в процесі контролю визначають усереднену, виражену числом, складність. Складність завдання № j визначають двома способами. Наприклад, так: $q_j=1-p_j$, де p_j - відносна кількість претендентів, які виконали дане завдання. Або, як «логіт», рівний $\ln(q_j/p_j)$. Претендент, правильно виконує складні завдання, з великою ймовірністю правильно виконає і легенькі. Складність тесту відповідає змісту навчання і

підтримує високий рівень мотивації претендентів. Не слід нехтувати завданнями, які виконуються великою кількістю претендентів (від 90 до 100%). Іноді може здатися, що завдання зайве, оскільки занадто легке. Треба проконтролювати вибірку – чи увійшли до неї слабо підготовлені претенденти. Вони можуть свідомо не допускатися до тестувань керівництвом установи під будь-яким приводом. Аналогічна, але зворотна ситуація може трапитися і з найбільш складнішими завданнями. Перш, ніж відкидати завдання, які ніхто не виконав, необхідно перевірити, чи проходили тестування найбільш підготовлені учні. Вони могли в цей час бути відсутніми внаслідок участі в олімпіаді або в роботі фізико-математичного гуртка.

12. Трудомісткість. Для кожного ТЗ в процесі контролю визначають усереднену, виражену числом, трудомісткість. Вона характеризує кількість елементарних операцій (і розумових, і фізичних, і розрахункових), які необхідно виконати в процесі роботи над завданням. Ставлення трудоємкості завдань можна визначити через усереднене ставлення часу їх виконання. Відносна кількість праці, яку тестовані витратять (у середньому) на виконання кожного тестового завдання, повинно бути пропорційно відносній значимості відображеного в завданні елементу курсу.

13. Виваженість. Для додання результату тестування більшої об'єктивності, всі ТЗ після сеансу «зважують». Потім загальний результат претендента обчислюють як суму ваг виконаних завдань. Іноді за вагу ТЗ приймають його відносну складність, тобто частку претендентів, які не зуміли його виконати. Іноді вагу завданням приписують спираючись на викладацький досвід. Обидва підходи до обчислення ваг не позбавлені недоліків. Наприклад, при першому підході можливий варіант, коли найвищу вагу отримає завдання, яке здалося всім претендентам занадто легким, недостойної одномоментної уваги, яку вони, в гонитві за рейтинговими балами, залишили «на потім» і не встигли до нього повернутися.

14. Системність. Завдання в БТЗ зв'язуються між собою загальною структурою знань. Такий зв'язок може бути визначений інтуїтивно або методами факторного аналізу.

15. Комплексність і збалансованість. Слід гармонійно поєднувати в тесті завдання на перевірку знань теоретичного матеріалу (поняття, закони, закономірності, гіпотези, факти, структурні компоненти теорії), методів наукової та практичної діяльності, умінь вирішувати типові завдання. Відношення кількості ТЗ перерахованих типів повинно відповідати відносним значенням та обсягам рекомендованої навчальної інформації.

16. Відповідність змісту і форми. Педагогічний тест характеризують як результат поєднання змісту завдань з найбільш придатною формою.

17. Роздільна здатність тесту. Складність ТЗ не може бути тільки легкою (всі виконали), середньою (половина групи виконала) або важкою (ніхто не виконав):

- Виражені чисельно складності сукупності ТЗ в тесті повинні рівномірно заповнювати той інтервал, який відповідає навченості претендентів;
- Трудомісткість ТЗ повинна бути такою, щоб у тимчасові рамки тесту могли вклатися всі претенденти, здатні їх виконати.

Від збалансованості ТЗ за складністю і трудомісткістю залежить здатність тесту диференціювати претендентів у відповідності з їх властивостями. Ця здатність тесту, виражена числом, називається роздільною здатністю.

18. Дидактична спрямованість тесту і технології тестування. Одних претендентів важкі ТЗ можуть підштовхнути до навчання, інших - відштовхнути від нього. Нудні, одноманітні, зайво-формальні сукупності ТЗ шкідливі у навчанні. ТЗ в своїй сукупності повинні включати елементи:

- привабливості (гумористичне забарвлення умов тощо);
- сюрпризу (різні і несподівані словникові звороти, рідко зустрічаються в підручнику, звернення в умові завдань безпосередньо до виконавця з використанням даних реєстрації тощо);

- Практичної спрямованості (приклади з життя, можливість застосувати результати виконаного завдання безпосередньо в діяльності тощо)

Сюрпризні завдання включаються в тест для протидії виникненню в процесі роботи претендентів ефекту «рівної дороги». Трудомісткість і складність ТЗ, у сукупності, повинна бути такою, що хоча б пару-трійку завдань повинні правильно виконати абсолютно всі претенденти. Це змусить слабких претендентів повірити у свої сили, захопить, зніщить змагання, дасть підстави для похвали.

Тест (технологія тестування), відповідно до принципів дидактики, повинен (повинна) не тільки диференціювати і вимірювати властивості претендентів, але і володіти властивістю ініціювання їх до самонавчання і виявляти їх прагнення до підвищення ефективності якості знань, умінь, навичок.

19. Оптимальність тесту. У тест, як правило, включаються тільки ті ТЗ, зміст яких не дублюється і рівномірно покриває, у своїй сукупності, навчальний матеріал. Частина ТЗ виключають з тесту, коли відомо, що:

- претенденти не знають відповідної частини матеріалу або, навпаки, знають її гарантовано добре;
- претенденти не мають навичок роботи і тому не впораються з трудомісткими завданнями.

Тест вважають більш оптимальним, якщо з його допомогою можна виміряти такі ж характеристики і провести диференціювання претендентів за менший час.

20. Темперування складності. Послідовність завдань тесту така, що кожне подальше ТЗ, на думку укладача, важче попереднього. Цей принцип, який застосовується в бланковому тестуванні, нерідко піддається критиці, оскільки діє лише «в середньому». Якщо претенденти вчилися у різних викладачів і по різних підручниках, передбачити ряд зростаючих труднощів для більшості з них - велика проблема. А «усереднений» ряд не приносить бажаного ефекту підвищення роздільної здатності. Тому, в комп'ютерному тестуванні частіше

використовують інший принцип: претендентам дозволяють самостійно вибирати порядок виконання завдань. (Див. «Свобода вибору ТЗ»)

21. Темперування трудомісткості. Послідовність завдань тесту така, що кожне подальше ТЗ, на думку укладача, більш трудомістке, ніж попереднє. Застосовується, коли оптимізують тести для вимірювання працездатності, уважності та подібних характеристик претендентів.

22. Свобода вибору. Принцип складання тесту, що прийшов на зміну принципу «зростання складності». Принцип свободи вибору в технологіях комп'ютерного тестування проявляється в тому, що претендентів не обмежують у виборі послідовності виконання завдань. Практика показує, що при цьому претенденти самі шукають та інтуїтивно знаходять для себе індивідуальний ряд ТЗ зростаючої складності та трудомісткості.

23. Варіативність тесту. Кожне звернення до БТЗ призводить до генерації нового варіанту тесту відповідно до алгоритму, ключ якого може формуватися в відповідності з реєстраційними даними, датою і часом сеансу. Алгоритм, який орієнтується при складанні варіанта тесту на особистість претендента, може виключати можливість повторення ТЗ при повторному тестуванні.

24. Адаптивність тесту. Зміст тесту адаптують для надання тесту більшої роздільної здатності у відомій групі претендентів. Для адаптації тесту до властивостей групи претендентів і кожного претендента окремо, необхідна наявність в БТЗ такого набору класифікованих завдань, щоб можливою була побудова варіантів тесту націлених на вимірювання конкретних властивостей:

- тест розділяють на субтести, а сеанс - на ступені;
- перед побудовою субтеста чергового ступеня проводиться аналіз результатів виконання попередніх субтестів;
- при аналізі конкретизуються результати проведених вимірювань властивостей претендентів, потім, якщо об'єктивна диференціація групи претендентів ще не можлива, складається черговий субтест;
- дії повторюються на черговому ступені сеансу до тих пір, поки не буде досягнутий необхідний ступінь диференціації.

25. Повнота. У банку ТЗ повинні бути відображені абсолютно всі структурні елементи змісту дисципліни і зв'язку між ними.

26. Доцільність. Зміст тесту залежить від мети тестування. Наприклад, можлива ситуація:

- Якщо потрібно відібрати найбільш підготовлених учнів на олімпіаду, то завдання повинні бути складні, бо лише з допомогою таких завдань можна відібрати кращих;

- Якщо претенденти навчалися за різними підручниками, то формулювання відібраних для тесту ТЗ повинні відповідати кожному з них.

27. Верифіковані результати. Тест може містити рівномірно розподілені серед всіх завдань такі ТЗ, які характерні різній формі і рівні за змістом. Встановивши кореляцію результатів їх виконання можна встановити, як претендент виконував завдання - вдумливо або методом «випадкового тикуну».

28. Збалансованість. Кількість розділів БТЗ і включених до них завдань повинні корелювати зі змістом, відносними обсягами інформації розділів і формою дисципліни.

29. Стилiстична несуперечність. ТЗ формулюються у звичній для претендентів формі, з використанням словникових зворотів з рекомендованих навчальних посібників. Винятки - сюрпризні завдання, що включаються в тест для протидії виникненню в поведінці претендентів ефекту «рівної дороги».

30. Логічна і семантична несуперечність. Серед логічних вимог - визначеність предмета вимірювання. Близькість дисциплін ускладнює визначення предметної належності ТЗ. Чим менше перетинань дисциплін та їх розділів, тим чистіше, чіткіше виявляється у тесті їх зміст. Логічно ТЗ можна асоціювати з твердженням. У разі правильної підстановки (Форма ТЗ № 2) твердження перетворюється на правильне, у разі неправильної - в помилкове. Звичні для педагога питання не є твердженнями, а тому зміст питання не може визначатися як істинне або помилкове. Питання застосовуються в ТЗ закритої форми. Слід уникати невідповідностей роду, числа, відмінювання в

формулюваннях завдань і варіантів їх виконання. У ТЗ немає місця двозначностям, наприклад:

- на запитання: «Де знаходиться Київ?», - можна відповісти: «На березі Дніпра», «В Європі » тощо;
- на питання: «Хто першим полетів у космос?», - можна відповісти: «Радянська людина», «Юрій Гагарін», «Лайка» тощо;
- на запитання: «Коли почалася Велика вітчизняна війна?», - можна відповісти: «В червні», «В ХХ столітті » тощо;
- у завданнях відкритої форми слід вказувати рід, відмінок тощо очікуваних слів.

Наприклад, на запитання: «Як називається судно, яке розколює лід?», - можна відповісти і «криголам», і «криголамом»;

- серед завдань на встановлення відповідності або на встановлення послідовності зустрічаються варіанти, які трактуються в різних джерелах по-різному (причини і послідовності історичних подій).

31. Дистрактність. Варіанти виконання завдань (Форма ТЗ № 1) формулюються так, щоб правильні варіанти були схожі на неправильні, і навпаки. Формулюванні правильних варіантів повинні точно відповідати матеріалу навчального курсу. Слід уникати характерної помилки початківців авторів, коли формулювання правильних варіантів більш довше від формулювань невірних варіантів. Дотримання принципу дистрактності значно підвищує валідність тесту.

32. Надмірність. Часто трапляється так, що до групи претендентів потрапляють учні, що навчалися за різними програмами та підручниками. Тест складається один для всіх і по ньому, щоб «зрівняти шанси», робляться наступні кроки:

- в тест включаються ТЗ, формулювання яких викладено в різних підручниках приблизно в рівних пропорціях;

- еквівалентні за інформацією завдання групуються, і учасників інформують про те, що необхідно і достатньо правильно виконати лише одне завдання групи;

- кількість ТЗ в тесті перевищує норму, яку міг би виконати усереднений учасник сеансу.

Іноді завищують кількість завдань тесту вище норми і з інших міркувань:

- щоб збільшити шанс слабким учням хоч в чомусь проявити себе і отримати похвалу;

- щоб виключити при проведенні освітньої олімпіади накладки пов'язаної з тим, що кілька учасників виконують правильно всі запропоновані ТЗ.

33. Ділова гра. Технології тестування конструюються виходячи з рекомендацій дидактики, яка каже, що найкращих успіхів можна досягти при використанні ігрового навчання. Ігрова технологія дозволяє зробити процес тестування привабливим і досягти високих успіхів за рахунок:

- природної індивідуалізації;
- виникнення змагання;
- колективності;
- публічності.

Нижче опишемо зміст і практику застосування перерахованих принципів.

14.3 Педагогічний тест

Ключовим поняттям тестології є поняття «педагогічний тест».

Педагогічний тест – це:

- система завдань, виконана групою суб'єктів тестування, результат яких можливо надійно ранжувати (надати їм порядкові номери) за якістю навчання, кількістю наявних знань;

- система стандартизованих завдань, результат виконання яких дозволяє із заданим ступенем точності виміряти знання, навички та вміння випробуваного.

Іноді під поняттям «тест» розуміють засіб педагогічного вимірювання або стандартизований метод дослідження, призначений для кількісних і якісних оцінок характеристик людини. Плутанина понять (метод - засіб - сукупність -

система) відбувається від читання перекладної літератури, а також від існування безлічі скорочень, різночитань і жаргонних виразів вжитих в спеціальній літературі з психології, комп'ютерної техніки, педагогіки.

Форма наведеного визначення педагогічного тесту передбачає деяку невизначеність, неможливість точного вираження, пошук шляху побудови ідеалу. Тестологи кажуть, що ідеального педагогічного тесту ще не створено. У цьому змісті і слід розуміти багато визначень з тестології.

Чому тестологи не довели визначення до чисел? Чому визначення тестологів, в один і той же час, такі схожі на визначення з новітніх математичних напрямів - теорії нечітких множин тощо? Напевно тому, що місце тестології в загальній картині наук «лежить» якраз на стику педагогіки і математики. Тестологія ще тільки стає на ноги. Тому зайво математизована педагогічна мова тут відлякує педагогів і смішить математиків.

Педагогічним тестам, а також результатам їх застосування (результатами тестування), приписують характеристики та властивості. Найважливіші властивості тесту - валідність і надійність.

Надійність тесту - ступінь стійкості (незмінності) результатів при повторному тестуванні тієї ж групи претендентів (тих же суб'єктів тестування).

Валідність тесту - ступінь придатності тесту для вимірювання саме тих якостей претендентів (суб'єктів тестування), які він за задумом повинен вимірювати.

Валідності та надійності зазвичай не надають точних кількісних значень, використовуючи при їх порівнянні, для різних тестів, лише відношення порядку:

- валідність педагогічного тесту тим вища, чим більші результати його застосування відповідають вирішенню проблеми ранжування претендентів за якістю навчання;
- надійність педагогічного тесту тим вища, чим більш стійкі (однакові) результати його застосування при тестуванні однорідних груп претендентів.

Фахівцями найбільш часто обговорюються наступні характеристики тестів:

- час вимірювання, тобто скільки часу необхідно тестувати групу претендентів для того, щоб результати тестування стали досить валідними і надійними;
- привабливість, тобто наскільки охоче претенденти обирають для свого тестування той чи інший тест;
- простота створення (дуже важлива характеристика для умов поточного виробництва і для забезпечення економічності виробництва);
- простота застосування (іншими словами, наскільки великі вимоги до кваліфікації фахівців, що його застосовують або обслуговуючих, а також до техніки і обладнання).

Цим характеристикам так само, як і наведеним вище властивостям, важко скласти точні кількісні критерії. Проте їх можна порівняти, ґрунтуючись на думках експертів, і, відповідно, ранжувати.

Диференціююча здатність тесту оцінюється з точки зору відповідності рівня його складності, до рівня підготовленості претендентів. Марно давати слабким учням важкі завдання і знаючим студентам легкі. В обох випадках претенденти не будуть відрізнятися результатами. Вимірювання не відбудеться через невідповідність рівня складності тесту рівнем підготовленості. Ефективний тест відповідає по складності завдань рівню підготовленості випробуваних. Якщо включити в тест завдання для вимірювання на всьому діапазоні зміни складності, то знижується точність виміру на окремій ділянці. І навпаки, для точного вимірювання знань претендентів однакового рівня підготовленості, потрібно мати більше завдань відповідного рівня складності. Ефективний тест (з мінімальною кількістю завдань) не може бути ефективним на всьому діапазоні підготовленості претендентів.

Релевантність тестового комплексу. Релевантність тестового комплексу рідко обговорюється. Це дуже важлива характеристика тестових комплексів, (а також технологій тестування та навчальних посібників), ступінь оптимальності

відповідно до навчальних цілей. Таку характеристику можна назвати релевантністю. Поняття «релевантність» за своїм змістом близьке поняттю «валідність», але охоплює більше смислове коло, оскільки ранжування претендента є лише малою (безумовно, важливою) частиною навчального процесу.

Педагогічний тестовий комплекс створюється з урахуванням єдності семи систем:

- системи знань перевірки навчальної дисципліни;
- системи завдань;
- системи подання завдань піддослідним (алгоритм, інтерфейс);
- сценарію тестування;
- системи методичного забезпечення;
- системи технічного забезпечення;
- системи статистичної обробки результатів піддослідних.

Інформація, яку містить педагогічний тестовий комплекс, змінюється з часом, включаються нові завдання, змінюються статистичні ваги завдань, змінюються сценарії тощо.

14.4 Класифікації та види комп'ютерних педагогічних тестів

Педагогічний тест можна визначити не тільки щодо групи випробовуваних, як це зроблено вище, але й щодо одного випробуваного, взятого окремо.

Педагогічні тести класифікуються відповідно до виділених груп.

ПІДСТАВИ ДЛЯ КЛАСИФІКАЦІЇ ТЕСТІВ № 1 (за тематикою завдань)

Тест гомогенний - сукупність стандартизованих завдань з однієї навчальної дисципліни (розділу дисципліни). Система завдань зростаючої складності для об'єктивної та ефективної оцінки рівня підготовленості учнів по одній з навчальних дисциплін. Гомогенні тести поширені більше інших. Вони створюються для контролю знань з однієї навчальної дисципліни або по одному розділу з об'ємного навчального курсу.

Тест гетерогенний - сукупність стандартизованих завдань, створених з метою виміру знань з кількох навчальних дисциплін (розділів дисциплін, але не вимагає одноразового залучення знань з різних дисциплін (розділів дисциплін). Гетерогенні тести використовуються для комплексної оцінки випускників шкіл, для відбору абітурієнтів при прийомі до ВНЗ. Гетерогенний тест складається з гомогенних. Обчислення результату ведеться для кожного гомогенного тесту окремо з наступною інтеграцією в єдиний рейтинг.

Тест інтегративний - система завдань для загальної діагностики підготовленості випускника освітньої установи. Завдання такі, що для їх виконання потрібні синтетичні знання з кількох навчальних дисциплін.

Тест адаптивний - система завдань з відомими властивостями і параметрами. Завдання пред'являються, за допомогою комп'ютера, в порядку, залежному від відповіді випробуваного на попереднє завдання. За допомогою різних стратегій пред'явлення завдань вдається значно скоротити час комп'ютерних тестувань без втрати точності вимірювань.

ПІДСТАВИ ДЛЯ КЛАСИФІКАЦІЇ ТЕСТІВ № 2 (щодо процедури формування черги подання)

Тест змішаний - сукупність стандартизованих завдань різної складності, подається у випадковому порядку.

Тест зростаючої складності - сукупність стандартизованих завдань, застосовується в порядку зростання складності.

Тест адаптивний - система стандартизованих завдань з відомими (заданими) параметрами складності. Такий тест диференціює здібності. Черговість пред'явлення кожного завдання залежить від успішності виконання випробуваним попереднього завдання. При правильному виконанні попереднього завдання - ступінь складності наступного зростає, і навпаки.

ПІДСТАВИ ДЛЯ КЛАСИФІКАЦІЇ ТЕСТІВ № 3 (за кількістю претендентів)

Тест індивідуальний - тест, створений для окремого випробуваного або окремого типу піддослідних.

Тест груповий - тест, який група піддослідних виконує спільно (колективно).

ПІДСТАВИ ДЛЯ КЛАСИФІКАЦІЇ ТЕСТІВ № 4 (за формою подання завдань)

Тест стандартний - тест для застосування найбільш поширеної методики діагностики випробовуваних, згідно з якою вони виконують одну або схожі сукупності завдань (варіанти гомогенного тесту) за однаковий обмежений період часу і в однакових умовах. При цьому зміст завдань випробуванним заздалегідь не оголошується, а умови проведення тестування виключають можливість спілкування з метою підказки. Перевірка результатів проводиться вручну або із залученням засобів автоматизації. Результатом діагностики (вимірювання) для кожного випробуваного є число - рейтинг, що дозволяє впорядкувати піддослідних у рамках групи відповідно до виявлених рівнів знань (умінь, властивостей, характеристик).

Тест прихований - тест створений для виявлення неактуалізованих здібностей і оцінювання, наскільки успішно випробуваний зміг би освоювати знання, вміння тощо. При прихованому динамічному тестуванні викладач активно взаємодіє з випробуванним, консультує його, видає конкретні рекомендації щодо оптимізації дій.

Тест ігровий - тест призначений для виявлення ЗУН випробуваних у процесі комп'ютерної гри. Форма ділової гри застосовується на стандартних навчальних заняттях у різних видах.

Тест мультимедійний - тест, в якому інформація різної природи присутня рівноправно і взаємопов'язана - тексти, зображення, звуки, анімація, фрагменти фільмів. Як приклад тестуючих і навчальних комп'ютерних ігор можна назвати тренажери для осіб військових професій - льотчиків, танкістів.

ПІДСТАВИ ДЛЯ КЛАСИФІКАЦІЇ ТЕСТІВ № 5 (по процедурі аналізу результатів виконання)

Тест бланковий - тест, що проводиться з використанням бланків - стандартизованих паперових, картонних, пластмасових листів, на яких

нанесена тестова інформація і присутні місця для занесення інформації випробуваного. Інформація випробуваного, що символізує результати виконання ним тестових завдань, може наноситися на бланк шляхом перфорації отворів, нанесення знаків друкарською ручкою, вирізування фрагментів ножицями. Інформація аналізується шляхом механічного порівнювання зміни матеріалу бланка зі зразком. Механічне порівняння може здійснюватися шляхом протягування спиць в технологічні отвори бланку тощо.

Тест машинний - тест, результати якого обробляються автоматично за допомогою комп'ютерних програм і спеціальних технічних пристроїв

ПІДСТАВИ ДЛЯ КЛАСИФІКАЦІЇ ТЕСТІВ № 6 (за процедурою інформаційного захисту)

Тест безваріативний - тест, що складається з фіксованої кількості тестових завдань.

Всі претенденти отримують для виконання єдиний тест.

Тест багатоваріативний - сукупність безваріативних тестів (варіантів). Завдання з однаковими номерами близькі між собою у всіх варіантах. Варіантів складається так багато, щоб всі сусіди кожного учасника тестувань мали різні варіанти.

Тест з випадковою вибіркою - тест, що формується з бази тестових завдань безпосередньо перед початком тестування. Це робиться за допомогою якої-небудь процедури, яка генерує випадковий список номерів завдань.

ПІДСТАВИ ДЛЯ КЛАСИФІКАЦІЇ ТЕСТІВ № 7 (за метою)

Тест досягнень - тест контролю успішності навчальної діяльності випробуваного. Тести досягнень поділяються за цілям і завданням виміру:

- загальної результативності;
- ступеня засвоєння знань;
- загального розвитку;
- інтелектуальних здібностей;
- функціональних можливостей;
- професійної придатності.

Тест засвоєння знань - тест для визначення рівня засвоєння навчального матеріалу суб'єктів тестування з конкретної навчальної дисципліни, теми. Тестування проводиться або відразу після вивчення, або через великий проміжок часу (тест залишкових знань).

Тест залишкових знань - тест для визначення рівня засвоєння навчального матеріалу учнями з конкретної навчальної дисципліни, теми. Тестування проводиться відразу після вивчення або через великий проміжок часу (тест залишкових знань).

14.5 Недоліки педагогічних тестів

Тест - це сукупність або система кількох ТЗ. Характерні недоліки тестових завдань ми розглянемо нижче, а тут зосередимося на недоліках саме тестів. Перелічимо основні недоліки:

- нерівномірність «покриття» навчального матеріалу, відсутність рівності відображення подібних за інформаційною вагою і важливістю розділів дисципліни;

- необ'єктивність ваг ТЗ, тобто кількостей балів, які претенденти отримують в свою «скарбничку» за виконання кожного завдання; вага не відображає середньостатистичну складність і важливість завдання;

- мала кількість ТЗ в тесті або одноваріантному тесті, які передбачають простоту користування шпаргалкою і тестування не знань, а здібностей запам'ятовувати послідовності букв і цифр. Завдання для складання варіанта тесту повинні вибиратися з БТЗ випадковим чином, а БТЗ повинна бути настільки велика, щоб заходи щодо забезпечення її таємності були не потрібні (більше 300 завдань для цього достатньо);

- відсутність орієнтації завдань на врахування притаманних людині властивостей і якостей (настрій, втома, темперамент, вік, стать, національність), тобто відсутність у тесті елементів гуманістичності або толерантності;

- якщо в тестуючу програму не закладені методики випадкового вибору порядку пред'явлення ТЗ і тест створювався однією людиною, то номери правильних варіантів іноді можна досить надійно визначити за допомогою

калькулятора і простих алгоритмів, створених кібернетиками ще в середині минулого століття. Людині тільки здається, що вона розподілила варіанти у випадковому порядку;

- присутність зв'язку між послідовними завданнями, коли для вірного виконання одного завдання учень користується або змушений користуватися інформацією з інших завдань. Такий, закладений автором, зв'язок легко простежується в бланковому варіанті тесту, але зникає в комп'ютерному;

- питання та відповіді одноманітні, нудні. Це негативно впливає на випробуваних, викликаючи так званий ефект «рівної дороги» - відомий в практиці водіїв-далекобійників ефект засинання за кермом, відволікання і розсіювання уваги;

- інструкцію для учасників дали лише в одному ТЗ і припустили, що учасник виконує їх послідовно. Це призводить до плутанини та помилок при відхиленні;

- Відсутність елементів адаптивності;
- Відсутність елементів психофізичної розвантаження.

Часто ці недоліки взаємопов'язані, це легко побачити при деякому досвіді складання завдань. Є й інші недоліки тестів, які виявляються вже в процесі математичної обробки результатів тестування. Недоліки, пов'язані з валідністю і надійністю тесту, обговорені в спеціальному розділі цього посібника.

14.6 Валідність і надійність тесту

Основними властивостями і параметрами якості тестів вважаються валідність та надійність. Ці властивості, на відміну від багатьох інших, що обговорюються в педагогіці, виражаються числами.

Валідність відображає придатність тесту для вимірювання того, що він за задумом повинен вимірювати. Вимірювання валідності тесту може включати ряд процедур, головною з яких є валідація.

Валідація проводиться шляхом математичного порівняння результатів тестування з успішністю виконання відповідної практичної діяльності випробуваними. При цьому враховується шкальована думка керівників групи

суб'єктів тестування, що пройшли тестування, про їх конкретні виконавські якості.

Валідність вважається досить високою, якщо коефіцієнт кореляції буде більшим 0,6. При значенні коефіцієнта кореляції 0,45 - 0,65 валідність вважається цілком задовільною.

Вважається, що стовідсотково валідних тестів немає. Розрізняють такі види валідності:

- змістовну - ступінь відповідності тесту програмам навчання та освітнім стандартам;
- критеріальну - ступінь відповідності результатів тестування до зовнішнього, що не відноситься до тесту критерію;
- кваліметричну - ступінь зв'язку результатів математичної обробки результатів тестування та їх інтерпретації;
- прогностичну - ступінь повноти досягнення мети тестування; та ін.

Найбільшої «ваги» надають змістовій валідності.

14.7 Надійність тесту. Контрольовані та неконтрольовані фактори впливу

Надійність тесту визначається як стійкість результатів при повторному тестуванні на тій же вибірці суб'єктів тестування. Вважається, що при коефіцієнті кореляції результатів повторного тестування більше 0,75 рівень надійності тесту прийнятний.

Надійність тесту перевіряється щодо

- тимчасових змін;
- вибору конкретних завдань;
- конкретних індикаторів;
- ролі індивідуальності персоналу при накопиченні та обробці даних;
- аспектів процедури тестування.

Обчислення помилки вимірювань - ймовірних меж коливань вимірюваної величини, яке засноване на понятті надійності.

Жоден тест не є абсолютно надійним. Дисперсія помилки (міра відхилення результатів сеансів тестування) відображає випадкові коливання, викликані неконтрольованими факторами:

- тренінгом і донавчанням претендентів протягом періоду виміру надійності;
- змінами погоди;
- випадковими відволікаючими моментами;
- здатністю до навчання претендентів протягом сеансу і в перервах між сеансами;
- вмотивованістю претендентів і її змінами;
- адаптацією, звиканням претендентів до форми проведення тестування;
- змінами стану претендентів.

Перший з цих факторів, найбільш впливає на результати і «плутає всі карти», - звичайне явище, створюване батьками, керівництвом навчальних закладів, добросовісними педагогами, що не бажають, щоб їхні вихованці постали перед експериментаторами в поганому вигляді.

Стандартизація та комп'ютеризація тестів зрівнюють умови тестування і сприяють підвищенню надійності.

Стандартний набір даних про тест, що призначений для широкого вжитку, обов'язково повинен включати відомості про міру його надійності.

Надійність звичайно обчислюється за допомогою коефіцієнта кореляції похідної моментів К. Пірсона (його можна знайти в статистичних довідниках).

При вимірі надійності слід звертати увагу на критерій доцільності, пам'ятати про ставлення «витрати - користь». Слід контролювати, в якому інтервалі часу вимірювалася надійність, чи відбулися за цей період будь-які події, здатні вплинути на результати.

Існують різні типи надійності та підходи до їх обчислення. Педагогу, що апробує новий тест на своїх класах, важливо володіти технікою вимірювання:

- ретестової надійності;
- надійності еквівалентних форм.

Надійність не обов'язково передбачає високу валідність. Серед тестологів існує думка, що на практиці завжди виконується такий вираз: *валідність < надійність*. Іншими словами, значення валідності тесту не може перевищувати значення його надійності, які б процедури їх визначення не використовувалися. Важко у це повірити, поки не проведено математичне порівняння процедур обчислення цих значень.

14.8 Типи надійності тесту та способи її визначення

Надійність і валідність тестів можуть бути значно підвищені, якщо їх якісні та змістовні характеристики будуть пов'язані зі статистичними даними, отриманими при обробці великих масивів результатів тестування суб'єктів тестування.

Існує кілька способів визначення і підвищення надійності тесту.

- Метод повторного тестування (метод ретест). Дворазове або багаторазове використання одного і того ж тесту в одній групі піддослідних. Перевага методу полягає в простоті його використання, ясності основних посилок, простоті порівнянь і розрахунків. До недоліків відносять невизначеність вибору тимчасового інтервалу між опитуваннями. Цей інтервал може коливатися від декількох хвилин до декількох днів, місяців і навіть років. Природно, що при цьому по-різному виявляються чинники: запам'ятовування або, навпаки, забування, вплив досвіду, отриманого в першому опитуванні на другий, вплив спілкування випробовуваних між собою після першого опитування.

- Метод паралельного тестування (метод еквівалентних форм). Одній і тій ж групі випробовуваних дається спочатку одна форма тесту, а після перерви - інша. Потім обчислюється величина коефіцієнта кореляції правильності виконання завдань, яка і приймається за значення коефіцієнта надійності. Якщо між поданням обох форм є значний часовий інтервал, то коефіцієнт надійності називають по-іншому: коефіцієнтом еквівалентності або коефіцієнтом стабільності.

- Метод роздільної корельованості. В основі лежить припущення про паралельність не тільки окремих форм, а й частин всередині однієї форми тесту.

Для отримання величини коефіцієнта надійності порівнюються результати виконання частин тесту. У залежності від способу розподілу тесту можуть змінюватися значення коефіцієнта. Часто застосовується процедура поділу тесту - це зведення в одну частину результатів респондентів у непарних номерах висловлювань, а в іншу - парних. Підсумовування балів у цих половинах тесту дає два вектори, коефіцієнт кореляції між якими і служить коефіцієнтом надійності тесту. Його називають ще коефіцієнтом внутрішньої спроможності тесту. Цей метод має перевагу перед іншими, оскільки дозволяє оцінити надійність при одноразовому тестуванні.

Найчастіше розглядають три типи надійності:

- ретестова надійність. Вимірюється при повторному проведенні того ж самого тесту, на тому ж контингенті, та в тих же умовах. Її аналогом є перевірна контрольна робота, що проводиться, коли результати першої контрольної здаються сумнівними. Ретестова надійність зазвичай підраховується з використанням коефіцієнта кореляції моментів Пірсона. Щоб підвищити ретестову надійність, треба відібрати з початкового апробуючого варіанту тесту ті завдання, на які випробувані дають стійкі відповіді. Чим вища ретестова надійність, тим менш чутливі результати до впливу неконтрольованих факторів.

- Надійність еквівалентних форм. Вимірювання надійності еквівалентних форм простіше за вимірювання ретестової надійності. Це обчислення кореляції результатів виконання двох форм одного і того ж тесту, що вважаються рівноцінними. Їх називають також паралельними, взаємозамінними, порівнянними, подібними. Тут коефіцієнт надійності одночасно відображає тимчасову стабільність тесту та узгодженість результатів сеансів по двох формах. Тільки якщо варіанти сеансів йдуть один за одним, можна точно виміряти надійність еквівалентних форм тесту. Для зниження впливу ефекту самонавчання претендентів, змінюють спосіб подання еквівалентних форм подібних завдань. На еквівалентність також перевіряють:

- інструкції;

- персонал;
- тимчасові рамки роботи;
- формат бланків.

Надійність тесту на швидкість. Для перевірки надійності тестів на швидкість вважаються найбільш ефективними метод ретест і метод еквівалентних форм. Є прийоми поділу повного часу виконання тесту на чотири частини з реєстрацією результатів окремо для кожної чверті.

14.9 Розвиток поняття валідності та надійності тесту

Властивості валідності та надійності зазвичай приписують тесту, набору тестових завдань. Це правильно, поки використовується одна-єдина процедура тестування. Але таких процедур вже багато: є адаптивні процедури, є тестуючі ділові ігри. Вони відрізняються від стандартів п'ятирічної давності та безперервно розвиваються. Цей напрям діяльності розвивається і не можна забувати про це, будуючи понятійний апарат тестології.

Сама процедура тестування, тобто спосіб представлення та організації тесту, сильно та по-різному впливає на ступінь вирішення освітніх завдань, на вимірювані величини валідності та надійності тесту. Тому доцільно приписувати властивості валідності та надійності не тесту, як такому, а всій використовуваній тестовій технології, що включає в себе і тест, і інтерфейс програмного комплексу, і процедуру публікації результатів тощо. У цьому випадку результати вимірювання валідності та надійності стануть більш об'єктивними.

Крім цього, у зв'язку з розвитком технологій, слід ввести в розгляд нові властивості, що характеризують ступінь відповідності освітнім завданням не тесту, як такого, а всієї технології тестування.

Часто вважають, що тестування призначене для виконання єдино значущої функції: вимірювання рівня (кількості, якості) ЗУН учня (групи учнів). У зв'язку з цим, найважливішою характеристикою процедури тестування вважають надійність. Іншими словами, вважається, що тестовий комплекс

хороший тоді, коли результат тестування при повторному проходженні не змінюється.

Вимога відтворюваності (розуміється як синонім надійності) обумовлена не освітніми потребами або нормами, а специфікою мислення (освіти) творців тестуючих комплексів - технарів, як часто підсвідомо ототожнюють людину з будь-яким пристроєм.

Насправді ж освітні потреби та дидактичні принципи спрямовані саме на те, щоб результати тестувань не повторювалися, а підвищувалися. З цієї точки зору, процедура тестування і тесту є комплексом хорошим лише тоді, коли вона (процедура) виконує головну освітню задачу: спонукає учня до самостійного підвищення рівня якості ЗУН і відповідної самоперевірки. Рецепт відомий: процедура тестування повинна включати не тільки механізм вимірювання, але й елементи змагальності та привабливості (ділової гри, сюрпризу тощо). А кожне проходження цієї процедури має спонукати в учня прагнення до самовдосконалення, тобто до самоосвіти, і приводити до підвищення результату вимірювання. Міра цього зростання - коефіцієнт релевантності (відповідності освітнім цілям і принципам) - одна з найважливіших характеристик процедури.

Ця точка зору викликає чимало суперечок, у тому числі і в середовищі професійних педагогів. Іноді вважають, що в такій серйозній справі, як освітній процес, немає місця грі. Є автори, які серйозно стверджують, що в освіті немає місця змагання, людина повинна самовдосконалюватися сама, не озиратися на колег. Нібито, це - вищий ступінь розвитку учня. Існування таких суперечок дивне - адже класики педагогіки і психології в один голос висловлюються за важливість і необхідність застосування саме ігрових елементів у навчанні (див. напр.: Хуторський В. А. Сучасна дидактика, Спб: Питер, 2001.), А значить , і сюрпризу і змагальності.

Проведені в Росії експерименти серед школярів (2-11 класи), абітурієнтів, студентів (1-5 курс), вчителів, викладачів вузів, показали наступне:

- Використання в процедурі тестування елементів гри, таких як напівінтерактивний діалог учня з комп'ютерними персонажами, замість сухої демонстрації табличок з питаннями і списками відповідей, спонукає бажання учнів до неодноразового тестування;

- Включення в процедуру тестування елементів змагальності (наприклад, публікація результатів, символічний приз) з паралельним дозволом неодноразового повторення спроб призводить до зростання популярності тестування та зростання рівня результатів учасників;

- Вчителі та викладачі ставляться до тестування своїх ЗУН, як правило, насторожено. Вони побоюються процедури реєстрації, вони побоюються торкатися до клавіатури і миші. Розвіяння (чудове перетворення) наступає, тоді коли виявляється, що «тестуються знання не самого учасника, а керованого ним комп'ютерного персонажа». При такій постановці педагоги годинами готові сидіти за комп'ютером, обговорюючи з колегами нюанси, що виникають на екрані, прагнучи «заробити» якомога більше балів;

- Допущення сценарієм гри комп'ютерної «шпаргалки» призводить до інтенсифікації запам'ятовування навчальної інформації. Інформація, що отримується «обхідним» шляхом, виявляється, краще запам'ятовується. Учні (і педагоги) із задоволенням діляться один з одним почерпнутими знаннями, допомагають один одному їх знайти, перестають бути «зжатими» при колективному спілкуванні на навчальні теми.

Повторимо: описані явища відбуваються як серед дітей, так і серед дорослих. Більш того, в ігровому тестуванні з цікавістю беруть участь навіть присутні батьки та шкільні «двієчники». Вони з інтересом обговорюють з однокласниками і вчителем перипетії своїх комп'ютерних діалогів в рамках навчальної дисципліни, звертаються до підручника.

Останнє з перерахованих вище спостережень, ймовірно, може призвести в майбутньому до незвичайної технології освітнього процесу, коли учні самостійно, групою і поодиночці, «видобувають» навчальну та додаткову інформацію, що стає колективним надбанням.

Проведені дослідження спроб тестувань школярів показали, що, у разі зняття кількісних обмежень на набір тестових завдань, «крива зростання» результатів учасників в середньому близька експоненті. При цьому обчислюваний експоненціальний параметр (коефіцієнт) стійкий і, мабуть, може використовуватися в якості характеристик як учасника, так і колективу учасників тестувань. На форму кривої «зростання» повинні впливати такі характеристики учня як швидкість адаптації до ситуації, здатність швидко засвоювати навчальний матеріал (вчитися), психічний стан під час сеансу тестування, обізнаність. Важливим показником, який можна зробити з результатів повторних тестувань, разом з сумарною кількістю набраних учасником балів, є різниця кількостей балів, набраних при повторних тестуваннях. У залежності від часу, який пройшов між спробами, цю різницю можна інтерпретувати і як показник звикання, і як показник навченості (в перерві учасник міг за своєю ініціативою спілкуватися з колегами, читати підручник тощо). Відхилення цього показника в позитивну (негативну) сторону цілком можуть характеризувати учня з позитивної (негативної) сторони.

На підставі сказаного, доцільно обговорити введення в науковий обіг такої чисельної характеристики процедури тестування (ми її тут називаємо «релевантність»), яка прямо пов'язана з валідністю, але характеризує не ступінь стійкості результату вимірювання, а ступінь його зростання при повторних вимірах .

Уточнимо спеціально для фізиків: описана характеристика має цікавий аналог в статистичній фізиці - виробництво ентропії.

14.10 Інформаційний граф

Технологія створення електронного навчального посібника і педагогічного тесту - це оснований на досягненнях тестології складний і трудомісткий процес. Додаткові складнощі тут створюють своєрідні педагогічні традиції. Цитуємо:

Методологической основой разработки и использования тестовых средств В. С. Идиатулин считает «представления когнитивной психологии о пропозициональной репрезентации усвоенной информации в человеческой

памяти, о существовании ее наименьших значимых единиц, которые могут быть выражены отдельным утверждением, имеющим самостоятельное смысловое содержание, описывающее объекты и отношения. При этом выделение достаточно однородных элементов знания становится предпосылкой измерения его совокупности». (Кречетников К. Г. Задания в тестовой форме и методика их разработки: Учебно-методическое пособие – Владивосток: Изд-во Дальневост. ун-та, 2002.- 36с).

На жаль, багато педагогів люблять викладати свої думки саме так.

Досвідчені укладачі спочатку завжди «конвертують» матеріал дисципліни, тобто переводять його в послідовність коротких тверджень - текстових блоків. При цьому «виливається водичка» і в якості блоків залишаються:

- положення та поняття;
- визначення та формулювання;
- факти;
- класифікації;
- закони та закономірності;
- висновки;
- тезаурус;
- важливі ілюстративні матеріали.

Інформація блоків повинна бути достатньою для контролю ЗУН. Поруч «на полях» мають бути відмічені внутрішні і зовнішні зв'язки блоків між собою і з суміжними дисциплінами. Блоки пов'язані між собою в загальній структурі матеріалу дисципліни, включені один в одний та ієрархічно підпорядковані. Відносини блоків можуть бути представлені у вигляді орієнтованого графа. У вершинах графа розташовуються блоки, а їх взаємозв'язки позначаються спрямованими ребрами. Такий граф, оснащений деякими додатковими атрибутами, іноді називають моделлю знань. У такому вигляді простіше:

- шукати і встановлювати нові зв'язки;
- виробляти впровадження нового матеріалу;
- знаходити і усувати дублювання матеріалу;

- виробляти розподіл інформації на великі модулі - підграфи;
- проектувати інтелектуальні електронні системи.

Граф знань зручно використовувати при складанні тестових завдань і гіпертексту.

14.11 Типи помилок вимірювань в тестології та їх джерела

Так само, як і в будь-якій технічній галузі, тестологічні вимірювання супроводжуються помилками. Ці помилки можуть бути:

- систематичними, що виникають, наприклад, з вини творців тестових завдань, тестуючої програми, навчального посібника, і
- випадковими, котрі виникають через те, що учасники намагаються «схитрувати», виконати частину завдань методом «випадкового тикуну».

Систематичні помилки виникають через:

- різницю у тлумаченнях визначень в ТЗ і в підручнику;
- неясності в інтерфейсі;
- логічних суперечностей у текстах ТЗ;
- неправильних пояснень, котрі написав супроводжуючий тестування оператор тощо.

На величину помилок впливає людський фактор. Найбільш ймовірні величини випадкових помилок визначення рейтингу, якщо їх відкласти на графіку, утворюють невеликий інтервал значень або, говорячи мовою статистики, - довірчий інтервал. Величина цього інтервалу залежить від того, в яку область рейтингів потрапив досліджуваний рейтинг. Простіше кажучи, - якщо учасник правильно виконав багато завдань, ймовірність того, що він натискав клавіші «зі стелі», мала. І навпаки, ця ймовірність велика для учасників, результати яких потрапили в область «найбільш ймовірного» результату.

У загальному випадку, величину довірчого інтервалу можна представити функцією двох змінних:

- 1) відстані від досліджуваного результату до «ідеального»;

2) відстані від досліджуваного результату до центру області «найбільш ймовірного» результату.

Очевидно, зі зміною цих змінних, величина області «довірчого інтервалу» буде змінюватися в певних межах. Ці межі вкладені між величиною області «найбільш ймовірного» результату і нулем.

Знання величин довірчих інтервалів допоможе на практиці при розборі екзаменаційних апеляцій, а також дуже важливо для правильного застосування методів факторного аналізу.

14.12 Педагогічне тестове завдання

Не кожне завдання, створене для подальшого включення в тест, можна назвати тестовим або педагогічним. ТЗ, які складають педагогічний тест, називають педагогічними.

Ключове завдання - націлене на перевірку знання конкретних блоків навчальної дисципліни. У випадку адаптивного тестування, тест повинен містити не менше 3-х ключових завдань різних рівнів складності для кожного блоку.

Педагогічне завдання - засіб інтелектуального розвитку, освіти і навченості, що сприяє активізації навчання, підвищенню якості знань, а також підвищення ефективності педагогічної праці (В. С. Аванесов).

Конкретне педагогічне завдання є результатом праці по відборі (редагуванню, зміні) сукупності тестових завдань для педагогічного тесту. До того, як завдання будуть апробовані та включені в педагогічний тест, щоб уникнути плутанини, їх називають завданнями у тестовій формі.

Завдання в тестовій формі - це одиниця навчально-контрольного матеріалу. Будучи включеним у педагогічний тест, завдання у тестовій формі стає педагогічним тестовим завданням.

Тестове завдання - самостійний елемент тесту (тобто сукупності тестових завдань), що складається з восьми компонент:

1. інструкції;
2. постановки завдання (власне завдання);

3. демонстраційного матеріалу;
4. еталонів результату виконання завдання серед яких є, як мінімум, один абсолютно правильний;
5. спеціально організованого місця або області для введення результату виконання;
6. емпіричних параметрів - ваги, складності завдання;
7. номер і вказівки приналежності завдання до конкретної теми навчальної дисципліни;
8. шкали оцінювання результату виконання.

В інструкції ТЗ пояснюють, що саме суб'єкт тестування повинен зробити, яким чином, де і які робити позначки і записи. Іноді, для економії часу і ресурсів, дається тільки одна інструкція для групи однотипних завдань, яка міститься на початку групи завдань. В інструкцію прийнято включати приклади.

Текст постановки завдання зазвичай складається із запитання, на яке слід відповісти, або твердження, яке слід підтвердити, виправити або спростувати.

Демонстраційний матеріал завдання повинен бути таким, щоб для знайомства з ним претендент витратив не більше 3-5 хвилин.

Еталони виконання ТЗ визначаються явно (переліком) або неявно (вказівкою можливостей). Еталони відповідають обраній автором якісній або кількісній шкалі оцінювання - правильні, неправильні, неповні, неточні тощо. Частіше, для простоти, автори вибирають двозначну шкалу: правильно - неправильно. Еталонів виконання повинно бути не менше двох. Оптимальним вважається кількість - 4-6. Еталони не демонструються у разі ТЗ відкритої форми, де претендент самостійно вводить знаки в спеціально виділених місцях. Зазвичай вважається, що ТЗ закритої форми повинні містити хоча б один правильний еталон виконання. Їх може бути і декілька.

Наявність певного місця для відповідей обов'язково як при бланковому, так і при комп'ютерному тестуванні:

- У ТЗ закритої форми - це код (цифри або букви) відповіді з числа продемонстрованих. У ТЗ відкритої форми відповідь пишеться замість прочерку;

- У ТЗ на встановлення відповідності відповіді або пишуться в спеціально відведеному для цього рядку нижче тексту завдання, або фіксується за допомогою миші;

- У ТЗ на встановлення правильної послідовності випробуваний ставить ранги в спеціально відведеному для цього місці.

Кожне ТЗ взаємопов'язане з іншими. Цей зв'язок буває явний і неявний. ТЗ часто неявно пов'язане з іншими просто в силу структури тесту, подібно до атомів в кристалі.

Завдання в рамках тесту пов'язані явно, якщо (наведені приклади - образи):

- всі ТЗ присвячені одній різко окресленій темі, і будь-яке додаткове завдання буде зайвим (тест - кристал);

- всі ТЗ мають ваги, ваги є складовими формули розрахунку результату тестування, і будь-яке додаткове ТЗ зруйнує рівновагу (тест - ваги);

- важлива послідовність викладу та перевірки знань (тест - конвеєр);

- важлива послідовність зростання складності ТЗ (тест - еволюція);

- кожне ТЗ займає певну нішу у структурі тесту, як архітектурний елемент у споруді, і без його виконання не можна судити про будь-яку якість претендента (тест - палац).

ТЗ в рамках тесту пов'язані неявно, якщо:

- кожне ТЗ присвячено окремому блоку інформації, не пов'язаного з іншими відношенням порядку, а сам тест формується з безлічі ТЗ випадковим чином;

- ТЗ склалися без урахування того, скільки і яких завдань вже увійшло до складу тесту.

ТЗ можуть бути розмежовані, тобто результати їх виконання свідчать про різні складові знань претендента та взаємопроникаючими, коли результат виконання одного ТЗ підтверджує результат виконання іншого.

Якщо з будь-якого тесту зробити вибірку меншої кількості ТЗ, то може утворитися система, яка не поступається помітно за своїми властивостями початковому тесту. Тест з меншою кількістю ТЗ тоді можна називати більш ефективним.

14.13 Чотири форми тестових завдань

У тестологічній практиці прийнято вважати, що існують всього чотири форми тестових завдань. Все інше розмаїття можна класифікувати як їх варіації.

Разом з тим, існує думка, що чотирма типами все описати не можна. Тестологія розвивається. Наприклад, не так давно з'явилася методика диференційованої оцінки виконання завдань. Вона відома під назвою «grading» (градування, розподілення). У цій методиці передбачається, що оцінка виконання завдань може бути не тільки двозначною «правильно - неправильно», а й приймати цілий спектр значень. Вважається, що найбільше ця методика може бути застосована в математиці.

Ось назви (характеристики) всіх чотирьох традиційних форм тестового завдання:

- завдання з вибором відповідей (випробуваний вибирає правильні відповіді з числа готових, пропонованих в завданні тесту) - завдання в закритій формі;
- завдання на доповнення (випробуваний сам дає коротку або розгорнуту відповідь) - завдання у відкритій формі;
- завдання на встановлення відповідності між елементами двох множин;
- завдання на встановлення правильної послідовності у низці елементів.

Вибір форми залежить від:

- цілі тестування;
- змісту тесту;
- технічних можливостей;
- рівня підготовленості викладачів та персоналу.

Наведемо (з деякими скороченнями) об'ємну цитату, що прояснює взаємовідношення форми і змісту тестових завдань:

«Овладение формой является необходимым, но недостаточным условием создания полноценных тестов. Форма придает заданиям лишь структурную целостность и определенность, внешнюю организованность. Задания в тестовой форме только внешне похожи на тестовые задания, но это недостаточно для их включения в тест; нужна еще проверка свойств, позволяющих включить их в тест. Вместе с тем, меняющееся от дисциплины к дисциплине содержание позволяет абстрагировать форму в качестве самостоятельного предмета и рассматривать ее как подобие инварианта, независимого, в значительной степени, от содержания конкретной учебной дисциплины. Тем самым здесь подтверждается общий философский тезис о сравнительном постоянстве формы, остающейся устойчивой при изменении содержания.

Начиная с Аристотеля, форма понимается как идеальный принцип расположения элементов. Другое истолкование выдвинул И. Кант, согласно которому форма есть принцип упорядочения. Эти две идеи – организации и упорядочения, плюс идея существования содержания в каких-либо формах легли в основу нашего определения формы тестовых заданий: это способ организации, упорядочения и существования содержания теста. Исследованием научных основ разработки тестов занимается педагогическая тестология. В западных странах получили развитие эконометрика, биометрика, социометрия и другие науки, имеющие своим предметом изучение количественных свойств и отношений, разработку объективных показателей состояния интересующих явлений.

Общие принципы разработки формы тестовых заданий рассматриваются в тестологии как важный предмет исследования, в то время как содержание, будучи не менее важным компонентом хорошего теста, относится все же не столько к тестологии, сколько к той науке, знание которой проверяется.»

14.14 Закрита форма завдання

Основою цієї форми є закон виключеного третього, сформульований Аристотелем. Вибір правильного варіанта породжує істину, вибір неправильного - брехню. Третього не дано. З цього закону випливає: кожне завдання повинно мати один вірний варіант виконання.

Найчастіше в літературі зустрічаються завдання в закритій формі, що мають питання і чотири варіанти відповіді, з яких один - правильний. Це - вже класика.

Для вибору правильного варіанту випробовуваний повинен зробити одну з дій:

- ввести код (цифри або букви) необхідної відповіді в спеціальному вікні;
- клацнути мишею по «радіокнопці» з номером варіанту;
- клацнути мишею по тексту варіанту.

Головними недоліками класичної закритої форми є:

- висока ймовірність вгадування вірного варіанту;
- непродуктивні витрати часу на прочитання всіх варіантів;
- можливе мимовільне підсвідоме запам'ятовування піддослідним неправильних варіантів як правильних.

Форма накладає логічну заборону на застосування таких варіантів, як:

- «правильної відповіді немає»;
- «всі відповіді правильні»;
- «всі відповіді неправильні».

Вона часто порушується як у зарубіжних, так і в російських тестах. Порушення ці свідчать про неухважність або про відсутність культури логічного мислення серед укладачів тестових завдань.

У публікованих у пресі завданнях закритої форми часто зустрічаються логічні казуси, невідповідності, багатозначності. Це відбувається в тому випадку, коли:

- питання містить приховане питання, яке деякі, особливо спокушені в тестуванні суб'єкти тестування можуть прийняти за основний;

- питання містить додаткову контекстну інформацію, спроможну здатися невірною частині суб'єктів тестування (особливо це стосується історії, політології тощо);

- питання багатослівні, містять додаткові звороти;

- у завданні відсутнє завдання або питання (тобто є тільки текст в області завдання і чотири варіанти тексту в області відповідей);

- завдання містить твердження з трьома крапками і варіанти на заміщення крапки.

Найчастіше логічна неузгодженість зустрічається в четвертому із цих випадків.

Зазначені та ряд інших недоліків є у всіх тестах. Є три шляхи боротьби з ними:

- повторне авторське редагування;

- самотестування автора;

- проведення тестування в референтних групах, що складаються не тільки з тих хто навчається, але й з викладачів.

Найбільш швидким, економічним і ефективним слід визнати третій шлях.

Є два досить відомих шляхи розвитку закритої форми тесту, що заповнюють зазначений недолік:

- завдання з оцінюванням правильності, тобто вибором одного, найбільш правильного варіанта з декількох правильних різного ступеня;

- завдання з вибором кількох вірних варіантів з декількох запропонованих.

Обидва вони так само вважаються завданнями в закритій формі.

Є третій і четвертий шляхи розвитку завдання в закритій формі, що ще більше знижують вірогідність вгадування, опису яких приділено особливу увагу:

- завдання з оцінюванням ступеня вірності кожного з варіантів;

- завдання з оцінюванням ступеня вірності одного з варіантів, де випробований не може бачити всі варіанти одночасно.

В останньому випадку випробуваному надається можливість переглядати варіанти, що виводяться у випадковому порядку.

Третій і четвертий шляхи, крім наявності очевидних переваг в оцінюванні ступеня володіння знаннями, дуже перспективні для:

- розвитку евристичного мислення;
- навчання студентів педагогічних спеціальностей.

Ситуації, в яких необхідний вибір, особливо множинний, як в останніх випадках, часто викликає у випробовуваних відчуття невпевненості. Джерелом сумнівів є:

- відсутність інформації про альтернативи;
- суб'єктивною цінністю альтернатив;
- конфліктний (безконфліктний) характер альтернатив.

Досвід показує, що невпевненість суб'єктів тестування зникає після перших успіхів.

14.15 Відкрита форма завдання

Завдання у відкритій формі є зазвичай твердженням або великим текстом. У тексті є спеціальні місця, куди слід вносити інформацію, обумовлену інструкцією. Зазвичай це «клітки», в кожен з яких вписуються літери, цифри або залишаються прогалини. Після заповнення клітин, текст перетворюється на правильне або хибне висловлювання.

Зазвичай використовується три види тестових завдань відкритої форми:

- завдання доповнення, де випробовуваний повинен сформулювати доповнення до пред'явлених текстів;
- завдання - кросворди, орієнтовані на з'ясування знання претендентами визначень, що можуть містити підказки на перетинах слів та мати ігровий елемент;

- завдання вільного викладу, де потрібно викласти повне рішення задачі, твір або переказ тексту, де майже ніякі обмеження на виконання завдання не накладаються.

Правильно сконструйовані завдання відкритої форми повністю виключають здогадку - головний недолік завдань закритої форми. Це - головне достоїнство таких завдань.

Завдання відкритої форми часто викликають нерозуміння серед педагогів у зв'язку з частотою однакових випадків: претенденти, які виконують завдання відкритої форми з географії (історії, літератури, математики ...) роблять орфографічні помилки. Чи вважати варіанти з орфографічними помилками помилковими? Чи ставити учню, що відмінно знає теореми і вміє вирішувати завдання, двійку з математики за те, що він неправильно пише слова «пряма» і «площина»?

Однак завдання вільного викладу мають істотний недолік - важко прогнозована багатоваріантність. Результати їх виконання неможливо повноцінно перевірити автоматично. Принаймні, на сьогоднішньому етапі розвитку комп'ютерних інтелектуальних систем. Можна з цим боротися:

- приділяти більше уваги формулюванню завдань, щоб істотно скоротити кількість варіантів;
- вказувати кількість літер у допустимому варіанті виконання;
- передбачити всі можливі правильні варіанти.

Всі ці спроби призводять до різкого обмеження довжини можливих текстів, що вводяться претендентами як варіанти відповіді.

Враховуючи, що кількість цифр і букв обмежена, можна вважати завдання у відкритій формі еквівалентним завданням в закритій формі, тільки з дуже великою кількістю (мільйонами, мільярдами) неправильних варіантів виконання, прихованих від претендента.

14.16 Завдання на встановлення відповідності

Це завдання, де до елементів однієї множини потрібно правильно підібрати елементи другої. Варіантом виконання завдання є перелік зв'язків

(відповідностей) елементів. Для позначення цих зв'язків діють так. Елементи однієї множини позначаються цифрами, іншої - літерами. Відповідь конструюється як послідовність пов'язаних цифр і букв. Наприклад, для двох п'яти елементних множин, потрібно ввести як результат виконання завдання кодове слово: 1P2M3B4E5H6L.

Результати виконання завдання на встановлення відповідності:

- записують у відведеному для цього рядку за обговореним в інструкції правилом;
- можуть вказуватися мишкою графічні зображення зв'язків (відповідностей).

Завдання на встановлення відповідності рекомендуються для перевірки асоціативних і алгоритмічних навичок користування знаннями. Пошук аналогій на основі навчального матеріалу, дозволяє робити висновки не тільки про володіння учнем тими чи іншими поняттями, а й про інші характеристики структури знань.

Завдання на встановлення відповідності дозволяють вести перевірку засвоєння відразу з декількох дисциплін. Тому вони можуть мати великий діапазон для перевірки.

Виконання таких завдань легко перевіряти шляхом порівняння з єдиним правильним варіантом. Це жорсткий підхід. Але претендент може зробити одну помилку - невже через це всі завдання він виконав неправильно? Перевірити виконання таких завдань «м'яко», диференційовано - вручну не можна, а з комп'ютером - можна.

Часто зустрічається недолік завдань третьої форми - «особливо начитані» претенденти вбачають у таких завданнях не тільки ізоморфні (однозначні) зв'язки, а й гомеоморфні (багатозначні). Це призводить до необґрунтованого зниження оцінок і справедливо поданих апеляцій.

Оскільки результат виконання завдання третьої форми - кодове слово з чітко обмеженою кількістю літер, в цьому розумінні ця форма збігається з другою (відкритою).

14.17 Завдання на встановлення послідовності

У цих завданнях потрібно встановити правильну послідовність обчислень, дій, кроків, операцій, термінів.

Встановивши правильну, зі своєї точки зору, послідовність, випробуваний вводить свій варіант у спеціально відведеному для цього місці шляхом:

- або клацання мишею по послідовності мнемонічних зрозумілих елементів (цифр тощо);
- або введення послідовності цифр з клавіатури.

Ця форма завдань викликає часто нарікання не на себе, а через недоліки інтерфейсу та інструкції, що призводить до різночитань алгоритму введення варіанту.

Цю форму застосовують для перевірки правильності розуміння піддослідними ходу рішення.

Четверта форма завдань стає еквівалентною третій, якщо уявити одну з елементних множин завдання третьої форми простою послідовністю цифр.

14.18 Вага тестового завдання. Чинники які знижують цінність поняття ваги ТЗ

Можна «зважити» завдання одне до одного всередині кожного рівня, якщо це необхідно. Однак тут слід мати на увазі наступні міркування.

Обчислення відносної ваги завдань, тобто їх відносної складності для учнів - досить складне завдання. Воно вимагає проведення великої кількості тестувань за участю багатьох учнів. При цьому на величини обчислюваних Вами терезів можуть непередбачувано вплинути наступні фактори:

- різні учні в різний час можуть користуватися при підготовці до тестування різними підручниками, де одне і теж питання висвітлено з різним ступенем ясності;
- зміна (або тимчасова підміна) викладача дисципліни призводить до висвітлення різних питань у різних навчальних групах з різним ступенем ясності. Тому при зважуванні можна вагу одного і того ж питання для одних обрати надто високою, для інших - низькою;

- учні, які пройшли тестування першими неминуче передадуть почерпнуту інформацію наступним, і ті придуть тестуватися більш підготовленими. Неодночасність сеансів тестування і неконтрольованість поширення інформації серед учнів про тестові завдання зроблять Ваші спроби «скрупульозного зважування» безглуздими;

- учні краще за всіх знають навчальний матеріал, краще за всіх прогнозують для себе час, необхідний для виконання завдань. Через те, вони виконають в першу чергу прості завдання, а складні залишать на «потім». Це призведе до помилкового розуміння складності завдань - за «складні» Ви помилково приймете «довгі»;

- робочі навчальні програми з дисциплін інколи змінюються у відповідності з поглядами керівництва. Кожна така зміна буде руйнувати Ваші результати «зважувань» і примушувати Вас прийматися за обрахунок цього завдання знову.

- величини ваг ТЗ, обчислені на різних вибірках претендентів, вийдуть різні. Отже, претендент, результати тестування якого враховуються в сукупності з різними вибірками, буде мати різний результат.

Таким чином, для учнів, які навчалися у різних навчальних групах і в різні періоди часу, ваги завдань, обчислені за результатами тестувань інших груп, будуть відносно необ'єктивні. Крім того, досвід роботи журі олімпіад говорить, що за якою б системою ваги не обчислювалися, додаткової об'єктивності рішенням журі це ніколи не додає і список призерів від цього змінюється мало.

Тому «складність» завдань треба вираховувати по-іншому. Наприклад, шляхом переходу до багатовимірної векторної оцінки, векторному рейтингу. Компоненти вектора - рейтингу характеризують відносний успіх претендента у виконанні конкретних груп ТЗ.

14.19 Трудомісткість і складність тестового завдання

У класичній тестології поняття «важкість» і «складність» використовували як синоніми. Різні автори пропонували розрізняти їх. У сучасній тестології

термін «важкість» замінюють на більш відповідний - «трудомісткість» (запропоновано В. П. Беспалько).

- Трудомісткість завдання - міра кількості та тривалості передбачуваних операцій, необхідних для успішного виконання завдання.

- Складність завдання - міра об'єму і складності інформації, яку повинен освоїти претендент для того, щоб виконати завдання.

Якщо претенденти мають достатньо часу, щоб продумати варіанти виконання та не змагатися в швидкості, визначальним чинником стає складність завдань. Тому складність – слово яке найбільш часто використовується для характеристики завдань.

Складність завдання залежить від шляху вивчення навчальної теми суб'єктами тестування (за будь-яким підручником), від їх підготовки та кмітливості. Складність приймає різні значення для різних вибірок претендентів. Одне й те саме завдання для студента одного вузу може стати нездоланим, для студента іншого - «наслідком». Чим більше претендентів, правильно виконали завдання, тим менша його складність для даної групи. На величину складності можуть впливати артефакти, наприклад – втома, що накопичується в процесі довгого сеансу або недостатність часу сеансу. Просте завдання можуть не виконати більшість претендентів тільки тому, що поки вони до нього дісталися, вони втомилися. Або, взагалі не встигли дістатися. Так просте завдання з великим номером автоматично формально стає дуже складним.

І в школі, і у вузі часто буває ситуація, коли:

- учні однієї паралелі навчаються у різних викладачів і по різних підручниках;

- викладачі змінюються в навчальній групі протягом навчального року.

А якщо так, то зусилля по емпіричному визначенню складності завдань тесту не приведуть до результату, і складність завдань буде приймати різні значення.

Обчислена складність завдань може реально використовуватися в подальших тестуваннях за умови збереження трьох зовнішніх параметрів:

- особистість викладача і матеріал лекцій, контрольних робіт тощо;
- комплект навчальних посібників;
- параметри суб'єктів тестування (тобто групи учнів, котрі піддаються тестуванню, повинні статистично не відрізнятися за характеристиками).

В іншому випадку, складність завдань доведеться перераховувати після кожного сеансу тестувань перед визначенням рейтингів. А тоді може статися казус:

- суб'єкти тестування, які тестувалися в складі різних груп і отримали однаково високий рейтинг, будуть значно відрізнятися своїми характеристиками;
- суб'єкт тестування, який пройшов тестування по одному тесту, але в різних групах, отримує різний рейтинг.

Один із шляхів подолання цих проблем - перехід до векторного рейтингу.

14.20 Підстави для типології тестових завдань

Створювані в даний час в Росії БТЗ, на перший погляд, такі різні, що важко їх зіставити для подальшого спільного використання. Ряд дослідників, які займаються цією проблемою, запропонували способи класифікації БТЗ.

ПЕРША ДЮЖИНА

підстав визначена комбінованою типологією ТЗ і запропонована Є.А. Михайличевим (Михайлычев Е. А. Дидактическая тестология. М.: Народное образование, 2001. - 432 с.) на основі типологій Г.А. Балла, В.І. Андрєєва, В.П. Беспалько.

1. Орієнтація на діагностичний рівень засвоєння (за концепцією В. П. Беспалько):

- I рівень;
- II рівень;
- III рівень (нетипова задача, пов'язана з графічними і модельними побудовами, не має еталона відповіді);

- IV рівень (евристична задача, не має еталона відповіді).
2. Характер узагальненості предмета завдання:
 - індивідуальний;
 - родовий.
 3. Характер предмета вирішення задачі:
 - матеріальний;
 - інформаційний.
 4. Характер розв'язання тестової задачі:
 - принципово нерозв'язна;
 - принципово вирішувана.
 5. Суб'єктивна підготовленість претендента до вирішення тестового завдання:
 - рутинна;
 - не рутинна.
 6. Чіткість побудови тестової задачі:
 - чітка (алгоритмічна);
 - нечітка.
 7. Націленість формулювання на адекватність сприйняття:
 - орієнтована на адекватне сприйняття;
 - орієнтована на неадекватне сприйняття.
 8. Можливість зміни предмета завдання:
 - теоретична;
 - практична.
 9. Особливості формулювання завдання:
 - завдання на виконання;
 - завдання на перетворення;
 - завдання на відновлення;
 - завдання на побудову;
 - завдання на використання процедури;
 - завдання на використання наявного стану, умови.

10. Логічна схема побудови завдання:

- індуктивна;
- дедуктивна;
- змішана (індуктивно-дедуктивна).

11. Домінуючі прийоми методу пізнання, необхідного для вирішення задачі:

- логічна;
- аналітична;
- експериментальна;
- графічна.

12. Форма організації виконання завдання:

- індивідуальна;
- групова;
- фронтальна.

На основі цієї схеми можна описувати основні аспекти орієнтації ТЗ. Вона ж може допомогти в плануванні структури новостворюваної БТЗ.

У деяких випадках вдається домогтися ще більш детальної класифікації завдань. Для цього доповнимо наведену першу дюжину підстав ще однією.

ДРУГА ДЮЖИНА ПІДСТАВ

може бути побудована на підставі роботи (Васильев В. И., Тягунова Т. Н. *Теория и практика формирования программно-дидактических тестов. М.: Издательство МЭСИ, 2001. - 130 с.*):

13. Форма:

- на зазначення правильних варіантів (закрита форма);
- на виправлення або введення інформації (відкрита форма);
- на зазначення відповідності;
- на зазначення послідовності.

14. Тип оцінювання:

- з конструктивною оцінкою
- з неконструктивною оцінкою

15. Шкала (логіка) оцінювання

- з двозначною шкалою оцінювання (вірно - не вірно)
- з тризначною шкалою оцінювання (вірно - не зовсім вірно - не вірно)
- з чотиризначною шкалою оцінювання (вірно - не повно - не точно - не вірно)

16. Форма демонстраційних матеріалів

- не мультимедійне
- мультимедійне
- графічне
- текстове

17. Тип носія

- бланкове;
- комп'ютерне.

18. Величина емпіричних параметрів:

- вага;
- трудомісткість;
- складність.

19. Зміст:

- інформаційно-відображаючий;
- проблемно-евристичний;
- розвиваючий.

20. Спрямованість на з'ясування характеристик претендента:

- знань;
- умінь;
- навичок.

21. Загальні вимоги:

- технологічності;
- форми;
- змісту

22. Статистичні вимоги:

- відомої складності;
- достатньої варіації тестових балів;
- позитивної кореляції балів завдання з балами по всьому тесту.

23. Орієнтація:

- нормативні тести - орієнтовані на статистичні норми, підставою для порівняння в яких служать статистично отримані дані про виконання тесту репрезентативною вибіркою претендентів;

- критеріальні тести - орієнтовані на критерії, призначені для визначення рівня досягнень претендента щодо заданого критерію;

- прогностичні тести - орієнтовані на прогноз успішності діяльності.

24. Вид тестування:

- статичне тестування - направлено на вимірювання результатів навчання, але не може оцінити потенціал претендента;

- динамічне тестування - спрямовано на виявлення прихованих здібностей претендента, а також оцінку - наскільки успішно він може далі навчатися.

Найчастіше автори тестів не користуються типологіями, складають тести так, як вони вважають за потрібне. Їх виправдовує прагматична мета - не підвищення якості вимірювання, а надійна диференціація груп претендентів. Тестування з такими тестами схожі на звичайну екзаменаційну лотерею, в якій гарантовано виграє тільки підготовлений претендент. Найчастіше, властивостей таких тестів цілком достатньо для практичного застосування.

У згаданій книзі Є. А. Михайличева наведений великий огляд різних типологій тестових завдань. В даний час найбільш поширеною типологією ТЗ є диференціація за чотирма формами.

14.21 Характеристики ЗУН, що виявляються шляхом тестування

Наведемо класифікацію критеріїв суб'єктів тестування, виділених різними авторами. Розіб'ємо їх на чотири великі групи.

1. Кругозір:

- знання основних фактів;
- знання історії науки;

- методологічна підготовка (класифікації, внутрішньопредметні і міжпредметні зв'язки);

- вміння застосовувати філософські методи, розуміння протиріч, розрізнення форми та змісту;

- знайомство з основною літературою;

- словниковий запас.

2. Знання основ дисципліни (теми):

- знання законів;

- знання визначень і ключових понять, номенклатури;

- знання правил;

- знання основних концепцій.

3. Уміння вирішувати стандартні завдання в рамках дисципліни (теми):

- вміння застосовувати знання у вирішенні простих завдань;

- вміння захищати свої погляди;

- вміння логічного викладу;

- вміння застосувати теорію до практики, описати і пояснити просте явище;

- наявність стандартних умінь і навичок (побудова таблиць і графіків, формул, проведення експериментів, аналіз текстів та зображень);

- вміння перевести стандартний текст на іншу мову.

4. Вміння вирішувати проблеми:

- розуміння взаємозв'язку тем і розділів курсу, вміння застосувати взаємозв'язки;

- знайомство з періодичною і не основною літературою;

- спостережливість, вміння виокремити головне;

- вміння розпізнати структуру тексту (проблеми, зображення, формули, явища);

- вміння застосувати міжпредметні знання, «наскрізні» навчальні вміння і навички.

14.22 Вимоги до тестових завдань

Тестові завдання мають відповідати наступним вимогам:

1. Адекватність - відповідність інструкції, форми та змісту завдання;
2. Виваженість - наявність ваги, тобто певного, вираженого у вигляді числа або відношення порядку рівня складності або трудомісткості завдання;
3. Дистрактності - ступінь схожості неправильних варіантів на правильні і навпаки;
4. Диференціююча здатність - властивість окремих завдань і тесту диференціювати претендентів на групи відповідно до ступеня їх підготовки;
5. Доступність - потенційна можливість виконання для більшості претендентів;
6. Інтегрованість, синтетичність - наявність або відсутність міжпредметних зв'язків;
7. Коректність - відсутність малозрозумілих і рідковживаних виразів, а також невідомих для претендентів символів і слів, що ускладнюють сприйняття змісту;
8. Стислість - ретельний підбір слів, символів, графіків, дозволяє мінімумом засобів домогтися максимуму ясності смислового змісту завдання;
9. Критеріальність - наявність можливості пред'явити претендентам не тільки двозначну шкалу оцінки варіантів виконання («вірно - не вірно») а й інші, у відповідності з критеріями (наприклад, «вірно - не точно - не повно - не вірно»);
10. Багатофункціональність - наявність можливості діагностики і пам'яті, і мислення;
11. Оригінальність - мінімізує ймовірність користування шпаргалкою;
12. Відкритість - претенденти мають можливість познайомитися з подібними ТЗ завчасно;
13. Зрозумілість - відсутність подвійних заперечень, прихованих питань, того, що мається на увазі тощо;

14. Привабливість - присутність частки гумору та подібних хитрощів, сприяння розвитку бажання претендентів до самоосвіти;

15. Технологічність - простота складання та коригування, наявність чітких правил складання та оцінювання ваг;

16. Толерантність - відсутність фрагментів, що викликають нерозуміння і обурення, пов'язані з відмінностями людей за об'єктивними ознаками (національність, стать тощо);

17. Універсальність - можливість використання в різних тестових комплексах;

18. Формальність - відповідність обраній формі;

19. Ергономічність виконання і обробки - мінімальність затрат часу на прочитання, розуміння, відповідь; мінімальність витрат ресурсів і часу на перевірку і аналіз результатів.

20. Доцільність.

14.23 Недоліки тестових завдань та їх причини

У вітчизняній і зарубіжній науковій тестологічній літературі неодноразово відмічалось, що більшість укладачів тестових завдань здійснюють ряд однакових помилок. Перелічимо основні помилки в порядку частоти появи.

1. У списку відповідей правильна відповідь, як правило, найдовша. Тому його неважко знайти. Це основна помилка початківців укладачів тестових завдань. Часто учні про це знають, і вміло користуються цими знанням.

2. Неправильні варіанти відповідей бувають не узгодженими з питанням завдання у відмінках, родах, числах тощо. Автори та редактори звертають на них уваги менше, ніж на правильні варіанти.

3. У варіантах виконання завдань зустрічаються перехресні логічні та інформаційні посилання, що ускладнює знаходження правильного варіанту, якщо відповіді розставлені комп'ютером за випадковим алгоритмом.

4. Правильний варіант виконання завдання можна з великим ступенем ймовірності «вчислити» логічно, шляхом виключення явно невірних.

5. У питанні основного завдання міститься ще одне приховане питання, яке, особливо спокушені в тестуваннях («особливо обдаровані») випробовувані, шукаючи логічний підхід, можуть прийняти за основне.

6. Питання містить додаткову контекстну інформацію, яке може здатися невірною частині суб'єктів тестування (особливо це стосується історії, політології тощо).

7. Питання завдання багатослівні, двозначні, містять додаткові звороти.

8. У деяких пунктах тесту відсутнє завдання або питання, а є тільки текст в області завдання і чотири варіанти тексту в області відповідей.

9. Завдання містить твердження з трьома крапками і варіанти заміщення крапки.

10. Завдання передбачає непропорційно високі витрати часу на знайомство випробуваного з усіма варіантами виконання.

11. Завдання передбачає можливе мимовільне підсвідоме запам'ятовування випробуваними неправильних варіантів як правильних.

12. Завдання містить логічно неможливі варіанти виконання. Наприклад, в літературі часто обговорюється відому логічну заборону на застосування таких варіантів відповіді у завданнях закритого типу, як:

- «правильної відповіді немає»;
- «всі відповіді правильні»;
- «всі відповіді неправильні».

13. У завданні відсутня інструкція і важко визначити, що ж потрібно від випробуваного на цей раз. Наприклад, у завданнях третьої і четвертої форм, коли на екрані комп'ютера присутні два-три однаково нумерованих елементи, а в завданні пропонується «вказати» один з них, випробуваний змушений звертатися за тлумаченням до чергового в класі, який може не мати потрібної інформації.

14. Порушення ізоморфізму. У завданнях третьої і четвертої форм, де потрібно перенумерувати або зіставити один одному елементи, зустрічаються

елементи, які можна зіставити по-різному або які можна одночасно зіставити з кількома елементами.

15. Присутність логічного зв'язку між різними варіантами виконання завдання, що не дозволяє розглядати їх по-окремо, наприклад, якщо комп'ютер демонструє їх по черзі, або у випадковому порядку.

16. Питання завдання не однозначне, містить в собі більше одного питання (тобто кілька підпитань). Техніка складання таких питань використовується в соціологічних анкетах для «підштовхування» людини до потрібної відповіді, але в шкільних тестових завданнях вони з'являються виключно через недосвідченість укладачів. Підпитання часто логічно приховані та видні тільки фахівцеві. Тому іноді важко однозначно визначити, яку відповідь автори вважають повною і правильною.

17. Питання завдання містить подвійне заперечення. У цьому випадку часто можна побудувати не менше двох правильних і абсолютно протилежних за змістом відповідей на таке питання.

18. Питання завдання не має на увазі повну однозначну відповідь, не окреслює коло обов'язкової інформації, яку має містити відповідь. Така помилка укладачів часто зустрічається у завданнях всіх чотирьох форм.

- Наприклад, на запитання: «Де відбулася Бородінська битва?», можна абсолютно точно відповісти:

- «Під Москвою», - «У Росії», - «У селі Бородіно», - «У полі».

- Наприклад, у завданнях з питаннями типу: «Хто першим винайшов ...?» не пояснюється, чи повинна відповідь містити прізвище, ініціали, національність вченого, а також один чи декілька вчених мають на увазі. Не враховується, що питання «Хто?» застосовується в українській мові як для означення однієї людини, так і для означення групи людей.

- Наприклад, питання сформульоване так: «Назвіть клички кращих англійських скакунів ...». Питання не містить пояснення, скільки кличок потрібно назвати.

- Підмет у називному відмінку важко відрізнити від прямого додатка у знахідному відмінку, наприклад: «Мати любить дочку» (приклад В. С. Аванесова);

- На запитання: «Коли почалася Велика вітчизняна війна?», - можна відповісти:

- «В червні», «В ХХ столітті» та ін;

- У завданнях відкритої форми слід вказувати рід, відмінок та інші граматичні категорії поданих слів.

Наприклад, на питання: «Як називається судно, яке розколює лід?», можна відповісти як «криголам» так і «криголамом».

- Серед завдань на встановлення відповідності або на встановлення послідовності зустрічаються варіанти, що трактуються в різних джерелах по-різному (причина і послідовність історичних подій).

19. Питання неоднозначне або варіанти відповіді не настільки повні й однозначні, як мається на увазі у завданні. Наприклад, не можна в завданні ставити питання: «Коли почалася Велика вітчизняна війна?». Слід завдання сформулювати так:

«Назвіть дату початку Великої вітчизняної війни». Адже правильна відповідь для першого формулювання неоднозначна: «У червні», «У ХХ столітті» тощо.

20. Варіанти відповіді не узгоджені з питанням в числі та відмінку, дають привід для протирічного тлумачення.

21. ТЗ містять орфографічні та інші мовні помилки. Ці помилки, ставлять в логічний глухий кут учня, оскільки йому не відомо, випадкова це, чи навмисна помилка.

22. В інструкції завдання відкритого типу не вказується, в якому числі і відмінку повинен бути написаний варіант словосполучення. У підсумку - введений принципово правильний варіант може бути прийнятий при автоматичній перевірці за невірний.

23. У завданні відкритого типу можливо кілька формулювань відповіді, а упорядником буває передбачена тільки одна. У підсумку - введений принципово правильний варіант, може бути прийнятий при автоматичній перевірці за невірний.

24. Завдання одноманітні, нудні. Послідовність таких завдань негативно впливає на учнів, викликають так званий ефект «рівної дороги» - відомий в практиці водіїв-далекобійників ефект засипання за кермом, відволікання і розсіювання уваги, коли дорога дуже рівна й безлюдна.

Часто учні знають про типові помилки, і вміло користуються цими знаннями, отримуючи високі бали. Відомі випадки, коли жартівник, обізнаний про такі помилки, впевнено здавав іспит, наприклад, з спеціальної медичної дисципліни.

Застосовуються в системі освіти БТЗ часто далекі від ідеалу. Наведемо вагомі причини цього об'єктивного положення:

1. Складати тексти так, щоб вони були легкі для читання і відповідали правилам української мови, це не так просто, як уявляє собі дехто. Не кожний учитель здатний сформулювати питання так, щоб воно однозначно сприймалося всіма;

2. У спеціальних наукових дисциплінах вкоренилися жаргонні словосполучення, які можуть бути застосовані в колі однодумців для оптимізації часу спілкування, але не в друкованих матеріалах. Часто такі словосполучення суперечать мовним нормам, і зміст їх важко сприймається неспеціалістами;

3. Об'єктивно неможливо повністю виключити ймовірність вгадування правильної відповіді, хоча й існує цілий ряд методик, що дозволяють її мінімізувати. Порівняйте два завдання (приклад А. Анастасі):

- Погано складене завдання: «Нелегальний: несекретний, незаконний, сексуальний, безграмотний, зухвалий».

- Добре складене завдання: «Нелегальний: секретний, протизаконний, сексуальний, неграмотний, зухвалий».

У першому випадку слово «незаконний» передбачає правильну відповідь з префіксом «не», а в другому - слово «неграмотний» наводить на помилковий слід;

4. Можна стверджувати, що практично в будь-якому тесті можна знайти неправильні, з точки зору української мовної культури, фрази, котрі не узгоджені в роді, числі, відмінку та інше;

5. Відсутність культури критики. У вузівських тестах (так само, як і в навчальних посібниках) значно менше неправильних випадків вживання того, чи іншого слова, ніж у шкільних. Причина цієї культурно-мовленневої переваги в наступному. У вузівському середовищі, на відміну від шкільного, заохочується критика: критичні зауваження з приводу методичних посібників публікуються. Розгляньте критичні зауваження до збірників тестів для школярів і студентів, рекомендовані для підготовки до тестувань;

6. Немає межі досконалості. Ідеал важко досягти ще й через суперечність спірності розроблених вимог до тесту.

СПИСОК РЕКОМЕНДОВАНОЇ ЛІТЕРАТУРИ

Нормативно-правова база

1. Закон України “Про вищу освіту” // Законодавство України про освіту. Збірник законів. – К. : Парламентське вид-во, 2002.
2. Закон України “Про загальну середню освіту” // Законодавство України про освіту. Збірник законів. – К. : Парламентське вид-во, 2002.
3. Закон України “Про освіту” // Законодавство України про освіту. Збірник законів. – К. : Парламентське вид-во, 2002.
4. Наказ МОН від 29 січня 2007 року № 64 “ Про організацію та проведення міжнародних порівняльних досліджень якості освіти учнів загальноосвітніх навчальних закладів”.
5. Указ Президента України від 17 квітні 2002 р. № 347 «Про Національну доктрину розвитку освіти» // У кн.: Законодавчі акти України з питань освіти. – К. : Парламентське вид-во, 2004.
6. Указ Президента України від 4 липня 2005 року №1013 “Про невідкладні заходи щодо забезпечення функціонування та розвитку освіти в Україні”.
7. Указ Президента України «Про невідкладні заходи щодо забезпечення функціонування та розвитку освіти в Україні» // <http://www.president.gov.ua/documents/2928.html>.
8. Закон України «Про вищу освіту» // <http://zakon.rada.gov.ua/cgi-bin/laws/main.cgi?nreg=2984-14>.

Література

1. Аванесов В. С. Композиция тестовых заданий. Учебная книга. 3 изд., доп. – М.: Центр тестирования, 2002 г. – 240 с.
2. Адаптивное тестирование : учеб.-метод. пос. / [Н. М. Опарина, Н. Г. Полина, Р. М. Файзулин, И. Г. Шрамкова].– Хабаровск, 2007. – 95 с.
3. Анастаси А. Психологическое тестирование / А. Анастаси, С. Урбина. – СПб. : Питер, 2005. – 688 с.

4. Башмаков А. И., Башмаков И. А. Разработка компьютерных учебников и обучающих систем. – М.: Информационно-издательский дом «Филинь», 2003. – 616 с.
5. Беспалько В. П. Основы теории педагогических систем (Проблемы и методы психолого-педагогического обеспечения технических обучающих систем) / Владимир Павлович Беспалько. – Изд-во Воронежского ун-та, 1977. – 304 с.
6. Білоусова Л. І. Потенціал комп'ютерного тестування / Л. І. Білоусова // Вісник ТІМО. – 2008. – № 10. – С. 40 – 44.
7. Булах І. Є. Створюємо якісний тест: навч. посіб. / І. Є. Булах, М. Р. Мруга. – К. : Майстер-клас, 2006 – 160 с.
8. Васильев В. И., Тягунова Т. Н. Теория и практика формирования программно-дидактических тестов. – М.: Издательство МЭСИ, 2001. – 130 с.
9. Вісник ТІМО, 2005 – 2010 р.
10. Вопросы тестирования в образовании / Под ред. Хлебникова В. А., Неймана Ю.В. М.: «Век книги», 2001. – 115 с.
11. Гавриленко І. М. Соціологічний моніторинг і діагностика в освіті / І. М. Гавриленко // Освіта і управління. – 1998. – № 2. – Т. 2.– С. 7 – 9.
12. Гузеев В. В. Оценка, рейтинг, тест. Школьные технологии. 1998, №3. – С. 3-40.
13. Дзюбенко А. А. Новые информационные технологии в образовании. М.: 2000. – 104с.
14. Дистанционное обучение: Учеб. пособие / Под ред. Е . С. Полат. – М.: Владос, 1998. – 192 с.
15. Енциклопедія освіти / Акад. пед. наук України ; головний ред. В. Г. Кремень. – К. : Юрінком Інтер, 2008. – 1040 с.
16. Клайн П. Справочное руководство по конструированию тестов / Пол Клайн. – К., 1994. – 283 с.

17. Кречетников К. Г. Задания в тестовой форме и методика их разработки: Учебно – методическое пособие. – Владивосток: Изд-во Дальневост. ун-та, 2002. – 36 с.
18. Кречетников К. Г. Теоретические основы создания креативной обучающей среды на базе информационных технологий для подготовки офицеров флота.: Моногр. – Владивосток: Изд-во Дальневост. ун-та, 2001. – 360 с.
19. Локшина О. І. Моніторинг якості освіти: світовий досвід / О. І. Локшина // Педагогіка і психологія. – 2003. – № 1. – С. 108 – 116.
20. Майоров А. Н. Мониторинг как практическая система / А. Н. Майоров // [http://www.mto.ru/ children/monitoring/system.html](http://www.mto.ru/children/monitoring/system.html).
21. Майоров А. Н. Теория и практика создания тестов для системы образования / Алексей Николаевич Майоров. – М. : Интеллект-Центр, 2001. – 296 с.
22. Майоров А. Н. Тесты школьных достижений : конструирование, проведение, использование / Алексей Николаевич Майоров. – СПб : Образование и культура, 1996. – 304 с.
23. Макарова Т. Д. Тестирование в системе мониторинга качества образования // Стандарты и мониторинг в образовании. 1998, №1. – С. 60-61.
24. Малафіїк І. В. Дидактика : навч. посіб. / Іван Васильович Малафіїк. – К. : Кондор, 2009. – С. 173 – 175.
25. Машбиц Е. И. Психолого-педагогические проблемы компьютеризации обучения: (Педагогическая наука – реформе школы).– М.: Педагогика, 1988.– 192 с.
26. Михайлычев Е. А. Дидактическая тестология. – М.: Народное образование, 2001. – 432 с.
27. Моніторинг якості освіти: становлення та розвиток в Україні : Рекомендації з освітньої політики / [під заг. ред. О. І. Локшиної]. – К. : “К.І.С”, 2004. – 160 с.

28. Нардюжев В. И., Нардюжев И. В. Модели и алгоритмы информационно-вычислительной системы компьютерного тестирования. – М.: Прометей, 2000. – 148 с.
29. Нейман Ю. М., Хлебников В. А. Введение в теорию моделирования и параметризации педагогических тестов. – М.: Прометей, 2000. – 168 с.
30. Нестеров А.В., Тимченко В.В., Трапицын С.Ю. Информационные педагогические технологии. Учебно-методическое пособие. – СПб.: Издательство ООО "Книжный дом", 2003. – 340 с.
31. Осин А.В., www.rnmc.ru/ideas/ideas.html
32. Основы открытого образования. Т. 1 / Ответ. редактор В.И. Солдаткин. – РГ ИОО. – М.: 2002. – 676 с.
33. Основы открытого образования. Т. 2 / Ответ. редактор В.И. Солдаткин. – РГ ИОО. – М.: 2002. – 680 с.
34. Педагогічне оцінювання і тестування. Правила, стандарти, відповідність. Наукове видання / [Болюбаш Я. Я., Булах І. Є., Мруга М. Р., Філончук І. В.] – К. : Майстер-клас, 2007. – 272 с.
35. Поддубный А. В. Методические основы разработки и использования педагогических тестов / Поддубный А. В., Панина И. К., Ащепкова Л. Я. – Владивосток, 2003. – 296 с.
36. Равен Д. Педагогическое тестирование: Проблемы, заблуждения, перспективы / Джон Равен. – М. : «Когито-Центр», 2001. – 142 с.
37. Роберт И.В. Современные информационные технологии в образовании: дидактические проблемы; перспективы использования. – М.: "Школа–Пресс", 1994. – 205 с.
38. Рудницька О. П. Основи педагогічних досліджень / Рудницька О. П., Болгарський А. Г., Свистельнікова Т. Ю. – К. : Експрес-об'ява, 1998. – 144 с.
39. Самылкина Н. Н. Современные средства оценивания результатов обучения / Надежда Николаевна Самылкина. – М. : БИНОМ, 2007. – 172 с.

40. Софронова Н.В. Программно-методические средства в учебном процессе общеобразовательной школы. – М.: ИИО РАО, 1998. – 178 с.
41. Тесты для старшеклассников и абитуриентов. Оценка знаний / Под ред. Шмелева А. Г. – М.: «Первое сентября», 2000. – 132 с.
42. Хуторской А. В. Современная дидактика: Учебник для вузов. – СПб: Питер, 2001. – 544 с.
43. Чельшкова М. Б. Теория и практика конструирования педагогических тестов : учеб. пособ. / Марина Борисовна Чельшкова. – М. : Логос, 2002. – 432 с.
44. Шеншев Л. В. Основы адаптивного обучения языку (семиотические аспекты развития речи с помощью автомата). – М.: Наука, 1995. – 113 с.
45. Шмелев А. Г., Похилько В. И. Анализ пунктов при конструировании и применении тест-опросников: ручные и компьютерные алгоритмы. – Вопросы психологии, № 4, 1985. – С. 126-134.

Навчальне видання

КОМП'ЮТЕРНІ ТЕХНОЛОГІЇ В ТЕСТУВАННІ

Навчальний посібник

TEMPUS PROJECT
Educational Measurements
Adapted to EU Standards



Видано за підтримки програми TEMPUS IV на виконання міжнародного проекту «Освітні вимірювання, адаптовані до стандартів ЄС»

Відповідальний за випуск Гадяк Ж.В.

Надруковано з оригінал-макету, наданого авторами

Підписано до друку
Формат 60x84/16. Папір офсетний.
Друк цифровий. Гарнітура Times New Roman.
Ум.-друк.арк. . Обл.вид.арк. .
Тираж примірників.
Замовлення .

СПД Гадяк Жанна Володимирівна
Свідоцтво Держкомінформу України ДК №3585 від 22.09.09 р.
вул. Привокзальна, 12, м. Луцьк, 43007.
Тел.: (0332) 77 07 14, 77 05 02.
e-mail: vpolygraf@gmail.com
www.vpolygraf.com.ua