

## Вимірювання параметрів педагогічної моделі студента за допомогою тестів

Умовні позначення:

$\{J_c\}_i$  – множина номерів завдань, які  $i$ -й студент виконав правильно;  
 $\{J_u\}_i$  – множина номерів завдань, які  $i$ -й студент не виконав правильно;  
 $\{J_a\}_i$  – множина номерів завдань, які було запропоновано  $i$ -му студенту;  
 $\{J_c\}_j$  – множина номерів завдань, які правильно виконали  $j$ -те завдання;  
 $K$  – коефіцієнт;  
 $m$  – кількість завдань;  
 $n$  – кількість студентів;  
 $P_{теор}$  – ознака теоретичної спрямованості завдання;  
 $P_{прак}$  – ознака практичної спрямованості завдання;  
 $s$  – статистична оцінка стандартного відхилення;  
 $T$  – момент (дата і час) виконання завдання;  
 $X$  – оцінка відповіді на окреме завдання;  
 $Y$  – тестовий бал як сума оцінок за завдання;  
 $y$  – тестовий бал як середня оцінка за завдання;  
 $\beta$  – імовірність вгадування правильної відповіді;  
 $\delta$  – символ Кронекера (для будь-яких цілих величин  $i$  та  $j$ , якщо  $i = j$ , то  $\delta_{i,j} = 1$ , інакше  $\delta_{i,j} = 0$ );  
 $\tau$  – тривалість виконання завдання;

індекси:

$a$  – завдання було запропоновано студентові в процесі тестування;  
 $c$  – студент виконав завдання правильно;  
 $i$  – номер студента;  
 $j$  – номер завдання;  
 $L$  – рівень навчальних досягнень;  
 $L1$  – початковий рівень;  
 $L2$  – середній рівень;  
 $L3$  – достатній рівень;  
 $L4$  – високий рівень;  
 $u$  – студент не виконав завдання правильно;  
авт – автоматизм діяльності, навичка;  
гл – глибина знань;  
гн – гнучкість знань;  
норм – нормалізована шкала вимірювання;  
оп – оперативність знань;  
пов – повнота знань;  
прак – практична спрямованість завдання;  
сис – систематичність знань;  
теор – теоретична спрямованість завдання;

Необхідність розробки системи параметрів педагогічної моделі учня було усвідомлено ще у 60-х роках минулого сторіччя. Видатні науковці розробили систематизацію властивостей знань [1], запропонували поняття рівнів навчальних досягнень і визначили систему основних діагностичних параметрів педагогічної моделі [2]. Проблема визначення педагогічно виваженої системи параметрів і способів їх вимірювання залишається актуальною і сьогодні. Завдяки розвитку тестових технологій педагогічної діагностики і застосуванню інформаційно-комп'ютерної техніки з'являються нові перспективи вдосконалення педагогічної моделі студента і навчального процесу, практичного її застосування для визначення індивідуальних стратегій навчання кожного студента. Надії на вирішення цієї задачі пов'язані із застосуванням нового математичного апарату нечіткої логіки і класифікаційного аналізу на основі інтелектуальних систем, що «навчаються» [3]. Потрібна обґрунтована педагогічна теорія, яка б спрямувала зусилля розробників інтелектуальних систем педагогічної діагностики на реалізацію потенційних можливостей використання комп'ютерного тестування для визначення індивідуальних особливостей учнів.

Ключовим поняттям у сучасному вимірюванні навчальних досягнень є поняття **рівня засвоєння**. За означенням В.П.Беспалька [5, с. 116], рівень засвоєння – це майстерність оволодіння діяльністю як здатність до продуктивного перетворення знань для використання в нетипових умовах. Він визначає чотири рівні [5, с. 116]: упізнавання, репродуктивна діяльність, евристична діяльність, творча (дослідницька і винахідницька) діяльність. З точки зору діагностики, рівень засвоєння пов'язується з рівнем розумової діяльності, яку має здійснити учень для виконання діагностичного завдання. Слід зазначити, що рівень творчої діяльності за В.П.Беспалько передбачає здобування нових для людства відомостей і не може бути перевіреном за допомогою класичних тестів, оскільки еталона відповіді в таких завданнях ще не існує [5, с. 159]. За керівними документами Міністерства освіти і науки України [6] визначається чотири рівні навчальних досягнень: початковий, середній, достатній, високий. Для діагностики на кожному рівні можна застосовувати відповідні завдання у тестовій формі [7]: для діагностики на початковому рівні завдання передбачають репродуктивний характер діяльності учня, переважно розпізнавання; середній рівень – це завдання репродуктивного характеру, виконання яких передбачає відтворення знань, перевірку знань термінології та основних теоретичних відомостей, вміння виконати певні дії за зразком; за допомогою завдань достатнього рівня перевіряють вміння учнів застосовувати власні знання у стандартних ситуаціях за відомим алгоритмом; високий рівень відповідає

завданням проблемного характеру, застосування учнем знань у нових для нього ситуаціях. Якість оволодіння діяльністю на кожному рівні визначається за допомогою **коефіцієнта засвоєння** як частка правильно виконаних суттєвих операцій розумової діяльності відповідного рівня [5, с. 119]. Стосовно тестових технологій рекомендується враховувати поправку на можливість вгадування правильної відповіді [8]. В [9] запропоновано формулу обчислення тестового балу з врахуванням випадкового надання правильних відповідей:

$$y_i = \sum_{j=1}^m X_{i,j} / m, \text{ де } X_{i,j} = \begin{cases} 1, & \text{правильна відповідь} \\ 0, & \text{відмова від відповіді} \\ \frac{-\beta_j}{1-\beta_j}, & \text{неправильна відповідь} \end{cases} \quad (1)$$

де  $j$  – номер завдання;  $i$  – номер випробовуваного;  $m$  – кількість завдань у тесті;  $\beta_j$  – ймовірність випадкового надання правильної відповіді для  $j$ -го завдання. Практична реалізація вимірювання рівня засвоєння і коефіцієнта засвоєння здійснюється, наприклад, у автоматизованій системі педагогічної діагностики «Експерт» [10].

**Коефіцієнт усвідомленості** (за В.П.Беспалько [5, с.126]) – це відношення рівня усвідомленості знань студентом до максимально можливого: характеризує здатність студента обирати спосіб діяльності і обґрунтувати цей вибір. Рівень усвідомленості знань оцінюється за чотирибальною шкалою (від 0 до 3), він обмежений рівнем засвоєння. В [5, с.126] надається детальний опис умов, за якими здійснюється оцінювання усвідомленості: 0 – дії за алгоритмом без аргументації; 1 – аргументація вибору способу дій у межах однієї навчальної дисципліни; 2 – аргументація із застосуванням зв'язків між навчальними дисциплінами; 3 – системний рівень. Коефіцієнт усвідомленості, безумовно, є об'єктивною і діагностичною ознакою навчальних досягнень студента, але у практиці тестування складно запропонувати завдання, спрямовані на вимірювання саме цієї ознаки окремо від вимірювання рівня засвоєння. Тому коефіцієнт усвідомленості не є зручним параметром для систем автоматизованої діагностики, і врахування відповідних властивостей навчальних досягнень студента краще зробити на основі аналізу систематичності та системності знань за іншими підходами.

**Коефіцієнт навички** – відношення часу, який потрібен студенту для виконання діяльності, до часу, який потрібен для виконання цієї діяльності фахівцеві [5, с.125]. Таке вимірювання нескладно здійснити у автоматизованій системі діагностики. Діяльність студента під час тестування є переважно розумовою, характеризується підвищеною напруженістю. Це призводить до впливу на обчислені значення коефіцієнта навички додаткових факторів. По-перше, якість і швидкість відповідей студентів на завдання у тестовій формі у значній мірі залежатиме від його індивідуальної здатності до проявлення таких якостей уваги, як інтенсивність і стійкість. Процедура тестування характеризується високою напруженістю розумової праці студента. Як правило, час тестування складає від 20 до 60 хвилин. За такий проміжок часу обов'язково мають проявитися коливання уваги, які є природним захисним механізмом [11], [12, с.240]. Для визначення впливу додаткових факторів на швидкість виконання завдань у тестовій формі автором було проведено експериментальне дослідження у Харківському національному педагогічному університеті імені Г.С.Сковороди. Студентам було запропоновано завдання у тестовій формі на підрахунок кількості певних літер. Звісно, усі студенти мали достатні навички виконання цієї діяльності, тобто фактор навчальних досягнень не впливав на результати тестування. За результатами експерименту визначено, що середня тривалість виконання завдань значно відрізняється для різних студентів. Залежність якості виконання певним студентом завдань від середньої тривалості роботи над ними не є вірогідною. «Швидкі» та «повільні» студенти припускаються помилок однаково. Зменшення тривалості виконання завдання у порівнянні із середньою, характерною для конкретного студента, пов'язане з підвищенням уваги. Разом із тим, існує критична тривалість виконання завдань, нижче якої катастрофічно зменшується увага, збільшується кількість випадкових помилок. Таким чином, крім навички як результату навчальної діяльності є принаймні два фактори, які також впливають на швидкість виконання тестових завдань: темп розумової діяльності студента як властивість особистості (не впливає на якість виконання завдань) і напруженість уваги в момент виконання завдання (впливає на якість виконання завдання).

Для правильного вимірювання коефіцієнта навички (автоматизму) в умовах тестової діагностики навчальних досягнень потрібно:

- побудувати процедуру тестування так, щоб забезпечити стійкість уваги під час виконання завдань;
- застосовувати апарат математичної статистики для врахування коливань уваги;
- розробити відповідну математичну модель для врахування індивідуальних властивостей студента щодо темпу розумової діяльності.

Розглянемо властивості знань за класифікацією І.Я.Лернера [1] та їх зв'язок з параметрами, що запропоновані В.П.Беспальком.

**Повнота** знань (об'єм) – кількість елементарних фактів, зв'язків тощо відповідно до навчальної програми [1, с.14]. Повноту знань можна вимірювати як коефіцієнт засвоєння на середньому рівні навчальних досягнень. При цьому банк тестових завдань середнього рівня має будуватися так, щоб репрезентативно відобразити зміст навчального матеріалу відповідно до освітнього стандарту [13].

**Глибина** (сукупність усвідомлених суттєвих зв'язків між співвідносними знаннями [1, с.16]) і **гнучкість** (готовність до самостійного знаходження способу застосування знань у змінній ситуації або різних способів у одній і тій же ситуації [1, с.27]) відповідають високому рівню навчальних досягнень. Безпосереднє вимірювання цих властивостей знань за допомогою тестів ускладнено. Але можна ґрунтовно припустити, що вони у значній мірі пов'язані з результатами тестування за допомогою компетентнісних тестів, за якими визначається ефективність застосування знань у новій для студента

ситуації. Основою для такого припущення, наприклад, може служити висока кореляція між результатами вимірювання навчальних досягнень за спеціальними компетентнісними тестами з інформатики і результатами Всеукраїнської олімпіади з програмування [14]. При цьому гнучкість скоріше характеризує практичну розумову діяльність щодо створення нових фактів у новій ситуації, а глибина – відтворення відомих знань, але новим для студента є розуміння того, що саме ці знання відповідають запропонованій йому у завданні ситуації. Поділимо за цією ознакою завдання високого рівня на «практичні» і «теоретичні». Припустимо, що коефіцієнт засвоєння, який розраховано за «практичними» завданнями високого рівня, характеризує у деякій мірі гнучкість знань, а відповідний коефіцієнт за «теоретичними» завданнями характеризує глибину знань. Така наближена оцінка дозволить, все ж таки, спиратися на результати тестування під час аналізу глибини і гнучкості знань.

Аналогічним чином будемо розглядати коефіцієнт засвоєння, який обчислено за «практичними» завданнями середнього і достатнього рівнів, як міру **оперативності** знань (кількості ситуацій, в яких учень може завідома застосувати знання [1, с.16]).

Розрізнити систематичність (усвідомлення складу деякої сукупності знань, їх ієрархії і послідовності за певним заданим кутом зору [1, с.18]) і системність знань (усвідомлення знань за їх місцем у науковій теорії [1, с.22]) за тестовими результатами недоцільно, оскільки тестове завдання завжди передбачає порівняння відповіді з еталоном і цей еталон може бути пов'язаний з суб'єктивним поглядом під проблему за певним кутом зору. Тому розглядатиме тільки **систематичність**. В цьому зв'язку звернемо увагу на підхід до опрацювання тестових результатів, який запропоновано В.П.Пустобаєвим [15]. Структура навчального матеріалу подається у вигляді семантичної мережі. Вузли мережі відповідають елементам навчального матеріалу, а ребра – зв'язкам між ними. Мережа будується так, що матеріал вузлів нижчих рівнів спирається на матеріал вузлів-предків. Кожне завдання тесту віднесено до певного вузла. «У ситуації, коли тестований має певні знання за підрозділами (вузли-нащадки) деякого розділу (предка), але не знає змісту самого узагальнюючого розділу (предка), вся гілка результатів заповнюється нулями, оскільки уявлення учня про даний розділ відображає набір деяких незв'язних даних, а не систему знань» ([15]). Відношення тестового балу  $Y_{\text{сис}}$ , який обчислено за таким підходом, до загального тестового балу  $Y$  є кількісною мірою систематичності знань.

**Міцність** знань – стійка фіксація у пам'яті системи студента суттєвих знань і способів їх застосування або готовність вивести необхідні знання на основі інших опорних знань [1, с.22]. Природною мірою міцності є відношення відповідних коефіцієнтів засвоєння за даними попереднього і поточного тестування. Якщо в математичній моделі, що застосовується у автоматизованій системі діагностики, розглядаються параметри навчальних досягнень студента у динаміці, як функції часу, то окремий параметр «міцність знань» не потрібний. Його замінює функціональна залежність усіх інших параметрів від часу, що, безумовно, несе у собі більше відомостей.

Сьогодні не відомі способи вимірювання за допомогою тестів таких якостей знань, як усвідомленість, конкретність і узагальненість, згорнутість і розгорнутість. Для вирішення цієї задачі можуть застосовуватися різні форми педагогічного контролю: співбесіда, спостереження, виконання практичних завдань тощо.

Для здійснення окреслених вище вимірювань, для кожного завдання, яке включено до тесту, за даними апробації на репрезентативній вибірці та експертних оцінок визначаються параметри – властивості:

- $L$  – рівень завдання – набуває значень 1 (початковий), 2 (середній), 3 (достатній), 4 (високий);

- $\bar{\tau}_j = \frac{\sum_{i \in K_j} \tau_{i,j}}{\sum_{i \in K_j} 1}$  – середня тривалість правильного виконання завдання студентами;

- $s_{\tau_j} = \sqrt{\frac{\sum_{i \in K_j} (\tau_{i,j} - \bar{\tau}_j)^2}{\left(\left(\sum_{i \in K_j} 1\right) - 1\right)}}$  – оцінка стандартного відхилення тривалості виконання

завдання;

- $P_{\text{прак}}$  – ознака практичної спрямованості завдання – набуває значення 1, якщо під час виконання завдання студент відтворює способи діяльності, щось обчислює, знаходить, відносить до деякого класу, і значення 0 в інших випадках;

- $P_{\text{теор}}$  – ознака теоретичної спрямованості завдання – набуває значення 1, якщо виконання завдання пов'язано з відтворенням з відтворенням фактів і зв'язків між ними або створенням нових зв'язків між знаннями і фактами, і значення 0 в інших випадках;

- відповідність завдання вузлу семантичної мережі, що описує структуру навчального матеріалу.

Таким чином, з точки зору комп'ютерного моделювання навчальних досягнень учня, кожне завдання тесту можна розглядати як об'єкт, що має певні властивості.

За результатами тестування для кожної відповіді також формується об'єкт, який включає властивості:

- $X_{i,j}$  – оцінку правильності відповіді, яку розраховано за формулою (1) ;
- $i$  – умовний номер студента;
- $j$  – умовний номер завдання;
- $\tau_{i,j}$  – тривалість виконання завдання;
- момент (дата і час) виконання завдання для визначення динаміки параметрів як функцій часу.

Спільне опрацювання властивостей відповідей і завдання дозволяє визначити властивості об'єкта «студент» щодо його знань за результатами тестування у конкретній момент часу:

- коефіцієнти засвоєння  $K_L$ , для рівнів навчальних досягнень  $L=1, 2, 3, 4$ , визначаються як середнє арифметичне оцінок за завдання відповідного рівня;

$$K_{L,i} = \frac{1}{m_L} \sum_{j \in \mathcal{A}_i} X_{i,j} \delta_{L_j,L}, \text{ де } m_L = \sum_{j \in \mathcal{A}_i} \delta_{L_j,L}.$$

- $K_{нов,i} = K_{L2,i} = \frac{1}{m_{L2}} \sum_{j \in \mathcal{A}_i} X_{i,j} \delta_{L_j,2}$ , де  $m_{L2} = \sum_{j \in \mathcal{A}_i} \delta_{L_j,2}$ , – характеризує повноту знань;
- Коефіцієнт навички  $K_{авт,i} = \frac{\sum_{j \in \mathcal{A}_i} \tau_{норм,i,j}}{\sum_{j \in \mathcal{A}_i} 1}$  визначається як середня тривалість правильного виконання завдань, яку подано в одиницях стандартного відхилення:  $\tau_{норм,i,j} = \frac{\tau_{i,j} - \bar{\tau}_j}{s_{\tau_j}}$ .

Застосування нормалізованої шкали дозволяє порівнювати між собою відповіді на різні завдання, які, безумовно, мають різну тривалість виконання, якщо застосовувати звичайну шкалу часу. Таке визначення коефіцієнта навички не змінює його природу як показника відношення тривалості виконання завдань студентом до деякої стандартизованої тривалості, у даному випадку застосовано  $\bar{\tau}_j$ . Якщо замість  $\bar{\tau}_j$  застосувати тривалість виконання завдань експертами, то  $\tau_{норм,i,j}$  зміниться на адитивну константу, що не впливає на результати статистичного аналізу за більшістю методів. Але, як вже відмічено вище, такий підхід дискримінує студентів, які природно мають повільний темп діяльності. Тому, у випадках, коли швидкість прийняття рішення не є визначальною у фаховій підготовці, доцільно шукати альтернативні способи обчислення коефіцієнта навички. Пропонуємо обчислювати цей коефіцієнт як середнє квадратичне відхилення тривалості виконання студентом завдань у одиницях стандартного відхилення за формулою:

$$K_{авт,i} = \sqrt{\frac{1}{m-1} \sum_j (\tau_{норм,i,j} - \overline{\tau_{норм,i}})^2}, \text{ де } \overline{\tau_{норм,i}} = \frac{\sum_{j \in \mathcal{A}_i} \tau_{норм,i,j}}{\sum_{j \in \mathcal{A}_i} 1} = 0$$

Такий підхід виводить на перший план не швидкість, а рівномірність темпу роботи, що можна вважати ознакою навички. До речі, саме такий підхід застосовується при оцінюванні навички роботи на друкарській машинці. Для визначення ефективності кожного з наведених способів обчислення коефіцієнта навички потрібне додаткове експериментальне дослідження.

- $K_{зл,i} = \frac{\sum_{j \in \mathcal{A}_i} X_{i,j} P_{теор,j}}{\sum_{j \in \mathcal{A}_i} P_{теор,j}}$  характеризує глибину знань (середнє арифметичне оцінок X за «теоретичні» завдання високого рівня);
- $K_{зн,i} = \frac{\sum_{j \in \mathcal{A}_i} X_{i,j} P_{прак,j} \delta_{L_j,4}}{\sum_{j \in \mathcal{A}_i} P_{прак,j} \delta_{L_j,4}}$  характеризує гнучкість знань (середнє арифметичне оцінок X за «практичні» завдання високого рівня);
- $K_{он,i} = \frac{\sum_{j \in \mathcal{A}_i} X_{i,j} P_{прак,j} (\delta_{L_j,2} + \delta_{L_j,3})}{\sum_{j \in \mathcal{A}_i} P_{прак,j} (\delta_{L_j,2} + \delta_{L_j,3})}$  характеризує оперативність знань (середнє арифметичне оцінок X за «практичні» завдання середнього і достатнього рівнів);

• коефіцієнт систематичності знань – відношення суми оцінок за завдання, що входять до ненульових гілок семантичної мережі, до суми оцінок за всі завдання, які включено до семантичної мережі і було запропоновано студенту. Гілка семантичної мережі, яка виходить з деякого вузла, вважається нульовою, якщо середня оцінка за завдання у вузі менша деякого критичного значення. Наприклад, можна запропонувати критичне значення на рівні 0,5, як це робиться у [6] для визначення оцінок 1–3, які вважаються позитивними, але у практиці роботи навчальних закладів розглядаються як незадовільні.

Слід зазначити, що за результатами одного тестування, коли загальна кількість завдань невелика, надійність визначення деяких параметрів педагогічної моделі студента може виявитися невисокою, тому доцільно визначати такі параметри на основі систематичного моніторингу навчальних досягнень у рамках певної навчальної теми.

Об'єкт «студент» має розглядатися як динамічний, його властивості є функціями часу. Наведений список властивостей об'єкта «студент» не є повним, він має доповнюватися психолого-фізіологічними параметрами моделлю компетентностей студента, що включає його відношення до процесу навчання і власних досягнень. До загальної моделі навчального процесу обов'язково додаються моделі методів навчання і «соціального замовлення» [5].

За визначеними параметрами здійснюється класифікація студентів за рекомендованими стратегіями подальшого вдосконалення навчальних досягнень. Математичний апарат має підтримувати ймовірнісний характер вимірюваних величин і забезпечувати самонавчання інтелектуальної діагностичної системи [3].

## ЛІТЕРАТУРА

1. Лернер И.Я. Качества знаний учащихся. Какими они должны быть? – М.: «Знание», 1978. – 48 с.
2. Беспалько В.П. Основы теории педагогических систем: Проблемы и методы психолого-педагогического обеспечения технических обучающих систем. – Воронеж: Изд-во Воронеж. ун-та, 1977. – 304 с.
3. Красноясовський А.С. Інформаційний синтез інтелектуальних систем керування: підхід, що ґрунтується на методі функціонально-статистичних випробувань. – Суми: СумДУ, 2004. – 260 с.
4. Колгатін О.Г. Параметри педагогічної моделі студента. / Тези доповідей Першої міжнародної науково-технічної конференції «Інтелектуальні системи в промисловості і освіті – 2007» (Суми, 7-9 листопада 2007 року). – Суми: Видавництво СумДУ, 2007. – С. 80-81.

5. Беспалько В.П. Образование и обучение с участием компьютеров. – М.: МПСИ; Воронеж: МОДЕК, 2002. – 352 с.
6. Критерії оцінювання навчальних досягнень учнів у системі загальної середньої освіти./ За заг. ред. Віктора Огнев'юка, Олександр Савченко // Освіта України. – 2001 р. – № 6 (7 лютого 2001 р.). – С.3-16.
7. Білоусова Л.І., Колгатін О.Г. Методика обробки та інтерпретації результатів педагогічної діагностики. // Комп'ютер у школі та сім'ї. – №8, 2003. – С.28-31.
8. Кромер В.В. О некоторых вопросах тестовых технологий. // Тез. докл. Второй Всеросс. конфер. "Развитие системы тестирования в России", г. Москва 23-24 ноября 2000 г. – Ч. 4. – М: Прометей, 2000. – С. 59-61.
9. Колгатін О.Г. Статистичний аналіз тесту з різними за формою завданнями. //Засоби навчальної та науково-дослідної роботи./ За заг. ред. В.І.Євдокимова і О.М.Микитюка. – ХДПУ ім. Г.С.Сковороди. – Харків: ХДПУ, 2003. – Вип. 20. – С.50-54.
10. Білоусова Л.І., Колгатін О.Г., Колгатіна Л.С. Тестологічний аналіз у системі "Експерт". //Комп'ютер у школі та сім'ї. – №7, 2003. – С.41-43.
11. Добрынин Н.Ф. Внимание и его воспитание. – М.: «Правда», 1951. – 31 с.
12. Загальна психологія. /О.В.Скрипченко, Л.В.Долинська, З.В.Огороднійчук та ін. – К.: «А.П.Н.», 1999. – 463 с.
13. Белоусова Л.И., Колгатин А.Г., Колгатина Л.С. Принципы построения автоматизированной системы педагогической диагностики. // УСиМ. – №2, 2007. – С.75-81.
14. Білоусова Л.І., Колгатін О.Г. Завдання теоретичного туру Харківської обласної олімпіади 2004 року. // Комп'ютер у школі та сім'ї. – №6, 2005. – С.47-51.
15. Пустобаев В.П., Саяпин М.Ю. Формализация элементов диагностики знаний учащегося. // Информатика и образование. –2005. – №7. –С. 120-123.