

УДК 378.126

АЛЕКСЕЄВ О.М., КОНОВАЛОВА Н.А., ЛОЗОВА К.А.,  
ТРОФИМЕНКО П.Є.

Сумський державний університет

## ЗНАЧУЩІСТЬ І СКЛАДНІСТЬ ТЕСТОВИХ ЗАВДАНЬ ПІД ЧАС ОЦІНЮВАННЯ ДІЯЛЬНОСТІ ВИКЛАДАЧІВ ВНЗ

**Мета.** Розроблення способу кількісного визначення комплексного показника значущості та складності тестового завдання.

**Методика.** Аналіз існуючих підходів до побудови рейтингових систем, освітніх стандартів та навчальних програм для узагальнення теоретичних питань сутності, структури, функцій, особливостей застосування тестового контролю успішності навчання студентів під час оцінювання професійної діяльності викладачів ВНЗ.

**Результати.** У ході дослідження підтверджено можливість оцінювання педагогічної діяльності викладача за результатами контролю успішності навчання студентів. Розроблено спосіб безекспертного розрахунку вагомості тестового завдання, що дозволяє підвищити достовірність результатів тестового контролю успішності навчання. Запропоновано призначаючи вагомості завдання, враховувати його складність і значущість.

**Наукова новизна.** Новий спосіб кількісного оцінювання вагомості тестових завдань з урахуванням показників значущості та складності

**Практична значущість.** Використання розробленого способу розширює можливості тестування як інструменту оцінювання діяльності викладачів.

**Ключові слова:** тестовий контроль, складність тестових завдань, значущість тестових завдань, вагомості тестових завдань, якість діяльності викладачів, рейтинг.

**Вступ.** В існуючих системах управління якістю освіти найважливіше значення справедливо надають оцінюванню діяльності професорсько-викладацького складу ВНЗ [1, 2]. Не заперечуючи значущості інших показників, які визначають рейтинг вузу, відзначимо, що такий підхід виправданий, оскільки рівень надаваних освітніх послуг багато в чому визначатиметься саме тим фактором, наскільки правильно сформований професорсько-викладацький корпус. Водночас зазначимо, що в застосовуваних рейтингових системах (від національних, до конкретного вузу), як і раніше, проблемним є правильний вибір показників, використання яких дозволить оперативне оцінити успішність професійної діяльності науково-педагогічних працівників ВНЗ.

У методиках рейтингового оцінювання діяльності викладача все частіше не виправдано відмовляються від такого показника якості, як успішність студентів, досягнута під час вивчення навчальних дисциплін під його керівництвом. Співвідношення успішності студента на початкових і завершальних етапах засвоєння навчального матеріалу дисципліни, порівняння успішності з даної дисципліни із середніми оцінками студента впродовж поточної сесії або в цілому за весь час навчання у вузі, інші аналогічні індикатори могли б широко використовуватися для оцінювання діяльності викладачів. Перешкодою цьому, на наш погляд, є недосконалість методів вимірювання успішності навчання студентів у ході

проведення контрольних заходів як у процесі проведення усного іспиту, так і під час тестування.

Традиційний усний іспит, як правило, є експертною системою з одним експертом і тому не може бути об'єктивним. За сучасних ринкових відносин, коли вузи для підвищення конкурентоспроможності зменшують вартість навчання, залучення кількох викладачів-експертів обмежене і, як правило, застосовується у виняткових випадках під час прийняття рішення про можливість студенту продовжити навчання, в деяких спірних ситуаціях і тощо. Відтак використання результатів усних іспитів у системі рейтингових оцінок обмежене.

У працях [3, 4] та інших вказується на об'єктивний характер тестового контролю як інструменту педагогічних вимірювань. У той же час тестуванню властиві недоліки, які знижують можливість його застосування для оцінювання знань та вмінь студентів і через них успішності роботи викладача. До таких недоліків ми можемо віднести помилки призначення числових значень вагомості завдань, від точного визначення якого багато в чому залежить достовірність оцінки, яку отримує студент за відповідь. При необґрунтованому значенні вагомості окремих тестових завдань неправильно визначається кількість балів, нарахованих за їх виконання, і в кінцевому підсумку – сумарний бал за виконання всього тесту. Як наслідок, підсумкова оцінка студента виявиться хибною і не в повному обсязі відображатиме досягнутий рівень знань і вмінь, що були сформовані під керівництвом викладача.

**Постановка завдання.** Мета статті – проаналізувати існуючі підходи до кількісного оцінювання вагомості тестових завдань; запропонувати комплексний показник, що враховує значущість і складність тестових завдань під час оцінювання правильності відповіді студента; розробити спосіб кількісного визначення комплексного показника значущості та складності тестового завдання.

**Результати дослідження.** Аналіз науково-теоретичних і практико-орієнтованих публікацій, присвячених проблемі обґрунтованого призначення вагомості тестового завдання (у низці праць описується як ваговий коефіцієнт, індекс складності, складність, вартість та ін.), показує, що досі не існує загальноприйнятого підходу до визначення того, які властивості завдання повинен відображати цей показник, і не розроблена методика розрахунку його числового значення.

При безмашинному, в деяких випадках і при комп'ютеризованому тестовому контролі вважається, що успішність виконання будь-якого з тестових завдань, із яких складається тест, повинна характеризуватися такою самою кількістю балів. Виходячи з цього, вважаємо, що всі завдання, виконані правильно і в повному обсязі, потрібно оцінювати однаково. Для цього, на нашу думку, помилково вважати, що вагомості завдань повинні мати однакове числове значення незалежно, наприклад, від складності для засвоєння студентами або значущості для правильного розуміння контрольованого навчального матеріалу.

При комп'ютеризованому контролі, визнаючи, що таке зрівняння оцінок тестових завдань неправомірне, тим не менш, вагомості можуть диференціювати відповідно до неістотних для забезпечення точності педагогічних вимірювань класифікаційних ознак завдань. При цьому може встановлюватися зв'язок між оцінкою, що проставляється студентові під час виконання тестового завдання, і конструкцією завдання, тобто тим, у якому вигляді в завданні сформульоване питання, в якій послідовності та як перераховуються правильні й неправильні відповіді тощо [5]. Однак, на наш погляд, такий

підхід, як правило, не обґрунтований, оскільки відсутня переконлива аргументація, наприклад, чому повинні відрізнятися числові значення вагомості завдань, які відрізняються лише тим, що містять одне, два чи більше альтернативних варіантів відповідей.

Найбільш поширеним при комп'ютеризованому тестовому контролі є пов'язування вагомості зі складністю тестового завдання. При цьому складність завдання може задаватися на основі передбачуваного числа і характеру розумових операцій, необхідних для успішного виконання завдання. Однак такий підхід унаслідок обмежених можливостей оцінки розумової діяльності людини більшою мірою є теоретичним і найближчим часом не зможе широко використовуватися для обчислення рейтингу викладачів з успішності студентів.

Інший підхід полягає в тому, що складність тестового завдання визначається емпірично – за результатами калібрувального (попереднього) тестування. У нормативних документах для розроблення складових системи стандартів вищої освіти МОН України [6] рекомендовано під час використання тестування як інструменту об'єктивного контролю якості освіти випускників вузу враховувати складність тестових завдань за індексом складності  $I_c$ :

$$I_c = \frac{H + L}{n} \cdot 100 \quad (1)$$

де  $H$  – кількість правильних відповідей у сильній групі;  
 $L$  – кількість правильних відповідей у слабкій групі;  
 $n$  – загальна кількість екзаменованих в обох групах.

Відзначимо, що використовуються й інші емпіричні формули для кількісного вираження складності тестових завдань. Однак при деяких варіаціях у математичних виразах, в загальному випадку вагомість завдання визначається за співвідношенням правильних і неправильних результатів його виконання. З урахуванням того, що студентів для участі в калібрувальному оцінюванні спеціально не відбирають, результати виконання ними одного конкретного завдання не детерміновані й багато в чому випадкові, виразимо вагомість тестового завдання через відношення числа подій  $N(A)$ , в яких умова  $A$  виконана (відповідь не оцінена максимально можливою кількістю балів), до загального числа подій  $N$ :

$$P_c = \frac{N(A)}{N} \cdot 100 = \frac{\sum_{i=1}^k (B_{max} - B_i)}{k \cdot B_{max}} \cdot 100 = \left( 1 - \frac{\bar{B}}{B_{max}} \right) \cdot 100 \quad (2)$$

де  $B_i$  – результат виконання тестового завдання  $i$ -м студентам, оцінений у балах;  
 $k$  – кількість студентів, у яких результат виконання тестового завдання був оцінений;  
 $B_{max}$  – максимально можлива кількість балів за виконання тестового завдання;  
 $\bar{B}$  – середня оцінка за виконання тестового завдання, виражена в балах.

Під час оцінювання успішності виконання тесту, крім складності тестових завдань, необхідно враховувати їх значущість. Значущість тестових завдань, на наш погляд, більш вузьке поняття, ніж значущість навчального матеріалу, яке за їх допомогою контролюється. Значущість навчального матеріалу охоплює і категорії цінності, і категорії знання.

Навчальний матеріал може бути цінним для вирішення завдань, з якими матимуть справу одні студенти, і не важливим для досягнення життєвих цілей або в професійній діяльності інших студентів. Значущість може бути і гностичною (пізнавальною), і практичною (діловою), й етичною (моральною), й естетичною (художньою), і соціальною (суспільною), й емоційною (привабливою), і виховною (розвивальною) і т. д. [7].

На відміну від цього значущість тестового завдання визначатиметься насамперед тим, наскільки знання контрольованого навчального матеріалу важливе для розуміння матеріалу всієї теми, дисципліни, розкриття змісту спеціальності тощо. Відстежити такий зв'язок під час тестування дозволяє порівняння результатів виконання поточного завдання і оцінки успішності здачі тесту загалом. Зробити це можна за величиною коефіцієнта парної кореляції  $r_{xy}$ , що характеризує тісноту лінійного зв'язку між ознакою-фактором (знання матеріалу, контрольованого даними тестовим завданням, виражене кількістю отриманих балів) та ознакою-результатом (знання навчального матеріалу в цілому, оцінене балами  $B_{\Sigma}$ ):

$$P_s = r_{xy} \cdot 100 = \frac{\sum_{i=1}^k (B_i - \bar{B}) \cdot (B_{\Sigma i} - \bar{B}_{\Sigma})}{\sqrt{\sum_{i=1}^k (B_i - \bar{B})^2} \cdot \sqrt{\sum_{i=1}^k (B_{\Sigma i} - \bar{B}_{\Sigma})^2}} \cdot 100 \quad (3)$$

Вагомість тестового завдання, виражена через його значущість, не є величиною сталою і залежить від того, для контролю якого навчального матеріалу повинен використовуватися спроектований тест. Наприклад, якщо контролюється знання окремої теми дисципліни, то необхідно порівнювати результати виконання студентами поточного завдання з їхніми ж оцінками, отриманими при відповідях на це й усі інші питання тесту з цієї теми. Під час атестації випускників вузу із залученням тестування для діагностування відповідності рівня підготовки студентів кінцевим цілям освітньо-професійної підготовки вагомість завдання повинна визначатися заставленням результатів його виконання з оцінками, отриманими за виконання всіх завдань, що входять до тесту державного іспиту. Оскільки в порівнянні беруть участь різні масиви даних, то й результати обчислень за формулою 3 збігатимуться лише в деяких окремих випадках.

Для переходу від двох часткових показників складності  $P_c$  і значущості  $P_s$  тестового завдання до загального показника, за допомогою якого можна охарактеризувати узагальнений показник вагомості завдання, використовуємо методику розв'язання багатокритеріальних задач із головним критерієм [8].

Порівнюючи показники складності й значущості, головним (більш важливим) при призначенні вагомості тестового завдання вважаємо показник значущості. Складність завдання характеризується тим, наскільки важко студентові його виконати. Причиною труднощів виконання завдання можуть бути не лише недоліки роботи студента, а й погано прочитана викладачем лекція, обмеження за часом доступу до лабораторного обладнання під час виконання практичних робіт або ж неякісне методичне забезпечення занять, тобто чинники, що не залежать від його здібностей і зусиль, докладених для освоєння навчального

матеріалу. Умови навчання студентів різних академічних груп можуть різнитись і тоді їм буде потрібно докласти різних зусиль для засвоєння навчального матеріалу. У той же час значущість тестового завдання завжди визначається тим, наскільки важливе оволодіння контрольованим навчальним матеріалом для відповідності досягнутих результатів цілям і завданням вивчення окремої дисципліни або взагалі вимогам освітньо-кваліфікаційної характеристики спеціальності.

Тому під час визначення вагомості тестового завдання як і цільової функції, на нашу думку, необхідно взяти показник значущості тестового завдання, а показник складності – обмежувальним критерієм. Для визначення числових значень обмежень пропонуємо використовувати характеристики недопустимої ймовірності отримання за результатами виконання завдання переважно тільки найвищих або лише незадовільних оцінок.

Таким чином, вагомість завдань, що входять до тесту, повинна відповідати параметру значущості і його числове значення може бути визначене розрахунком за формулою 3. Завдання, під час виконання яких студенти отримують переважно погані оцінки ( $P_c = 95\%$ ), не ґрунтуються на змісті контрольованого навчального матеріалу, неточно сформульовані або містять неправильну еталонну відповідь. Такі завдання потрібно переформулювати або виключити з тесту. Також неприпустимі й завдання, під час виконання яких студенти отримують здебільшого лише максимально допустимі оцінки ( $P_c = 5\%$ ). Швидше за все, вони тривіальні і їх виконання не становить труднощів.

**Висновки.** Підвищення достовірності тестового контролю як інструменту педагогічних вимірювань розширює можливості використання його результатів для оцінювання якості професорсько-викладацького складу ВНЗ щодо успішності студентів. Цьому сприяє реалізований авторами підхід, що базується на визначенні вагомості тестових завдань за показниками складності та значущості. Перевагами такого підходу є підвищення об'єктивності запропонованого способу визначення вагомості завдань за показниками складності та значущості завдань і використання безекспертного розрахунку показника значущості. Подальше вдосконалення тестового контролю, проведення досліджень, спрямованих на збільшення точності оцінювання успішності навчання студентів, повинні допомогти правильно визначити ефективність педагогічної діяльності викладача і на цій основі вдосконалити методику розрахунку його рейтингу.

### Список використаних джерел

1. Васильева Е. Ю. Концепция системы оценки качества деятельности профессорско-преподавательского состава в вузе : дис. ... д-ра пед. наук : 13.00.08 / Васильева Е. Ю. – СПб., 2005. – 499 с.
2. Рейтинг вищих навчальних закладів III-IV рівнів акредитації М-ва освіти і науки, молоді та спорту України за 2010/2011 навчальний рік [Електронний ресурс]. – Режим доступу: <http://www.mon.gov.ua/ua/activity/education/58/rejting/>
3. Данилова С. Д. Оценивание результатов тестирования в адаптивной системе автоматизированного тестирования / С. Д. Данилова // Вестник Восточно-Сибирского государственного технологического университета. – 2008. – № 1. – С. 12-20.

4. Linn R. L. Measurement and assessment in teaching / R. L. Linn, N. E. Gronlund. – Upper Saddle River, N. J. : Merrill, 2000. – 574 p.
5. Вимоги до навчально-методичних матеріалів дистанційної форми навчання та критерії їх оцінювання. – Суми : СумДУ, 2013. – 13 с.
6. Комплекс нормативних документів для розробки складових системи стандартів вищої освіти. – К. : [б. в.], 2001. – 126 с.
7. Есина Е. В. Педагогическая психология / Е. В. Есина.– М. : Эксмо , 2008. – 158 с.
8. Зінзура В. В. Методи розв'язку задачі багатокритеріальної оптимізації регулювання напруги в електричних мережах / В. В. Зінзура // Кіровоградський національний технічний університет : зб. наук. праць. – 2012. – Вип. 25,ч. 1. – С. 350-359.

### References

1. Vasil'eva E. Ju. (2005) *Koncepcija sistemy ocenki kachestva dejatel'nosti professorsko-prepodavatel'skogo sostava v vuze : dis. ... d-ra ped. nauk : 13.00.08 [The concept of quality assessment activities of the teaching staff at the university]* / Vasil'eva E. Ju. – SPb., – 499 p.
2. *Reytynh vyshchyykh navchal'nykh zakladiv III-IV rivniv akredytatsiyi M-va osvity i nauky, molodi ta sportu Ukrainy za 2010/2011 navchal'nyy rik [Rating of higher educational institutions of III-IV accreditation of the Ministry of Education and Science, Youth and Sports of Ukraine for the 2010/2011 academic year]* [Electronic resource]. - Access: <http://www.mon.gov.ua/ua/activity/education/58/rejting/>
3. Danilova S. D. (2008) *Ocenivanie rezul'tatov testirovanija v adaptivnoj sisteme avtomatizirovannogo testirovanija [Evaluation of test results in the adaptive system of automated software testing]* / S. D. Danilova // *Vestnik Vostochno-Sibirskogo gosudarstvennogo tehnologicheskogo universiteta.* – № 1. – P. 12-20.
4. Linn R. L. (2000) Measurement and assessment in teaching / R. L. Linn, N. E. Gronlund. – Upper Saddle River, N. J. : Merrill, 2000. – 574 p.
5. *Vymohy do navchal'no-metodychnykh materialiv dystantsiynoyi formy navchannya ta kryteriyi yikh otsinyuvannya [Requirements for distance learning programs and its evaluation].* – Sumy : SumDU, 2013. – 13 p.
6. *Kompleks normatyvnykh dokumentiv dlya rozrobky skladovykh systemy standartiv vyshchoyi osvity [The higher education regulatory complex for quality standards and processes development].* – К. : [b. v.], 2001. – 126 p.
7. Esina E. V. (2008) *Pedagogicheskaja psihologija [Educational Psychology]* / E. V. Esina.– М. : Jeksmo , 2008. – 158 p.
8. Zinzura V. V. (2012) *Metody rozv'yazku zadachi bahatokryterial'noyi optyimizatsiyi rehulyuvannya napruhy v elektrychnykh merezhakh [Methods of solving of the multi-objective optimization problem for voltage regulation in electric networks]* / V. V. Zinzura // *Kirovohrads'kyy natsional'nyy tekhnichnyy universytet : zb. nauk. prats'.* – Vyp. 25,ч. 1. – P. 350-359.

## **ЗНАЧИМОСТЬ И СЛОЖНОСТЬ ТЕСТОВЫХ ЗАДАНИЙ ПРИ ОЦЕНИВАНИИ ДЕЯТЕЛЬНОСТИ ПРЕПОДАВАТЕЛЕЙ ВУЗА**

АЛЕКСЕЕВ А.Н., КОНОВАЛОВА Н.А., ЛОЗОВАЯ Е.А., ТРОФИМЕНКО П.Е.

*Сумский государственный университет*

**Цель.** Разработка способа количественного определения комплексного показателя значимости и сложности тестового задания.

**Методика.** Анализ существующих подходов к построению рейтинговых систем, образовательных стандартов и учебных программ для обобщения теоретических вопросов сущности, структуры, функций, особенностей применения тестового контроля успешности обучения студентов при оценивании профессиональной деятельности преподавателей вузов.

**Результаты.** В ходе исследования подтверждена возможность оценивания педагогической деятельности преподавателя по результатам контроля успешности обучения студентов. Разработан способ безэкспертного расчета веса тестового задания, позволяющий повысить достоверность результатов тестового контроля успешности обучения. При назначении веса задания предложено учитывать его сложность и значимость.

**Научная новизна.** Новый способ количественного оценивания веса тестовых заданий с учетом показателей значимости и сложности

**Практическая значимость.** Использование разработанного способа расширяет возможности тестирования как инструмента оценивания деятельности преподавателей

**Ключевые слова:** *тестовый контроль, сложность тестовых заданий, значимость тестовых заданий, вес тестовых заданий, качество деятельности преподавателей, рейтинг.*

## **IMPORTANCE AND COMPLEXITY OF TEST TASKS AT EVALUATION OF ACTIVITY OF UNIVERSITY PROFESSORS**

ALEKSEYEV A. N., KONOVALOVA N. A., LOZOVA S. A., TROFIMRNKO P. E.

*Sumy State University*

**Purpose.** Development of method of quantitative measurement of importance and complexity of test task complex index.

**Methodology.** Analysis of existing approaches to building of rating systems, educational standards and academic programs for generalization of theoretical issues of essence, structure, functions, peculiarities of using test control of success in students' teaching at evaluation of professional activity of university professors.

**Findings.** In the course of research the possibility of teacher's pedagogic activity evaluation according to control results of students' training success is confirmed. The method of non-expert calculation of test task weight is developed allowing to increase the certainty of results of training success test control. It is offered to consider task complexity and importance at assignment of task weight.

**Originality.** New method of quantitative evaluation of test tasks weight with account for importance and complexity indexes

**Practical value.** Using of the developed method promotes opportunities of testing as instrument of teachers' activity evaluation.

**Key words:** *test control, test tasks complexity, importance of test tasks, test tasks weight, quality of teachers' activity, rating.*