



АПРОБАЦІЯ ТЕСТОВИХ ЗАВДАНЬ ТА АНАЛІЗ ЇХ ЯКОСТІ ДЛЯ НЕСТАНДАРТИЗОВАНИХ ТЕСТІВ

Андрій Каленський¹, Олександр Ямковий²

- 1 доктор педагогічних наук, доцент, завідувач лабораторії науково-методичного супроводу підготовки фахівців у коледжах і технікумах, Інститут професійно-технічної освіти НАПН України, м. Київ, Україна
<http://orcid.org/0000-0001-9034-5042>, e-mail: kalenskyandrii@gmail.com
- 2 кандидат педагогічних наук, молодший науковий співробітник лабораторії науково-методичного супроводу підготовки фахівців у коледжах і технікумах, Інститут професійно-технічної освіти НАПН України, м. Київ, Україна,
<https://orcid.org/0000-0001-6919-3515>, e-mail: yamkovoyasasha@ukr.net

Реферат.

Актуальність. Питання підвищення якості підготовки фахових молодших бакалаврів передбачає значне поліпшення контролю результатів навчання студентів як важливого засобу управління освітнім процесом. За допомогою педагогічного контролю розкривають рівень знань випускників та отримують інформацію про стан знань студентів у освітньому процесі. Шляхом контролю здійснюється систематичний зворотний зв'язок, що дає змогу будувати адаптивну компетентісно орієнтовану програму навчання та своєчасне коригування дій викладачів і студентів у освітньому процесі. Тестовий контроль знань як метод вимірювання й оцінювання знань, умінь та навичок студента найбільш повно відповідає вимогам законодавства у сфері освіти щодо забезпечення релевантності, прозорості, надійності та об'єктивності оцінювання результатів навчання.

Мета: висвітлення методики апробації тестових завдань та аналіз їх якості для нестандартизованих тестів.

Методи: теоретичного аналізу – для дослідження існуючих підходів до методики розроблення тестів, тестових завдань та перевірки якості тестових завдань; експертної оцінки – для визначення правильності формулювання тестових завдань, точності запитань і терміну виконання тестових завдань; апробація (пілотажне дослідження) – для перевірки розробленого інструментарію; коефіцієнта кореляції Пірсона – для визначення коефіцієнта кореляції балів завдання із сумарними балами тесту.

Результати: розкрито зміст понять "тест" та "тестовий контроль", визначено відмінність між стандартизованими та нестандартизованими тестами, охарактеризовано особливості здійснення експертної оцінки ефективності тесту та його перевірки в освітньому процесі в рамках пробного тестування, розкрито алгоритм аналізу якості тесту, визначено основні показники відбору завдань для нестандартизованих тестів та їх розрахунок.

Висновки: У статті розглянуто апробацію нестандартизованих тестових завдань, їх експертну перевірку на правильність формулювання завдань, точність формулювання запитань, умови здійснення пробного тестування; критерії відбору тестових завдань: потенціал складності; дисперсію балів; коефіцієнт кореляції балів завдання із сумарними балами тесту. Проте подальшими напрямками розвідок є робота з проведення спеціального експерименту для створення стандартизованих тестів, а саме: перевірка надійності та валідності на основі репрезентативної вибірки.

Ключові слова: *тестовий контроль, тестування, тест, тестове завдання, валідність, надійність, трудність, дисперсія балів, коефіцієнт кореляції.*

Вступ. Якість підготовки фахівців – це відповідність результатів навчання у закладах фахової передвищої освіти стандартам та вимогам законодавства, а також потребам роботодавців і суспільства. Цю якість забезпечують системи зовнішнього та внутрішнього забезпечення якості фахової передвищої освіти. Одним із завдань, згідно Закону України "Про фахову передвищу освіту", внутрішньої системи забезпечення якості цієї освіти є: забезпечення релевантності, прозорості, надійності та об'єктивності оцінювання результатів навчання.

Упровадження тестового контролю у процес оцінювання знань студентів може задовільнити вищезазначені вимоги. Однак педагогічна практика свідчить, що в структурі тестів, що розробляються викладачами, домінують одновибіркові завдання, тести розробляються без перевірки на валідність, надійність тощо. Отже, можна зауважити, що необхідні конкретні методичні рекомендації щодо конструювання педагогічного тесту та тестових завдань.

Джерела. Проблема використання тестів в освітньому процесі не є новою. Питаннями тестового контролю навчальних досягнень студентів займалися багато вітчизняних та зарубіжних науковців. Зокрема, теоретико-методологічні аспекти тестового контролю висвітлені в працях В.Аванесова, І.Булах, В.Безпалька, В.Бочарнікової, Н.Гулюкіної, В.Ільїна, П.Лузана, Е.Лузік, А.Майорова, Л.Морської, І.Романюка, Л.Паращенко, І.Підласого, М.Челишкової, Дж.Фішера та інших. У цих роботах розглядаються різні форми педагогічних тестів та тестових завдань, рекомендації з їхнього використання, принципи та вимоги складання цих тестів. Однак, незважаючи на значну кількість робіт щодо тестового контролю, проблема методики тестового контролю досліджена недостатньо, зокрема відсутні праці, в яких було би обґрунтовано докладні підходи щодо конструювання різнорівневих тестових завдань, не розроблено доступну для пересічного викладача-практика методику перевірки тестових завдань і тесту в цілому на валідність.

Методи: теоретичного аналізу – для дослідження існуючих підходів до методики розроблення тестів, тестових завдань та перевірки якості тестових завдань; експертної оцінки – для визначення правильності формулювання тестових завдань та точності запитань, терміну виконання тестових завдань; апробація (пілотажне дослідження) – для перевірки розробленого інструментарію; коефіцієнта кореляції Пірсона – для визначення коефіцієнта кореляції балів завдання із сумарними балами тесту.

Метою даної статті є висвітлення методики апробації тестових завдань та аналіз їх якості для

нестандартизованих тестів.

Результати й обговорення. На сучасному етапі реформування системи фахової передвищої освіти проблема тестового контролю знань та умінь студентів набуває особливої актуальності. За визначенням А. Майорова (2001, с. 14), тест – це інструмент, що складається з кваліметрично вивірених системи тестових завдань, стандартизованої процедури проведення і заздалегідь спроектованої технології обробки і аналізу результатів, призначений для вимірювання якостей та властивостей особи, зміна яких можлива у процесі систематичного навчання. Л. Ярошук (2010, с. 51) розкриває поняття тесту як сукупності тестових завдань, які переважно потребують однозначної відповіді, що укладений за певними правилами та процедурами, передбачає попередню експериментальну перевірку і відповідає таким характеристикам ефективності, як валідність і надійність, а тестовий контроль знань – це метод вимірювання й оцінювання знань, умінь та навичок студента за допомогою спеціально підготовлених стандартизованих завдань (Лузан та Каленський, 2014, с.62).

За рівнем уніфікації розрізняють стандартизовані та нестандартні тести. Стандартизовані тести – це тести, які пройшли спеціальний експеримент з вироблення нормативів (систем критеріальних оцінок), перевірки надійності та валідності на основі репрезентативної вибірки, та які мають добре описаний інструктивно – методичний апарат (Булах, 2005; Ярошук, 2010, с. 52).

Проте в діяльності закладів фахової передвищої освіти, як правило, викладачі використовують нестандартні тести, тобто тести, складені ними самими із певних навчальних дисциплін. Це пояснюється, по-перше, тим, що кожний навчальний предмет, що викладається у закладі фахової передвищої освіти, постійно вдосконалюється, оновлюється і, відповідно, мають оновлюватися тестові завдання. По-друге, використання нестандартних тестів зумовлено такими чинниками: значна витрата часу викладачів на складання тестів та перевірка їх на валідність, трудомісткість процесу, відсутність для викладача-практика єдиної методики перевірки тестових завдань та тесту в цілому на валідність. Тому на практиці нерідко використовуються неефективні тестові завдання, які не дають можливості точно визначити рівень і якість засвоєного студентами навчального матеріалу. Таким чином, вирішення проблеми якості тестів ґрунтується на математичних розрахунках певних показників, а саме: складнощях тестового завдання, дисперсії балів, коефіцієнта кореляції балів завдання із сумарними балами тесту.

Коли комплект тестових завдань з модуля чи розділу дисципліни створений, необхідно при-

ступити до його апробації та аналізу. Апробація або, як її називає Л. Бурлачук (2006), пілотажне дослідження, проводиться із групою людей, що мають спільні особливості з тими, для кого цей тест призначений. Апробація тестових завдань має на меті:

- 1) визначення складності завдань та оцінювання їх придатності для студентів;
- 2) визначення завдань, що мають істотні недоліки;
- 3) виявлення помилок (недоліків орфографії та пунктуації тощо);
- 4) визначення часу на розв'язання завдання або тесту в цілому;
- 5) аналіз відповідей на відкриті запитання з метою можливого уточнення формулювання правильних відповідей;
- 6) виявлення недоліків в інструкціях.

Таким чином, для забезпечення об'єктивності даних щодо надійності результатів, отриманих під час тестового контролю, необхідно перевірити та підтвердити валідність тестових завдань, які будуть включені до складу тесту оцінювання результатів навчання студентів. Після складання базового набору тестових завдань появляється необхідність перевірки цих завдань на відповідність дидактичній меті, що має бути досягнута. Перевірка ефективності тесту здійснюється експертною оцінкою або порівнянням результатів тестування з оцінками інших видів атестації.

Першим кроком в апробації тесту є його експертна оцінка. Створюється група із викладачів за такими критеріями: науковий ступінь і вчене звання, стаж роботи в закладах фахової передвищої освіти з урахуванням їхнього досвіду діяльності як експертів. Перша експертиза проводиться для оцінки якості базового набору тестових завдань викладачами-експертами. Експерти-викладачі оцінюють правильність формулювання умов тестових завдань і точність запитань, хронометрують власне виконання тестових завдань. Відповіді викладачів-експертів порівнюють з еталонними, підготовленими розробниками тестових завдань. Зауваження щодо конструювання тестових завдань обговорюються спільно з укладачем та вносяться у їх зміст відповідні корективи й уточнення.

Для кінцевої перевірки умов конструювання тестових завдань проводиться друга експертиза у формі пробної перевірки із залученням групи студентів-експертів. При цьому визначається, чи розуміють студенти умови завдання, чи зрозумілий їм порядок його виконання. Крім цього, фіксується час, який вони витрачають на виконання того чи іншого завдання, та враховуються їх побажання.

Якщо експертиза визначила, що студенти не в змозі виконати тест за термін, який є в структурі занять для проведення тестової перевірки, то тест повинен бути перероблений – скорочується кількість, коригується складність завдань тощо. Якщо

попереднє обмеження в часі невідоме, витрати часу на виконання завдань з'ясовують експерти.

Після завершення роботи над першим варіантом тест перевіряється в освітньому процесі в рамках пробного тестування (пілотажне тестування). Пробне тестування є основою для збору емпіричного матеріалу про якість тесту. Статистична обробка цього матеріалу дає змогу чи розробникам оцінити тест щодо його валідності й надійності. Як вважає В. Аванесов (Аванесов, Хохлова та Потап, 2005), для отримання статистично достовірних результатів при пробному тестуванні важливо забезпечити такі умови:

1. Пробне тестування провести в декількох паралельних групах. При цьому рекомендується в кожній групі провести пробне тестування двічі, при умові, що досліджувані отримають варіанти тесту із завданнями, на які раніше не відповідали.

2. Кількість досліджуваних у групах має бути не менше 20 осіб.

3. Усі паралельні групи забезпечуються однаковими умовами (час, місце, тривалість тестування тощо).

4. Час, відведений на пробне тестування, визначається так, щоб найбільш підготовлені студенти встигли відповісти на всі запитання тесту.

5. Для отримання достовірних результатів потрібно звести до мінімуму можливість підказок серед досліджуваних.

Після пробного тестування потрібно визначитися з оцінюванням відповідей на запитання. У педагогічній літературі описані різноманітні моделі оцінювання. На наш погляд, найбільш зручною є дихотомічна модель оцінювання: за правильну відповідь нараховується один бал (студент позначив усі правильні рішення), а за неправильну – нуль балів (навіть якщо студент не відмітив хоча б одне з усіх правильних рішень, або, крім правильних, хоча б одне неправильне). Для перевірки тестових властивостей завдань результати пробного тестування необхідно подати у вигляді матриці, яку потрібно впорядкувати. У цій матриці проведено два впорядкування. Перше з них стосується студентів, які проходили тестування, а друге – це рядки з результатами випробувань, які розташовують у порядку зменшення суми балів.

У матриці результатів пробного тестування рядки з результатами випробувань розташовують у порядку зменшення суми балів, набраних студентами при тестуванні: у першому рядку – прізвище студента, який набрав найбільше балів, а в останньому – того, у кого балів найменше. Зазначений підхід використовуємо і для фіксування тестових завдань: від найлегшого завдання – до найважчого. На перше місце ставиться найлегше тестове завдання, на яке отримано найбільше правильних відповідей, на друге – менше і т.д., до останнього, в якого всього одна правильна відповідь. У нижньому рядку таблиці зазначено суму балів, набрану всіма студентами із кожного з 10 завдань. В

останньому стовбці таблиці зазначено суму балів, яку набрав кожен досліджуваний.

Матриця результатів тестування дає можливість виконати безпосередньо перший крок аналізу якості тесту. Наприклад, розглянемо матрицю результатів перевірки знань десяти студентів (табл. 1).

Основними показниками відбору тестових завдань до складу тесту є (Лузан, Каленський та Колісник, 2017): потенціал складності, дисперсія балів і коефіцієнт кореляції балів завдання із сумарними балами тесту.

Саме ці показники, як зазначають фахівці, є основними критеріями, за якими можна визначити, чи може дане завдання в тестовій формі вико-

ристовуватись у тесті, чи ні. Тому для перевірки тестових властивостей завдань тестової форми, перетворення частини з них у тестові завдання необхідно провести низку статистичних розрахунків (табл. 2).

Важливим кроком у створенні тесту є перевірка складності запропонованих завдань. Розглянемо спочатку спосіб визначення потенціалу складності завдань.

Потенціал складності в сучасній тестології вираховується як відношення кількості неправильних відповідей на завдання до кількості правильних відповідей за формулою:

$$P = q_i/p_i,$$

Таблиця 1

Приклад матриці тестових результатів групи із 10 студентів

№ пор.	Прізвища											Σ
		1	2	3	4	5	6	7	8	9	10	
1	Галета	1	1	1	1	1	1	1	1	1	0	9
2	Борисович	1	1	1	1	1	1	1	1	1	0	8
3	Варламенко	1	1	1	1	1	1	1	0	0	0	7
4	Авраменко	1	1	1	1	1	1	0	0	0	0	6
5	Дмитрук	1	1	1	1	1	0	0	0	0	0	5
6	Яненко	1	0	0	0	0	1	1	1	1	0	5
7	Захаренко	1	1	1	1	0	0	0	0	0	0	4
8	Курінна	1	1	1	0	0	0	0	0	0	0	3
9	Смашко	1	1	0	0	0	0	0	0	0	0	2
10	Донців	1	0	0	0	0	0	0	0	0	0	1
	Σ	10	8	7	6	5	5	4	3	2	0	-

Таблиця 2

Аналіз результатів тестування

Показники	Номери завдань									
	1	2	3	4	5	6	7	8	9	10
Кількість правильних відповідей	10	8	7	6	5	5	4	3	2	0
Кількість неправильних відповідей	0	2	3	4	5	5	6	7	8	10
Частка правильних відповідей p_i	1	0,8	0,7	0,6	0,5	0,5	0,4	0,3	0,2	0
Частка неправильних відповідей q_i	0	0,2	0,3	0,4	0,5	0,5	0,6	0,7	0,8	1
Потенціал складності $P = q_i/p_i$	0.00	0.25	0.43	0.67	1.00	1.00	1.5	2.33	4.0	-
Дисперсія балів $S_j^2 = q_i p_i$	0	0.16	0.21	0.24	0.25	0.25	0.24	0.21	0.16	0
коефіцієнт кореляції балів завдання із сумарними балами тесту, r_{xy}	-	0.41	0.62	0.75	0.82	0.82	0.75	0.62	0.41	-

де, P – потенціал складності, p_i – частка правильних відповідей досліджуваного, q_i – частка неправильних відповідей.

Із *таблиці 2.8* бачимо, що на перше тестове завдання дали правильну відповідь усі респонденти, а на десяте завдання, навпаки, правильно не відповіли жоден студент. Зробивши розрахунки за вищевказаною формулою, отримуємо такі результати: потенціал складності завдання №1 дорівнює – 0 (0:10), а завдання №10 не можна вирахувати (10:0). Таким чином, завдання №1 та №10 варто вилучити з переліку тестових завдань, які можуть бути використані при складанні тестів.

Другою важливою характеристикою якості тестових завдань є *дисперсія балів*, яка може слугувати показником диференціовальної здатності завдання, тобто здатності розділити групу студентів на "сильних" і "слабких". Відзначимо, що чим більша дисперсія балів, тим краща диференціовальна здатність завдання.

Для завдань, у яких використовується тільки дихотомічна оцінка (1 або 0), дисперсія визначається за формулою:

$$S_j^2 = p_j q_j,$$

де p_j , q_j – частки правильних і неправильних відповідей у кожному завданні. Варто пам'ятати, що чим більша дисперсія балів, тим краща диференціовальна здатність завдання.

Визначення коефіцієнта кореляції балів завдання із сумарними балами по всьому тесту є наступною статистичною вимогою до тестових завдань і являє собою точнішу і технологічну міру здатності завдання, що диференціює.

Для обчислення коефіцієнта кореляції найчастіше використовують коефіцієнт моментів Пірсона:

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{n\sum x^2 - (\sum x)^2} \sqrt{n\sum y^2 - (\sum y)^2}},$$

де r – коефіцієнт кореляції; x – результат по кожному завданню; y – бал (результат) по всьому тесту; n – кількість попарних добутоків; Σ – знак суми.

Для прикладу розрахуємо коефіцієнт кореляції балів по сьомому завданню (X_7). Для цього будеться допоміжна *таблиця 3*, у якій використані відповідні дані.

У першому стовпчику наводяться значення балів, отриманих студентами в сьомому завданні. Сума цих балів дорівнює 4 ($\Sigma X_7 = 4$). У другому стовпчику представлено тестові бали (Y_i), які набрав кожен досліджуваний за всі завдання, тобто сумарний бал по тесту ($\Sigma Y_i = 50$). У третьому стовпчику даються добутки балів кожного випробуваного по сьомому завданню (X_7) і по сумі балів (Y); $\Sigma X_7 \cdot Y = 29$. У четвертій і п'ятій колонках – квадрати значень X_7 і Y .

Відповідно, $\Sigma X_7^2 = 4$ і $\Sigma Y_i^2 = 310$. Для розрахунку коефіцієнта кореляції, використавши формулу Пірсона, отримуємо:

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{n\sum x^2 - (\sum x)^2} \sqrt{n\sum y^2 - (\sum y)^2}} = \frac{10 \cdot 29 - 4 \cdot 50}{\sqrt{10 \cdot 4 - 4^2} \sqrt{10 \cdot 310 - 50^2}} = 0,75$$

Ключовими вважаються завдання з коефіцієнтом кореляції більше 0,7. У таблиці 2 це завдання № 4, № 5, № 6, № 7. Зокрема, рекомендується включити в тест лише ті завдання, коефіцієнт кореляції яких дорівнює не менше 0,25 – 0,3.

Чим вище значення r_{xy} , тим більша ймовірність перетворення завдання в тестовій формі у тестове завдання, тобто можливість бути включеним у тест.

Таблиця 3

Таблиця розрахунку коефіцієнта кореляції

№ пор.	Прізвища студентів	Бал за завдання №7, X_7	Сумарний бал по тесту, Y_i	$X_7 \cdot Y_i$	X^2	Y^2
1	Галета	1	9	9	1	81
2	Борисович	1	8	8	1	64
3	Варламенко	1	7	7	1	49
4	Авраменко	0	6	0	0	36
5	Дмитрук	0	5	0	0	25
6	Яненко	1	5	5	1	25
7	Захаренко	0	4	0	0	16
8	Курінна	0	3	0	0	9
9	Смашко	0	2	0	0	4
10	Донців	0	1	0	0	1
	Σ	4	50	29	4	310

Особливо помітно ймовірність підвищується при $r_{xy} \geq 0.5$. Якщо обчислити $(r_{xy})^2 \times 100\%$, то одержимо значення так званого коефіцієнта детермінації, вираженого у зручній для інтерпретації процентній мірі зв'язку завдання із сумою балів тестованих. Для прикладу, коефіцієнт детермінації у сьомому завданні дорівнює $(0,8)^2 \cdot 100\% = 64\%$. Це можна інтерпретувати так: 64 % варіації суми тестових балів протестованих по всіх завданнях пов'язано з варіацією балів сьомого завдання. Це вказує на високу потенційну готовність сьомого завдання здійснити свій внесок у загальну дисперсію тесту.

Аналізу тестових властивостей завдання також сприяє розрахунок повної кореляційної матриці, в якій подаються кореляції кожного завдання з усіма іншими завданнями, а також кореляції із сумою балів. Цю роботу виконувати найкраще з викорис-

танням персональних комп'ютерів, за допомогою статистичних пакетів.

Після проведення експертизи та усунення виявлених недоліків базовий набір тестових завдань може бути представлений студентам для користування.

Висновки. У статті розглянуто апробацію нестандартизованих тестових завдань, їх експертну перевірку на правильність формулювання завдань, точність формулювання запитань, умови здійснення пробного тестування; критерії відбору тестових завдань: потенціал складності; дисперсію балів; коефіцієнт кореляції балів завдання із сумарними балами тесту. Подальшими напрямками розвідок є робота з проведення спеціального експерименту для створення стандартизованих тестів, а саме: перевірка надійності та валідності на основі репрезентативної вибірки.

Список посилань

Майоров, А. Н., 2001. *Теория и практика создания тестов для системы образования (Как выбирать, создавать и использовать тесты для целей образования)*. Москва: Интеллект-центр.

Ярошук, Л. Г., 2010. *Основи педагогічних вимірювань та моніторингу якості освіти* : навч. посіб. Київ : Видавничий Дім "Слово".

Лузан, П.Г. та Каленський А.А., 2014. *Методи і форми організації навчання фахівців аграрної галузі: навчальний посібник*. Київ : Видавництво "Логос".

Булах, І., 2005. *Основи педагогічного оцінювання. Частина I. Теорія: Навчально-методичні та інформаційно-довідкові матеріали для педагогічних працівників*. Київ: Майстер-клас.

Бурлачук, Л.Ф., 2006. *Психодіагностика: учеб. для вузов*. Санкт-Петербург: Питер.

Аванесов, В.С., Хохлова, Т.С., Потап, О.Е. и Плискановский, С.А., 2005. *Педагогические тесты. Методология разработки и применения: пособие для преподавателей*. Днепропетровск: Пороги.

Лузан, П.Г., Каленський, А. А. та Колісник, Н. В., 2017. *Технологія тестового контролю успішності навчання майбутніх молодших спеціалістів: методичні рекомендації*. Київ : ІПТО НАПН України.

Translation & Transliteration

Majorov, A. N., 2001. *Teoriya i praktika sozdaniya testov sistemy dlya obrazovaniya (Kak vybirat, sozdavat i ispolzovat testy dlya celej obrazovaniya) [Theory and practice of creating tests for the education system (How to select, create and use tests for educational purposes)]*. Moskva: Intellect-centr [Moscow: Intellect Center], [in Russian].

Yaroshchuk, L.H., 2010. *Osnovy pedahohichnykh vymiriuvan ta monitorynhu yakosti osvity: navch. posib. [Fundamentals of pedagogical measurements and monitoring of education quality: training manual]*. Kyiv : Vydavnychiy Dim "Slovo" [Slovo Publishing House], [in Ukrainian].

Luza, P.H. ta Kalenskyi, A.A., 2014. *Methody i formy organizatsii navchannia fakhivtsiv ahrarnoi haluzi: navchalnyi posibnyk [Methods and forms of training of specialists in the agricultural sector: training manual]*. Kyiv: Vydavnytstvo "Lohos" [Kyiv: Lohos Publishing House], [in Ukrainian].

Bulakh, I., 2005. *Osnovy pedahohichnoho otsiniuvannia. Chastyna I. Teoriia: Navchalno-metodychni ta informatsiino-dovidkovi materialy dlia pedahohichnykh pratsivnykiv [Fundamentals of pedagogical assessment. Part I. Theory: Educational methodological and information reference materials for pedagogical staff]*. Kyiv: Maister-klas [Kyiv: Maister-klas], [in Ukrainian].

Burlachuk, L.F., 2006. *Psyhodiagnostika [Psychodiagnosics] : ucheb. dlya vuzov*. Sankt-Peterburg : Piter, [in Russian].

Avanesov, V.S., Hohlova, T.S., Potap, O.E. i Pliskanovskij, S.A., 2005. *Pedagogicheskie testy. Metodologiya razrabotki i primeneniya: posobie dlya prepodavatelej [Pedagogical tests. Development and Application Methodology: Manual for teachers. Dnepropetrovsk: Porogi [Dnipropetrovsk, Porohy], [in Russian]*

Luza, P.H., Kalenskyi, A. A. ta Kolisnyk, N. V., 2017. *Tekhnolohiia testovoho kontroliu uspishnosti navchannia maibutnykh molodshykh spetsialistiv: metodychni rekomendatsii [Technology of test control of success of training of future junior specialists: methodological recommendations]*. Kyiv : IPTO NAPN Ukrainy, [in Ukrainian].

Approbation of test tasks and analysis of their quality for non-standardized tests

Andrii Kalenskyi¹, Oleksandr Yamkovyi²

1 Doctor of Pedagogical Sciences, Associate Professor, Head of laboratory of scientific and methodological support for specialist training in colleges and technical schools, Institute of Vocational Education of the National Academy of Pedagogical Sciences of Ukraine Kyiv, Ukraine

2 Candidate of Pedagogical Sciences, junior researcher of laboratory of scientific and methodological support of specialist training in colleges and technical schools, Institute of Vocational Education of the National Academy of Pedagogical Sciences of Ukraine Kyiv, Ukraine

Abstract.

Relevance. The issue of enhanced quality of professional junior bachelors training provides for a significant improvement in the control of student learning outcomes as an important means of educational process management. By means of pedagogical control, it is possible to discover the level of knowledge of graduates and obtain information about the state of students' knowledge in the educational process. The systematic feedback is provided through monitoring that enables to create an adaptive competence-oriented training program and in a timely manner correct activities of teachers and students in the educational process. Test control of knowledge as a method of measurement and assessment of knowledge, abilities and skills of a student most fully meets the requirements of legislation in the field of education to ensure relevance, transparency, reliability and objectivity of assessment of learning outcomes.

The purpose of this article is to highlight the methods of approbation of test tasks and analyze their quality for non-standardized tests.

Methods: of the theoretical analysis – to study existing approaches to the methodology of test development, test tasks and quality control of test tasks; expert evaluation – to determine the correct wording of test tasks and accuracy of the questions, time limit for fulfilling test tasks; approbation (pilot test) – to test the developed tools; Pearson correlation coefficient – to determine the correlation coefficient of task scores with the test total score.

Results: the concepts "test" and "test control" were defined, and the difference between standardized and non-standardized tests was determined, the specifics of expert evaluation of test effectiveness and its verification in the educational process in the context of trial testing was described, the algorithm for analysis of the test quality was revealed, the main indicators for selection of tasks for non-standardized tests and their calculation were singled out.

Conclusions. The article deals with approbation of non-standardized tests, their expert checks for correctness of the wording of the task, the accuracy of the wording of the questions, conditions for the implementation of trial testing; criteria for the selection of test tasks: the complexity potential; the variance of scores; the correlation coefficient of the task scores with the total test score. The further research will be focused on conducting a special experiment to create standardized tests, namely: check for reliability and validity based on a representative sample.

Keywords: test control, testing, test, test task, validity, reliability, complexity, variance of scores, correlation coefficient.

Стаття надійшла до редакції: 12.05.20

Прийнято до друку: 29.05.20