

Ярич Ігор Ярославович

Методист,
викладач інформатики та комп'ютерних дисциплін
Державного навчального закладу
«Вище професійне училище №34 м. Стрий»

Оцінювання якості та ефективності контрольних-оцінювальних матеріалів як засобу вимірювання рівня навчальних досягнень за професійною кваліфікацією “Кондитер (кондитер 4 розряду) – трудові функції А, Б, В” здобувачів освіти Державного навчального закладу “Вище професійне училище №34 м. Стрий”

Оцінювання якості та ефективності контрольних-оцінювальних матеріалів (тестових завдань) як засобу вимірювання рівня навчальних досягнень було проведено 22.02.2024 року здобувачами освіти групи КК-31 в кількості 26 (двадцять шість) здобувачів за професійною кваліфікацією “Кондитер (кондитер 4 розряду) – трудові функції А, Б, В” Державного навчального закладу “Вище професійне училище №34 м. Стрий”.

За результатами оцінювання якості та ефективності контрольних-оцінювальних матеріалів як засобу вимірювання рівня навчальних досягнень аналізуються такі параметри, як середня оцінка, медіана оцінки, мода оцінки, стандартне відхилення, значення асиметрії розподілу, ексцес розподілу, коефіцієнт внутрішньої узгодженості, співвідношення помилок, стандартна помилка, індекс легкості, індекс дискримінації:

Середня оцінка	88,87%
Медіана оцінки	90%
Мода оцінки	90%
Стандартне відхилення	9,14%
Значення асиметрії розподілу	-0,66301
Значення розподілу ексцесу	-1,337541
Коефіцієнт внутрішньої узгодженості	84,4%
Співвідношення помилок	12,8%
Стандартна помилка	6,4%
Індекс легкості	68,7%
Індекс дискримінації	56,9%

Середня оцінка - середня оцінка перших та усіх спроб.

Медіана оцінки – оцінка, що розташована посередині ранжованого ряду вибірки.

Мода оцінки – оцінка, що з'являється в усіх спробах найчастіше.

Стандартне відхилення – показник розсіювання значень випадкової величини відносно її математичного сподівання.

Значення асиметрії розподілу – числова характеристика розподілу ймовірностей дійсної випадкової величини. Цей критерій застосовують для приблизної перевірки гіпотези про нормальність емпіричного розподілу. Якщо асиметрія від'ємна, то в розподілі переважають елементи з величиною, меншою за середню оцінку. Якщо асиметрія додатна, то в розподілі переважають елементи з величиною, більшою за середню оцінку.

Значення розподілу ексцесу – міра розподілу одиниць сукупності учасників тестування за значенням ознаки (тестових балів). Коефіцієнт ексцесу характеризує «крутість», тобто даний показник показуватиме, наскільки велика різниця між кількістю здобувачів освіти для кожного значення отриманої ними оцінки. При додатному значенні дана різниця буде більшою, а при від'ємному значенні така різниця буде менша, ніж у нормальному розподілі.

Коефіцієнт внутрішньої узгодженості призначений для оцінювання внутрішньої узгодженості (окремих тестових завдань із тестом), має відображати структурну ієрархію моделі оцінювання якості та ефективності контрольно-оцінювальних матеріалів. Внутрішня узгодженість тесту – характеристика тесту, яка вказує на ступінь однорідності завдань тесту. Внутрішня узгодженість є «істотним» елементом конструктивної валідності тесту і характеризує, якою мірою завдання спрямовані на вимірювання потрібної ознаки, а також міру внеску кожного завдання в кількісну оцінку. Максимальна валідність тесту досягається за рахунок добору таких завдань, які, володіючи значущою кореляцією з результатом тесту, в той же час мінімально корелюють між собою (при невиправдано високій кореляції тест буде переобтяжений зайвими практично однозначними завданнями). Коефіцієнт внутрішньої узгодженості дорівнює 84,4%. Щоб оцінити даний коефіцієнт, потрібно перевести його у відповідність до таблички значень коефіцієнта α -Кронбаха.

Табличка значень α -Кронбаха

Коефіцієнт α -Кронбаха	Значення
$\alpha \geq 0.9$	Дуже добре
$0.9 > \alpha \geq 0.8$	Добре
$0.8 > \alpha \geq 0.7$	Прийнятно
$0.7 > \alpha \geq 0.6$	Сумнівно
$0.6 > \alpha \geq 0.5$	Погано
$0.5 > \alpha$	Недостатньо

Значення коефіцієнта α -Кронбаха при оцінювання якості та ефективності контрольно-оцінювальних матеріалів як засобу вимірювання рівня навчальних досягнень дорівнює 0,844. Порівнявши отримане значення із табличними ($0,9 > \alpha \geq 0,8$), можна зробити висновок, що внутрішня узгодженість тесту є доброю.

Співвідношення помилок – розбіжність значень досліджуваної ознаки визначається за кількома параметрами. Перший вказує на те, наскільки одні здобувачі освіти кращі, ніж інші, другий – деяка випадкова варіація. Співвідношення помилок показує, скільки є випадкових варіацій. Співвідношення помилок – 12,8%. Чим менший цей показник, тим кращий тест.

Стандартна помилка – це стандартне відхилення оцінок, які будуть отримані за багаторазової випадкової вибірки даного обсягу з однієї і тієї ж сукупності; це спадна функція обсягу вибірки. Стандартна помилка – 6,4%. Чим менша стандартна помилка, тим достовірнішою є отримана оцінка. Прийнятною вважається помилка величиною у 5%.

Індекс легкості вказує, наскільки конкретне тестове завдання є легким/важким для тих, хто його виконує. Вираховується як співвідношення середнього отриманого бала за завдання до максимально можливого бала. Щоб оцінити даний індекс легкості, потрібно перевести його у відповідність до таблички значень критеріїв індексу легкості завдань.

Табличка значень критеріїв індексу легкості тестових завдань

Значення індексу легкості (Pdiff)	Висновок
$Pdiff \leq 0,20$	Вгадування
$0,20 < Pdiff \leq 0,36$	Надто складне
$0,36 < Pdiff \leq 0,84$	Середньої складності
$Pdiff > 0,84$	Надто легке

Значення індексу легкості при оцінювання якості та ефективності контрольно-оцінювальних матеріалів як засобу вимірювання рівня навчальних досягнень дорівнює 0,687. Порівнявши отримане значення із табличними ($0,36 < Pdiff \leq 0,84$), можна зробити висновок, що тести є середньої складності.

Також визначалася оцінка навмання – оцінка, яку б міг отримати здобувач освіти, якби вгадував правильний варіант.

Індекс дискримінації – параметр, який характеризує, наскільки ефективно тестове завдання здатне розрізнити здобувачів освіти з різним рівнем підготовки. Під час обчислення даного індексу здобувачі освіти розподіляються на три групи з найвищим, середнім та найнижчим рівнем підготовки, проте порівнюються тільки здобувачі освіти сильної та слабкої груп. Даний показник набуває значень від -1 до 1. При значенні 1 усі учасники

сильної групи відповіли правильно, а учасники слабкої групи неправильно. Якщо значення близьке до 0 або від'ємне – це означає, що учасники слабкої групи відповідають на завдання краще, ніж із сильної, або про те, що завдання слабо диференціює здобувачів освіти з різними рівнями підготовки. Щоб оцінити даний індекс дискримінації, потрібно перевести його у відповідність до таблички значень критеріїв індексу дискримінації тестових завдань.

Табличка значень критеріїв індексу дискримінації тестових завдань

Значення індексу дискримінації (D_j)	Висновок
$D_j \geq 0,4$	Завдання функціонує задовільно
$0,3 < D_j \leq 0,39$	Потрібна невелика корекція
$0,2 < D_j \leq 0,29$	Завдання потрібно переглянути
$D_j \leq 0,19$	Завдання потрібно вилучити або повністю переробити
$D_j = 0$ або $D_j < 0$	Завдання потрібно вилучити

Значення індексу дискримінації при оцінювання якості та ефективності контрольних-оцінювальних матеріалів як засобу вимірювання рівня навчальних досягнень дорівнює 0,569. Порівнявши отримане значення із табличними ($D_j \geq 0,4$), можна зробити висновок, що завдання в тестах функціонують задовільно.

За допомогою педагогічного експерименту перевірялась якість та ефективність розробленого навчально-методичного комплексу контрольних-оцінювальних матеріалів для досягнення статистично значимого зростання рівня навчальних досягнень здобувачів освіти з професійної кваліфікації. Виміряти даний показник було вирішено за допомогою проведення тестування. Для того, щоб отримані в ході експерименту результати можна було вважати достовірними та адекватними, була проведена перевірка якості розроблених тестів у цілому, так і окремих тестових завдань. Тести проводились у формі електронного комп'ютерного тестування. Це дозволило забезпечити незалежну оцінку навчальних досягнень здобувачів освіти, автоматизацію перевірки знань, оцінити ефективність організації навчального процесу.

Ефективним засобом організації та проведення комп'ютерних тестувань є система Moodle. Дана система електронного навчання є безкоштовною та вільнорозповсюджуваною, простою в користуванні та має широкі можливості для організації та контролю навчання здобувачів освіти. Була створена база завдань на відповідність (відповідь на кожне із кількох завдань вибирається із списку можливих) та множинний вибір (дозволяє вибрати один або декілька варіантів із заданого списку). При проходженні здобувачем освіти тесту із бази випадковим чином вибиралися завдання, які розміщувались у випадковому порядку. Тестування проводилось у кабінеті

інформаційних технологій під наглядом викладача та було обмежене в часі. Оскільки дане тестування виконувало лише контрольну функцію і мало на меті оцінити рівень навчальних досягнень здобувачів освіти – для тестування їм була надана одна спроба, обмежена в часі. Крім того, після тестування здобувачу освіти був доступний для перегляду лише його кінцевий бал без розбору його правильних і неправильних відповідей. Перевірка якості тесту дозволяє виправити його недоліки, змінити певні запитання та часові рамки проходження тесту тощо.

Використані джерела

1. Кухар Л.О., Сергієнко В.П. Конструювання тестів. Курс лекцій: навч. посіб. / Л.О. Кухар, В.П. Сергієнко. – Луцьк, 2010. – 182 с.
2. Сергієнко В.П. Використання вбудованої системи аналізу тестових завдань в LCMS Moodle / В.П. Сергієнко, Л.О. Кухар, О.В. Галицький, П.В. Микитенко // Інформаційні технології і засоби навчання. – 2014. – Т. 41, Вип. 3. – С. 196-208.