

УДК 37.091.26:004.9

Диховичний Олександр Олександрович

доцент, кандидат фізико-математичних наук, доцент кафедри математичного аналізу та теорії ймовірностей

Національний технічний університет України «Київський політехнічний інститут», м. Київ, Україна
*a.dux@mail.ru***Дудко Анна Федорівна**

асистент кафедри математичного аналізу та теорії ймовірностей

Національний технічний університет України «Київський політехнічний інститут», м. Київ, Україна
*afdudko@gmail.com***ОЦІНЮВАННЯ ЗМІСТОВОЇ ВАЛІДНОСТІ ТЕСТІВ З ВИЩОЇ МАТЕМАТИКИ
З ВИКОРИСТАННЯМ GOOGLE DOCS ДОДАТКУ**

Анотація. У статті досліджено питання аналізу змістової валідності комп'ютерних тестів з вищої математики. Розглянуто технологію оцінки якості змісту тесту з використанням Google Docs додатку, яка передбачає організацію і проведення експертизи із встановлення відповідності між тестовими завданнями і змістовою галуззю тесту і приймання узагальнюючих висновків розробником. Експертиза включає три напрями роботи експертів: оцінювання завдань тесту, оцінювання тесту у цілому та оформлення узагальнюючих висновків і рекомендацій. Відзначено основні особливості та переваги використання Google Docs додатку як інструменту проведення експертизи змісту тесту.

Ключові слова: валідність тесту; змістова валідність; експертиза якості змісту тесту; специфікація тесту; Google Docs додаток; тестування з вищої математики.

1. ВВЕДЕННЯ

Постановка проблеми. У сучасних умовах реформування освітньої галузі актуальною є проблема покращення контролю знань студентів. У зв'язку з упровадженням в освітній процес інформаційно-комп'ютерних технологій (ІКТ), які відповідають викликам сучасного інформаційного суспільства, особливого значення набуває комп'ютерне тестування. Так в НТУУ «КПІ» кафедрою математичного аналізу та теорії ймовірностей організовано комп'ютерне тестування з вищої математики з використанням створеного викладачами кафедри комплексу дистанційних курсів «Вища математика» [7]. Вибір тестування як засобу оцінювання навчальних досягнень студентів вимагає вирішення низки питань, пов'язаних із процесом розробки тестів. Однією з найвагоміших задач є задача оцінювання змістової валідності тестів. Оцінювання змістової валідності, у свою чергу, породжує питання організації експертизи і вибору інструментарію для її проведення. Використання ІКТ для проведення експертизи дає можливість підвищити ефективність роботи експертів. Тому набувають значущості задачі вибору засобу ІКТ як інструменту проведення експертизи і розробки технології оцінювання змістової валідності тесту з його використанням.

Аналіз актуальних досліджень і публікацій. Аналізу валідності педагогічного тесту присвячені роботи з проблем педагогічних вимірювань А. Анастасі, В. С. Аванесова, К. Інгенкампа, А. С. Казарінова, Т. Кука, А. Н. Майорова, О. О. Михайличева, А. І. Саміловського та інших дослідників. Пошук адекватних засобів забезпечення змістовою валідністю тестів у тестології проводився В. А. Аванесовим, В. Ю. Биковим, Ю. М. Богачковим, Ю. О. Жуком, А. Н. Майоровим, Е. А. Михайличевим, С. А. Сафонцевим М. Б. Челишковою, В. А. Хлебніковим. У руслі педагогічної психодіагностики аналогічні проблеми розроблялися за кордоном

провідними фахівцями: Дж. Алгіною, А. Анастазі, Р. Готсданкером, К. Ингенкампом, П. Клайном, Д. Кемпбеллом, Л. Крокер. Автори одноставні в тому, що оцінка змістової валідності тесту зазвичай проводиться незалежними експертами з великим стажем роботи, які не брали участь в розробці тесту. Експертиза якості змісту тесту обов'язково має включати аналіз якості змісту окремих тестових завдань. Аналізуючи роботи вищезазначених фахівців, можна зробити висновок, що методи забезпечення змістової валідності базуються на аналізі змісту завдань у термінах вимог до рівня підготовки іспитників або до цілей тестування; оцінюванні рівня значущості змісту завдань тесту; оцінюванні відповідності змісту тесту до його специфікації; виявленні невдалих завдань тесту; обчисленні відсотка правильно виконаних завдань гіпотетичною групою іспитників.

Серед методів проведення експертизи, описаних зокрема в роботі [6], можна виділити такі:

- очне опитування;
 - ✓ вільне інтерв'ю експертів;
 - ✓ анкетне опитування експертів;
 - ✓ «Мозковий штурм»;
- заочне опитування;
 - ✓ поштове анкетне опитування експертів;
 - ✓ дельфійська техніка.

Обрана нами дельфійська техніка полягає у багаторазовому поштовому анкетуванні однієї і тієї ж групи експертів. Після першого опитування експертів і обробки його результатів, підсумки повідомляються учасникам експертної групи. Вони повинні або підтвердити свою точку зору, висловлену на попередньому етапі, і якщо вона значно відрізняється від думки більшості, розгорнуто її мотивувати, або змінити свою оцінку відповідно до думок більшості учасників. Потім знову обробляються висновки експертів, результати знову розсилаються експертам і так до тих пір, поки не будуть отримані узгоджені оцінки експертів.

Також досліджуючи методи проведення експертизи якості тесту, можна виділити обов'язкові етапи проведення експертизи.

1. Організаційний. Формується група експертів.
2. Підготовчий. Відбувається ознайомлення експертної комісії з тестом, його спеціалізацією, структурою та завданнями.
3. Експертне оцінювання тестових завдань. На цьому етапі експерти самостійно працюють із завданнями тесту, заповнюючи бланк оцінювання. Обговорення експертами завдань при цьому не допускається.

Узагальнення експертних оцінок. Після того, як кожен експерт визначився зі своїм рішенням, розробник тесту приймає узагальнюючі висновки щодо подальшої роботи з тестом.

Метою статті є розробити технологію оцінювання змістової валідності із застосуванням Google Docs додатку.

2. РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Наведемо декілька визначень валідності тесту.

Валідність тесту – це поняття, яке вказує нам, що тест вимірює і наскільки добре він це робить (А. Анастазі) [2].

Валідність означає придатність тестових результатів для тієї мети, заради чого проводилося тестування (В. Аванесов) [1].

Валідність – це характеристика здатності тесту служити поставленій меті

вимірювання (М. Челишкова) [8].

Валідність визначає, наскільки тест відображає те, що він повинен оцінювати (А. Майоров) [5].

Змістова валідність визначається як характеристика репрезентативності змісту тесту по відношенню до запланованих для перевірки знань й умінь. Найбільше значення цей вид валідності має для тестів досягнень (тестів успішності, екзаменаційних тестів).

Для забезпечення високої змістової валідності необхідно розв'язати такі задачі:

- розробити специфікацію тесту;
- відібрати групу експертів, компетентних у змістовій сфері;
- оцінити змістову валідність створеного тесту.

Розробка специфікації тесту. Специфікація тесту в розгорнутій формі має включати наведені пункти.

1. Введення, у якому вказують тему тесту, для кого і для чого призначений тест, яким стандартам і програмам підготовки він відповідає, які навчальні матеріали будуть покладені в основу тесту.
2. ПІБ автора специфікації.
3. Дата створення специфікації.
4. Предмет (галузь), до якої відноситься тест.
5. Мета створення тесту.
6. Перелік можливих рішень (суджень), що можуть бути прийняті за результатами тестування
7. Мова тесту.
8. Загальна кількість завдань тесту. Кількість завдань тесту різних форм із зазначенням кількості варіантів відповідей завдань закритої форми.
9. Час на виконання тесту.
10. Форми проведення тестування (бланкова, комп'ютерна, змішана тощо)
11. Організаційні умови проведення тестування.
12. Момент проведення тестування (наприкінці вивчення теми, проміжний, підсумковий тощо)
13. Спосіб аналізу результатів тестування.
14. Форма і структура представлення результатів.
15. Зміст (перелік тем), що виноситься на тест.
16. Детальний перелік і опис підтем тесту.
17. Матриця тесту, у якій наводиться відсоток тестових завдань, що мають бути включені до конкретної підтеми тесту.
18. Список вимог до рівня підготовленості іспитників.
19. Відсоток вимог до рівня підготовленості іспитників, охоплених специфікацією.

Відбір групи експертів, компетентних у змістовій галузі. Оцінка якості змісту тесту зазвичай проводиться незалежними експертами, які не брали участь у розробці тесту. Як правило, число експертів складає не менше трьох осіб по кожному тесту. До експертизи розробники залучають найдосвідченіших викладачів, які мають великий стаж роботи з тими студентами, для яких призначений тест. Основними процедурами процесу вибору експертів є складання списку можливих експертів і вибір з них експертної комісії відповідно до компетентності кандидатів. Детально ці процедури описані в роботі [6]. Для кількісної оцінки рівня компетентності використовується коефіцієнт компетентності, з урахуванням якого зважуються думки експертів. Цей коефіцієнт визначається за апіорними й апостеріорними даними. У разі використання апіорних даних оцінка коефіцієнта компетентності здійснюється до проведення

експертизи на основі самооцінки експерта і взаємної оцінки інших експертів. У разі використання апостеріорних даних оцінка коефіцієнта компетентності здійснюється на основі обробки результатів експертизи. Ми використовуємо апіорні дані для вибору експертів на основі взаємної оцінки інших експертів.

Оцінювання змістової валідності створеного тесту. Технологія оцінювання змістової валідності створеного тесту передбачає:

- організацію і проведення експертизи з установами відповідності між тестовими завданнями і змістовою галуззю виконання тесту.
- приймання узагальнюючих висновків розробником щодо поліпшення змісту тесту.

Організація експертизи. Правильна організації експертизи є однією з найвагоміших задач розробників тесту, оскільки від неї залежить ефективність роботи експертів. Як інструмент проведення експертизи тестів з вищої математики нами був обраний Google Docs додаток, використання якого має низку переваг:

- швидкі й ефективні збір і оброблення даних;
- мінімальні витрати;
- можливість проведення експертизи незалежно від часу і відстані;
- зручність спільної роботи з документами;
- можливість обговорення і коментування в он-лайн режимі.

За 2–3 дні до початку роботи кожен експерт повинен ознайомитися зі специфікацією рецензованого тесту, що містить пояснення щодо його структури і запланованого до перевірки змісту. Для цього створюється файл у Google Docs, у якому розміщується текст специфікації (список вимог до рівня підготовленості іспитників додається як додаток на окремій сторінці в кінці документа) і текст тесту. Доступ до документа відкривається всім експертам.

Проведення експертизи. Технологія експертизи якості змісту тесту зазвичай включає три напрямки роботи експертів:

- оцінювання завдань тесту;
- оцінювання тесту в цілому;
- оформлення узагальнюючих висновків і рекомендацій щодо поліпшення змісту тесту.

Оцінювання завдань тесту. Робота експерта по першому напрямку полягає в аналізі змісту окремих завдань тесту. Для цього експерт повинен заповнити табл. 1. Таблиця розміщується розробником тесту у файлі Google Docs. Для кожного експерта створюється окремий файл експертизи, до якого йому відкривається доступ. Не рекомендуємо створювати єдиний файл з дублюванням таблиці для всіх експертів, оскільки можливість бачити коментарі інших експертів впливає на вираження об'єктивної думки.

У таблиці перший рядок містить номери завдань тесту.

У другому рядку для закритих завдань тестів наводяться номери правильних відповідей, обраних експертом.

Відповіді до відкритих завдань наводяться експертом у третьому рядку таблиці. Тут необхідно звернути увагу розробників тесту на можливі випадки неоднозначності, коли виникають додаткові, частково правильні відповіді.

Таблиця 1

Результати експертизи змісту завдань

1.	№ завдання	1	2	3	...	L	Σ
2.	№ правильної відповіді в закритих завданнях						
3.	Правильна відповідь завдань з відкритою						

	відповіддю						
4.	№ вимоги до рівня підготовленості						
5.	Рівень базовості						
6.	Значущість змісту завдання						
7.	Очікуваний відсоток виконання іспитниками з задовільною підготовкою						
8.	Очікуваний відсоток виконання іспитниками вибірки						
9.	Очікуваний час виконання завдання, хв.						
10.	Завдання, зміст яких не відповідає специфікації в цілому						
11.	Завдання, зміст яких не відповідає підтемі специфікації						
12.	Завдання, умова яких є нечіткою і неоднозначною						
13.	Завдання, в умові яких міститься не вся інформація для надання правильної відповіді						
14.	Завдання, в умові яких міститься інформація, яка дає відповідь на інші завдання тесту						
15.	Завдання, у яких зустрічаються неправильно складені дистрактори						
16.	Інші невдалі завдання						
17.	Коментарії про якість завдання						

У четвертому рядку таблиці наводяться результати аналізу змісту завдань у термінах вимог до рівня підготовки іспитників. Для проведення аналізу експерту необхідно зіставити зміст кожного завдання тесту і перелік пронумерованих вимог. За результатами зіставлення експерт визначає номер вимоги, на яку переважно орієнтований зміст даного завдання, і виставляє цей номер у четвертому рядку для кожного завдання тесту. Тут у процесі аналізу можливі ситуації, коли в окремих завданнях контролюється не одне, а два або більше вимог до рівня підготовки іспитників. У цьому випадку експерт наводить тільки номер тієї вимоги, яка в основному перевіряється за допомогою даного завдання тесту.

У п'ятому рядку таблиці проставляється рівень завдання на основі зіставлення змісту завдань з вимогами до рівня підготовки іспитників. Завдання розділяють на три групи: Б (базовий), П (підвищений), Д (досконалий). У тому випадку, коли завдання перевіряє ступінь досягнення вимог на мінімальному рівні, достатньому для виставлення оцінки «задовільно», воно вважається відповідним групі Б. До групи П відносяться завдання, правильне виконання яких дозволяє виставити студентів оцінки «добре» і «відмінно». Отже, завдання групи П за складністю перевищують мінімально достатній рівень, але знаходяться в межах мінімуму змісту освіти з даного предмету. До групи Д включаються завдання, що виходять за межі мінімуму змісту і вимог до рівня підготовки іспитників.

У шостому рядку таблиці експерт оцінює рівень значущості змісту кожного завдання тесту. Оцінка значущості пов'язана з необхідністю включення до тесту лише тих елементів змісту, які є найважливішими, ключовими для освоєння навчального курсу в межах вимог до рівня підготовки.

Оцінку значущості змісту завдань експерти виставляють за чотирибальною шкалою. Бал 0 виставляється в тому випадку, коли експерт вважає невиправданим внесення даного завдання в тест. Бал 1 виставляється в тому разі, якщо зміст завдання

представляється експерту незначним, 2 – значущим, а 3 – найважливішим опорним елементом. Отримані оцінки завдань тесту необхідно усереднити, обчисливши сумарну оцінку завдань тесту:

$$\mu = \frac{\sum_{j=1}^L Z_j}{L}$$

де Z_j – бал j -ого завдання, L – кількість завдань в тесті. Сумарна оцінка наводиться в комірці, яка знаходиться на перетині останнього стовпчика і п'ятого рядка.

Для заповнення сьомого рядка експерту необхідно представити гіпотетичний випадок, коли завдання з вибором відповіді виконувала вибірка студентів із задовільним рівнем підготовки. Утім, необхідно уявляти порівняно невелику вибірку. Наприклад, можна уявити групу «трієчників» всього з 10 чоловік і потім підрахувати гіпотетичний відсоток студентів, які правильно виконали завдання тесту.

У восьмому рядку експерт також наводить передбачуваний відсоток іспитників, які правильно виконали завдання тесту, орієнтуючись у своїх гіпотетичних оцінках не на групу «трієчників», а на вибірку, де рівномірно представлені іспитники з різними рівнями знань.

Дев'ятий рядок містить очікуваний час виконання студентом кожного завдання тесту. Під час вибору часу експерту слід орієнтуватися на студента із середньою підготовкою. Перше уявлення про час виконання завдань експерт може скласти в процесі своєї роботи над тестом. Тому, приступаючи до виконання завдань, експерту потрібно підготувати все необхідне: годинник, бланк з номерами завдань, і строго фіксувати початок і кінець роботи над кожним завданням тесту. Отриманий час необхідно скорегувати з урахуванням того, що завдання виконував викладач, а не студент. Для корекції зафіксований час виконання кожного завдання зазвичай збільшують в 4–5 разів. Отриманий результат (у хвиликах) слід записати в графу для очікуваного часу виконання учнем кожного завдання тесту.

У десятому й одинадцятому рядках експерт відповідно проставляє зірочки в комірках тих завдань, зміст яких не відповідає специфікації в цілому або обраній підтемі.

У дванадцятому рядку зірочками відмічаються ті завдання, умова яких є нечіткою і неоднозначною, так що всі іспитники працюють над різним завданням, у тринадцятому – ті, в умові яких міститься не вся інформація для надання правильної відповіді.

У чотирнадцятому рядку експерт відмічає завдання, в умові яких міститься інформація, яка дає відповідь на інші завдання тесту (у дужках зазначає, на які саме).

У п'ятнадцятому рядку експерт проставляє зірочки напроти тих завдань, у яких зустрічаються неправильно складені дистрактори. До таких завдань можуть належати такі:

- у яких немає логічної і граматичної відповідності між умовою і кожним дистрактором;
- у яких один або більше дистракторів є неправдоподібними;
- у яких є дистрактори, які можуть бути прийняті іспитниками за правильну відповідь;
- у яких іспитник може віднайти правильну відповідь логічним шляхом, маючи досвід тестування;
- у яких довжина запропонованих відповідей різної довжини.

Позначки експертів, виставлені в 10–15 рядках таблиці, дозволяють виявити невдало складені завдання.

У шістнадцятому рядку експерт відмічає інші невдалі завдання тесту. До таких слід віднести, наприклад, відкриті завдання з короткою відповіддю, у яких можливі випадки появи неоднозначності, коли на місці планованої єдиної відповіді можна привести синоніми, різні числа або більш-менш правильні відповіді.

Усі пояснення щодо виявлення невдалих завдань і додаткові зауваження з якості змісту експерт повинен привести в сімнадцятому рядку. Якщо експерт вважає за потрібне дати розгорнуті рекомендації щодо поліпшення змісту завдань, то їх слід привести на окремих сторінці файлу експертизи із заголовком «Додаток 1. Зауваження по якості змісту завдань». У цьому ж рядку експерт характеризує якість формулювань завдань тесту, якість представлення графічної інформації, якщо вона є в завданнях тесту. Оцінюючи якість формулювань, експерту слід пам'ятати про те, що тестові завдання повинні бути сформульовані чітко, точно, лаконічно і коректно з точки зору загальноприйнятої в навчальному предметі термінології і символіки предмета.

Оцінювання тесту у цілому. Другий напрямок роботи експерта пов'язаний з аналізом якості змісту всього тесту.

Під час роботи за другим напрямом слід мати на увазі, що зміст тесту визначається як оптимальне відображення вимог до рівня підготовки іспитників у системі завдань тесту. Експерт, використовуючи дані четвертого рядка таблиці з рис. 1, повинен обчислити відсоток вимог, охоплених тестуванням:

$$\nu = \frac{s}{q} \cdot 100\% ,$$

де s – кількість вимог, охоплених тестом, q – загальна кількість вимог. Отриманий експертом відсоток охоплення порівнюється з наведеними в специфікації тесту. Потім обчислюється міра відхилення у вигляді різниці відсотків, яка заноситься в табл. 2, яка розміщується на окремій сторінці файлу експертизи із заголовком «Додаток 2. Відхилення відсотка охоплення вимог до рівня підготовки іспитників». Тут же для зручності попередньо записують відсоток вимог, охоплених усім тестом, заявлений у специфікації, і відсоток вимог ν , отриманий експертом.

Таблиця 2

Відсоток вимог, охоплених всіма варіантами тесту, заявлений у специфікації	Відсоток вимог, охоплених всіма варіантами тесту, отриманий експертом	Відхилення відсотка охоплення вимог до рівня підготовки іспитників

Експерт має оцінити правильність пропорцій змісту тексту. Необхідно перевірити, чи всі важливі аспекти предметної галузі і в правильній пропорції охоплює тест. Найчастіше під час розробки тесту можливе зміщення пропорцій, оскільки тест легко перенаситити тими розділами змісту, за якими легше скласти завдання. Для оцінки правильності пропорцій тесту експерт підраховує відсотки завдань у тесті, орієнтованих на матеріал кожної підтеми, які разом із своїм баченням оптимального співвідношення розділів наводить на окремій сторінці файлу експертизи із заголовком «Додаток 3. Оцінка правильності пропорцій тесту і перевірка відповідності змісту специфікації». Тут же записують різницю, що характеризує відхилення думки розробників від експертних оцінок

Експерт перевіряє відповідність змісту тесту його специфікації. Для цього експерт обчислює загальну кількість завдань, які не відповідають специфікації тесту, використовуючи дані 10 та 11 рядків таблиці 1. і записує це число у Додатку 3.

Використовуючи дані 9 рядка таблиці 1 експерт обчислює очікуваний час

виконання всього тесту і записує його в комірку, яка знаходиться на перетині 9 рядка і останнього стовпця.

Оформлення узагальнюючих висновків і рекомендацій щодо поліпшення змісту тесту. Третій напрям роботи експерта полягає в оформленні узагальнюючих висновків і рекомендацій щодо покращення змісту тесту. На окремій сторінці файлу експертизи із заголовком «Висновки та рекомендації щодо поліпшення змісту тесту» експерт наводить свій загальний висновок про зміст тесту і рекомендації щодо його покращення. При цьому обов'язково необхідно:

- зробити висновки щодо правильності пропорцій тесту і відповідності змісту специфікації;
- оцінити значущість змісту завдань і тесту в цілому;
- оцінити, на який рівень підготовки іспитників орієнтований тест;
- навести рекомендації щодо коректування часу виконання тесту;
- навести рекомендації щодо виявлених невдалих завдань тесту;
- виставити оцінку змістової валідності від 0 до 1, враховуючи, наступні інтерпретації її значень: 0,7 –1 – тест володіє високою змістовою валідністю; 0,4 –0,6 – необхідна деяка корекція тесту і специфікації; 0–0,3 – необхідно докорінно змінити структуру тесту.

Приймання узагальнюючих висновків розробником щодо поліпшення змісту тесту. На основі результатів експертизи розробник перевіряє узгодженість думок експертів і робить узагальнюючі висновки щодо якості змісту тесту. Узгодженість висновків експертів перевіряється на основі обчислення коефіцієнта рангової конкордації Кендалла-Сміта [6]. У випадку, коли висновки експертів не є узгодженими, то відповідно до методу Дельфі експертиза повторюється. Коли ж висновки експертів узгоджені, то усереднюючи думки всіх експертів, наприклад, за методом Джона Кемені [6], розробник приймає рішення щодо подальшої роботи з тестом:

- залишити тест без змін;
- необхідна деяка корекція тестових завдань або специфікації тесту;
- необхідно докорінно змінити структуру тесту, переглянути специфікацію і провести повторну експертизу якості тесту.

3. ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

1. Науково обґрунтований відбір змісту є найважливішою передумовою досягнення високої якості педагогічного тесту.
2. Оцінка змістовної валідності тесту передбачає проведення експертизи змісту тесту незалежними компетентними у змістовній області експертами.
3. Оцінка змістовної валідності тесту передбачає оцінювання змісту окремих тестових завдань, тесту в цілому й оформленні узагальнюючих висновків і рекомендацій щодо поліпшення змісту тесту.
4. Як інструмент проведення експертизи змісту тесту з вищої математики цілком логічним є вибір Google Docs додатку.
5. Вибір Google Docs як інструменту проведення експертизи якості змісту обґрунтований перевагами: швидкими й ефективними збором і обробкою даних, можливістю проведення експертизи незалежно від часу і відстані, мінімальністю витрат, можливістю зручної спільної роботи з документами, ведення обговорення й коментування в он-лайн режимі.
6. Пробна експертиза якості змісту тесту з вищої математики підтвердила ефективність використання Google Docs для оцінювання змістової валідності.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Аванесов В. С. Композиция тестовых заданий. Учебная книга для преподавателей вузов, учителей школ, аспирантов и студентов педвузов. 2 изд., испр. и доп. / В. С. Аванесов – М.: Адепт 1998. – 217 с.
2. Анастаси А. Психологическое тестирование / А. Анастаси, С. Урбина. – СПб. : Питер, 2006. – 688 с.
3. Биков В. Ю. Моніторинг рівня навчальних досягнень з використанням Інтернет-технологій: монографія / В. Ю. Биков, Ю. М. Богачков, Ю. О. Жук; за ред. В. Ю. Бикова, чл.-кор. АПН України, д. тех. наук, проф.; Ю. О. Жука, канд. пед. наук, доц. – К. : Педагогічна думка, 2008. – 128 с.
4. Клайн П. Введение в психометрическое проектирование. Справочное руководство по конструированию тестов. / П. Клайн. – Киев : ПАН Лтд, 1994. – 184 с.
5. Майоров А. Н. Теория и практика создания тестов для системы образования. / А. Н. Майоров. — М. : «Интеллект-центр», 2001. – 296 с.
6. Орлов А. И. Экспертные оценки. Учебное пособие. / А. И. Орлов. – М.: 2002. –31 с.
7. Про розвиток та досвід експлуатації комплексу дистанційної освіти «Вища математика» / [П. В. Алексеева, В. О. Гайдей, О. О. Диховичний та ін.] // Дидактика математики: проблеми і дослідження : міжнар. зб. наук. робіт. – Вип. 31. – Донецьк : Вид-во ДонНТУ, 2009. – С.49–56.
8. Чельшкова М. Б. Теория и практика конструирования педагогических тестов / М. Б. Чельшкова. – М. : Логос, 2002. – 431 с.

Матеріал надійшов до редакції 30.03.2016 р

ОЦЕНИВАНИЕ СОДЕРЖАТЕЛЬНОЙ ВАЛИДНОСТИ ТЕСТА ПО ВЫСШЕЙ МАТЕМАТИКЕ С ИСПОЛЬЗОВАНИЕМ GOOGLE DOCS ПРИЛОЖЕНИЯ

Дыховичный Александр Александрович

доцент, кандидат физико-математических наук, доцент кафедры математического анализа и теории вероятностей

Национальный технический университет Украины «Киевский политехнический институт»,

г. Киев, Украина

a.dux@mail.ru

Дудко Анна Фёдоровна

ассистент кафедры математического анализа и теории вероятностей

Национальный технический университет Украины «Киевский политехнический институт»,

г. Киев, Украина

afdudko@gmail.com

Аннотация. В статье исследован вопрос анализа содержательной валидности компьютерных тестов по высшей математике. Рассмотрена технология оценивания качества содержания теста с использованием Google Docs приложения, которая предусматривает организацию и проведение экспертизы по установлению соответствия между тестовыми заданиями и содержательной областью теста, а также приема обобщающих выводов составителем. Экспертиза включает три направления работы экспертов: оценивание заданий теста, оценивание теста в целом и оформление обобщающих выводов и рекомендаций. Отмечены основные особенности и преимущества использования Google Docs приложения как инструмента проведения экспертизы содержания теста.

Ключевые слова: валидность теста; содержательная валидность; экспертиза качества содержания теста; спецификация теста; Google Docs приложение; тестирование по высшей математике.

EVALUATION OF CONTENT VALIDITY OF TESTS IN HIGHER MATHEMATICS USING GOOGLE DOCS APPLICATIONS

Oleksandr O. Dykhovychnyi

Associate professor, PhD (physical and mathematical sciences), associate professor of Mathematical Analysis and Probability Theory Department

National Technical University of Ukraine “Kiev Polytechnic Institute”, Kyiv, Ukraine

a.dyx@mail.ru

Anna F. Dudko

assistant of Mathematical Analysis and Probability Theory Department

National Technical University of Ukraine “Kiev Polytechnic Institute”, Kyiv, Ukraine

afdudko@gmail.com

Abstract. The present paper studies the issue of content validity analysis of computer tests in higher mathematics. The evaluation technology of quality of test content using Google Docs is given. It provides for organization and realization of examination to establish coincidence between test items and test content, making decisions by test compiler. The examination consists of three guidelines of experts: analysis of test items, test analysis and drawing up of general conclusions and recommendations. The main features and advantages of using Google Docs app as a tool for the examination of the test content are noted.

Keywords: validity of a test; content validity; quality examination of a test content; test specification; Google Docs; testing in higher mathematics.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Avanesov V. S. The composition of test items. Uchebnaja kniga dlja prepodavatelej vuzov, uchitelej shkol, aspirantov i studentov pedvuzov. 2 izd., ispr. i dop. / V. S. Avanesov . – M. : Adept 1998. – 217 s. (in Russian)
2. Anastazi A. Psychological testing / A. Anastazi, S. Urbina. – Spb.: Piter, 2006. – 688 s. (in Russian)
3. Bykov V. Ju. Monitoring of learning achievements using Internet technologies: Monograph / V. Ju. Bykov, Ju. M. Boghachkov, Ju. O. Zhuk; za red. V. Ju. Bykova, chl.-kor. APN Ukrainy, d. tekhn. nauk, prof.; Ju. O. Zhuka, kand. ped. nauk, doc. – K.: Pedagogichna dumka, 2008. – 128 s. (in Ukrainian)
4. Klajn P. Introduction to psychometric designing. Manual for tests design. / P. Klajn. – Kiev : PAN Ltd, 1994. – 184 s. (in Russian)
5. Majorov A. N. Theory and practice of test construction for the educational system. / A. N. Majorov. – M. : «Intellect-centr», 2001. – 296 s. (in Russian)
6. Orlov A.I. Expert assessments. Manual. / A. I. Orlov. – M. : 2002. – 31 s. (in Russian)
7. On the development and operation experience of set of distance education “Higher Mathematics” / [I. V. Alekseeva, V. O. Gaydey, O. O. Dykhovychnyi etc.] // Didaktika matematiki: problemi i doslidzhennya : mizhnar. zb. nauk. robit.. — Issue 31. – Donetsk: Vid-vo DonNTU, 2009. – S.49–56. (in Ukrainian)
8. Chelyshkova M. B. Theory and practice of constructing of pedagogical tests / M. B. Chelyshkova. – M. : Logos, 2002. – 431 s. (in Russian).

Conflict of interest. The author has declared no conflict of interest.



This work is licensed under Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.