

УДК 378.14:004.9

Петков Олександр Олександрович, кандидат технічних наук, старший науковий співробітник, провідний науковий співробітник Науково-дослідного та проектно-конструкторського інституту "Молнія" Національного технічного університету "Харківський політехнічний інститут", доцент кафедри "Інженерна електрофізика" Національного технічного університету "Харківський політехнічний інститут", м. Харків, e-mail: apetkov@yandex.ru.

ДІАГНОСТИЧНІ ХАРАКТЕРИСТИКИ КОМП'ЮТЕРНИХ ТЕСТІВ, ЯКІ ФОРМУЮТЬСЯ ЗА МЕТОДОМ ВІДНОВЛЮВАНИХ ФРАГМЕНТІВ

Анотація

У статті розглянуто визначення валідності й надійності тестів, які формуються за методом відновлюваних фрагментів. Проаналізовано структуру контрольованого теоретичного матеріалу обмеженої галузі знань, мовних виразів, які описують предмети контролю, та надійності тесту. Наведено методику визначення найістотніших складових надійності розглянутих тестів: надійності кількісного визначення коефіцієнта засвоєння і технологічної надійності. Результати проведених педагогічних експериментів довели, що тести даного класу дозволяють з високою надійністю провадити контроль засвоєння теоретичного матеріалу на рівні відтворення в будь-якій галузі знань. Показано, що валідність тестів з відновлюваними фрагментами в основному обумовлюється ступенем структурування і методичною проробкою контрольованого матеріалу і може сягати наперед заданих показників, аж до рівня абсолютної валідності.

Ключові слова: комп'ютеризований контроль знань, діагностика, валідність тесту, надійність тесту.

Постановка проблеми. Вивчення будь-якої галузі знань включає оволодіння його поняттєво-категоріальним апаратом (ПКА), що є сукупністю поглядів цієї галузі на сутність і змістовні характеристики досліджуваних явищ і процесів. ПКА включає фундаментальні поняття, істотні зв'язки й відносини предметів і явищ у розглянутій галузі знань [1–3].

З погляду навчального процесу ПКА, у першу чергу, є теоретичним матеріалом, який підлягає вивченню учнями й студентами до виконання практичних завдань. Неповне володіння ПКА навчального курсу приводить до труднощів у практичному використанні професійних знань.

Цим визначається обов'язковість першочергового контролю ступеню засвоєння ПКА в професійній освіті будь-якого рівня.

У теперішній час такий контроль найефективніше здійснюється з використанням комп'ютерних тестів. Ефективність застосування тестів у процесі контролю визначається за їх відповідністю низці показників – діагностичних характеристик, які мають бути визначені до широкого впровадження тесту в навчальний процес.

Аналіз останніх досліджень і публікацій. У загальному випадку кожний тест описується такими показниками [4]: валідність, надійність, репрезентативність, об'єктивність, вірогідність й ефективність. У [5] зазначено, що стосовно оцінки якості тестів навчальних досягнень особистості, вони традиційно описуються основними показниками – надійністю й валідністю.

Однак, поняття "валідність тесту" по-різному використовується в психометрії [6]. Як узагальнення можна використовувати таке визначення валідності [7]: "Валидность теста определяет, насколько тест отражает то, что он должен оценивать, и выражается устанавливаемой экспертами степенью (высокая, средняя, низкая). Для оценки валидности педагогического теста эксперты отвечают на следующие вопросы. Насколько адекватно тестовые задания отражают смысл измеряемого явления? Насколько точно результаты тестирования согласуются с известными объективными критериями? Насколько хорошо результаты тестирования могут быть объяснены в терминах педагогических характеристик?"

Багатогранність валідності тексту відбита й у [8], де наведені показники: валідність тесту; валідність тесту критеріальна; валідність тесту прогностична; валідність тесту змістовна; валідність порівняльна; валідність тесту факторна.

Незважаючи на неоднозначність поняття "валідність тесту", у [5] запропонована його кількісна оцінка – *критеріальний коефіцієнт валідності*, який визначається виходячи з індивідуальних числових оцінок респондентів за виконання

тесту й числових оцінок цих же респондентів, виставлених експертами за той же контрольований матеріал незалежно від процедури тестування.

Менший ступінь неоднозначності визначення має показник "надійність тесту". Відповідно до [8], надійність тесту – це показник точності й усталеності результатів вимірювання за допомогою тесту за його багаторазового застосування. Надійність характеризує ступінь адекватності відображення тестом відповідної генеральної сукупності завдань. У [7] визначення надійності тесту більш конкретизоване й визначається як міра того, наскільки точно даний тест вимірює досліджуване явище, його усталеність до дії завад типу стану випробуваних, їхніх відносин до процедури тестування, випадковим помилкам і т. д. Джерело [5] указує на те, що в психометрії поняття "надійність тесту" має два значення. У першому значенні надійність тесту визначається узгодженням його частин. Друге значення пов'язане з повторним тестуванням респондентів. Залежно від розглянутого значення поняття "надійність тесту" існують різні методи визначення її кількісної оцінки – *коефіцієнта надійності* [5].

Незважаючи на всі складності визначення якості тестів, їх валідність і надійність мають бути визначені для кожного тесту, застосовуваного в навчальному процесі. У протилежному випадку, як відзначено в [9], "Від розмаїття підходів може піти кругом голова, бо широке використання тестових технологій – це дійсно інша культура освітнього процесу." Останнім часом ця розмаїтість тестів усе більше структурується як за типом тестів, так і галузям їхнього застосування, наприклад [9]. Одним із варіантів тестів, використовуваних для підсумкового контролю ступеня засвоєння теоретичного матеріалу обмеженої галузі знань, є тести з відновлюваними фрагментами (ТВФ) [10]. Поряд із широкими можливостями ТВФ з перевірки засвоєння теоретичного матеріалу на рівні відтворення [11], у теперішній час відсутні рекомендації з оцінки якості тестів даного типу.

Як видно із проведеного аналізу, значна спільність і неоднозначність визначення валідності й надійності вимагає конкретизації визначення їхньої кількісної оцінки для різних видів тестів.

Метою даної статті є дослідження діагностичних характеристик комп'ютерних тестів, які створюються з використанням методу відновлюваних фрагментів.

Виклад основного матеріалу дослідження. ПКА будь-якої галузі знань, поряд з термінологічною системою [10] включає (див. рис. 1): описи формул (у тому числі й словесний опис), формулювання аксіом, теорем і наслідків з них, законів і правил, які в загальному випадку є мовним виразом, що обов'язково містить набір відмітних (базових) термінів й опис (конкретизацію) взаємозв'язку між ними. За допомогою аналогічних мовних виразів у технічних галузях знань також формулюється опис множини об'єктів типу: технічні пристрої, їхні складові й елементи, прилади, інструменти.



Рис. 1. Структура ПКА обмеженої галузі знань

Для цілей контролю знань на даному етапі розвитку комп'ютерної техніки мовний вираз повинен бути представлений в текстовому форматі. Тоді, структурна модель перерахованих вище мовних виразів, які відповідають предметам контролю знань, аналогічна структурній моделі визначення терміна [10]: мовний вираз структурно складається із сукупності логічно пов'язаних базових термінів, конкретизуючих і допоміжних фрагментів (див. рис. 2а). Конкретизуючий фрагмент є текстовим фрагментом контрольованого мовного виразу, який конкретизує зв'язок базових термінів між собою. Допоміжний фрагмент – це текстовий фрагмент

контрольованого мовного виразу, який забезпечує стилістичну цілісність виразу, його "читаність".

Зазначимо, що вищесказане характерне для будь-якої галузі знань (див. приклади в додатку).

Виходячи з описаної вище структури, знати мовний вираз, який відповідає предмету контролю, означає – знати вхідні в нього базові терміни й конкретизуючі фрагменти. Отже, предметом діагностики засвоєння мовного виразу є встановлення респондентом зв'язків типу: "аналізований мовний вираз – базовий термін" і "аналізований мовний вираз – конкретизуючий фрагмент".

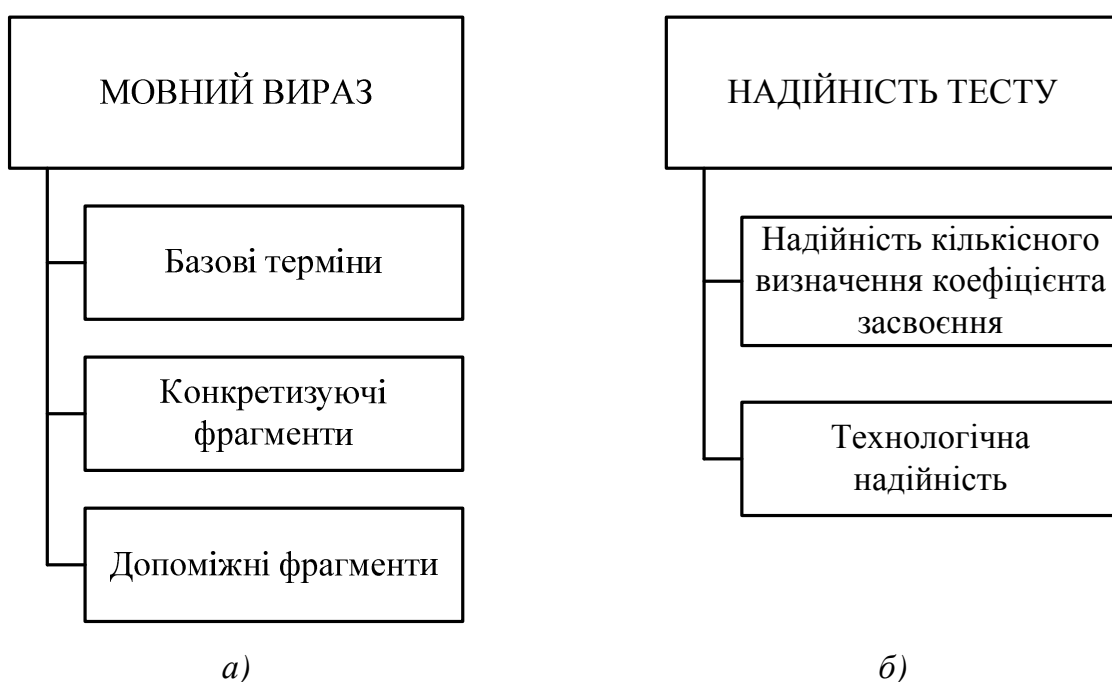


Рис. 2. Структурні моделі: а – мовного виразу; б – надійності тесту

Це дозволяє для контролю засвоєння формул, аксіом, теорем, законів, правил і призначення об'єктів використовувати метод відновлюваних фрагментів, описаний у [10]. Даний метод був реалізований у комп'ютерній програмі-оболонці [12].

Виходячи зі структури, показаної на рис. 1, ПКА обмеженої галузі знань (навчального курсу) є кінцевою множиною мовних виразів – M . Предметом контролю ТВФ є також кінцева множина мовних виразів – L (причому $L \in M$), склад якої визначається обсягом теоретичного матеріалу контрольованого курсу (його частини, розділу, окремої теми). Тому експертна оцінка валідності полягає у визначенні того, яку частину охоплює контрольна множина L від теоретичного матеріалу,

задекларованого в меті тесту, обумовленого множиною M . Тоді *коефіцієнт валідності* можна обчислити як

$$k_v = \frac{n}{N}, \quad (1)$$

де N – кількість мовних виразів, необхідних для вивчення за заявленої мети тесту; n – кількість мовних виразів, що контролюються в тесті.

З огляду на те, що $n \leq N$, значення коефіцієнта валідності змінюється у інтервалі $0 < k_v \leq 1$. Зіставлення значення коефіцієнта валідності k_v , визначеного за (1), зі ступенем, установлюваним експертами (високий, середній, низький), рекомендованого в [7], виходить за рамки об'єктивності й далі не обговорюється.

Особливо зазначимо, що за досить повної методичної розробки навчальної документації (детально пророблене питання про те, яким обсягом знань повинен володіти студент після вивчення навчального курсу) і контрольних питань (методично чітко пророблені формулювання всіх складових обсягу знань, зазначеного в меті), ТВФ можна вважати апріорі валідними, тому що в цьому випадку досягається 100 % охоплення завданнями тесту заявленого в меті контрольного матеріалу.

Надійність ТВФ, базуючись на визначеннях, наведених у [7, 8], можна розглядати як міру того, що результат тестування (оцінка за виконання тесту) сформований навчальною діяльністю [10], а не факторами, які безумовно впливають на процес контролю, але не стосуються безпосередньо процесу засвоєння знань, наприклад, ступенем володіння комп'ютером, індивідуальною реакцією респондента на співвідношення між загальним світловим фоном і колірними подразниками й т. п.

Як видно з рис. 2б, надійність ТВФ має два основні, найбільш істотні складові: надійність кількісного визначення коефіцієнта засвоєння й технологічна надійність.

Надійність кількісного визначення коефіцієнта засвоєння є мірою суттєвості відмінностей в значеннях коефіцієнта засвоєння, визначених програмою-екзаменатором і людиною-екзаменатором. Якщо відмінності не суттєві, то тест вважається надійним, а середній коефіцієнт засвоєння, визначений програмою, приймається відповідним рівню знань респондента.

Надійність кількісного визначення коефіцієнта засвоєння для ТВФ можна перевірити шляхом зіставлення результатів комп'ютерного тестування з результатами ідеальної моделі контролю – символічного відтворення респондентом контрольованого виразу [10]. За статистичного порівняння результатів вони не мають

значимо відрізнятись. Зіставлення результатів двох варіантів контролю може бути проведене з використанням відомої методики порівняння двох середніх нормальних генеральних сукупностей, дисперсії яких невідомі й однакові (незалежні вибірки) [13] або спеціально розроблених методик обробки даних у психології [14].

На базі програми-оболонки [12] автором розроблені такі тести: тест №1 "Введення в планування експерименту", тест №2 "Повний й дробовий факторні експерименти" (за курсом "Планування наукових експериментів") і тест №3 "Устрій і робота генераторів імпульсних напруг" [15] (за курсом "Розробка електрофізичних установок для випробувань і випромінювання високої щільності"). Кожний тест містить 30 завдань, які включають розглянуті вище предмети контролю. Програма-оболонка формує протокол контролю, у якому відбиті коефіцієнти засвоєння матеріалу кожного завдання тесту.

Для перевірки надійності кількісного визначення коефіцієнта засвоєння матеріалу тестів №1 і №2 був проведений педагогічний експеримент, суть якого полягала в такому. Респондентам після виконання комп'ютерного тесту пропонувалося письмово відтворити відповіді на окремі завдання тесту, які вибиралися випадково. Результати комп'ютерного тестування й письмового опитування на ті самі завдання тестів №1 і №2 зведені в таблицю 1.

Таблиця 1

Результати комп'ютерного тестування й письмового опитування

| Номер вибірки, i (Вид контролю) | Вибіркове середнє коефіцієнта засвоєння, \bar{x}_i | Виправлена дисперсія коефіцієнта засвоєння, s_i^2 | Кількість завдань (обсяг вибірки), n_i |
|--------------------------------------|---|--|--|
| 1 (Комп'ютерний тест) | 0,8587 | 0,044 | 21 |
| 2 (Письмове опитування) | 0,7883 | 0,089 | 21 |

Постановка задачі порівняння результатів. Є дві вибірки з нормальних генеральних сукупностей відповідей на тести №1 і №2: X_1 — вибірка результатів (коефіцієнтів засвоєння) комп'ютерного тестування; X_2 – вибірка результатів (коефіцієнтів засвоєння) письмового опитування. Обсяг вибірок n_1 і n_2 відповідно. Вибіркові середні рівні \bar{x}_1 й \bar{x}_2 , а виправлені дисперсії – s_1^2 і s_2^2 . При рівні значимості

$\alpha = 0,05$ перевіряється нульова гіпотеза $H_0: M(X_1) = M(X_2)$ – математичне очікування коефіцієнта засвоєння за комп'ютерного тестування дорівнює математичному очікуванню коефіцієнта засвоєння за письмового опитування, за конкуруючої гіпотези $H_1: M(X_1) \neq M(X_2)$.

Перевірка істинності гіпотези H_0 за методикою, наведеною в [13], показала, що немає підстав відкинути нульову гіпотезу H_0 , тобто немає підстави вважати, що вибіркові середні \bar{x}_1 й \bar{x}_2 , наведені в таблиці 1, відрізняються значимо. Іншими словами, результати комп'ютерного тестування й письмового опитування збігаються, що дає підставу говорити про надійність кількісного визначення коефіцієнта засвоєння тестів №1 і №2. Ступінь надійності даного висновку визначається ймовірністю того, що даний результат є істинним $p = 1 - \alpha$. Отже, для проведеного експерименту надійність кількісного визначення коефіцієнта засвоєння за допомогою програми-екзаменатора становить $p_1 = 0,95$.

Технологічна надійність є мірою можливості "обману" програми-екзаменатора, тобто відповідає на запитання – чи можна, використовуючи доступні респондентів засоби програми, одержати позитивну (залікову) оцінку, не знаючи у відповідному обсязі контрольований матеріал?

Технологічна надійність ТВФ визначається з урахуванням шкали оцінювання, обґрунтування якої наведене в [10]. Заліковим значенням коефіцієнта засвоєння є величина $K_3 = 0,75$. Єдиним доступним для респондента засобом модифікації стану програми-оболонки [12] є використання маніпуляторів (миша, клавіші управління курсором, клавіші Tab і Enter), які забезпечують вибір запропонованих програмою словосполучень. Тоді "обман" програми-екзаменатора може бути проведений тільки шляхом вибору правильних словосполучень, використовуючи випадковий вибір (модифікацію коду програми внаслідок її "злому" виключаємо). Для перевірки можливості появи такої події для кожного тесту 50 разів було проведено контрольне тестування з фіксацією середнього й максимального значень коефіцієнта засвоєння K_3 по всіх експериментах. На рис. 3 показані результати педагогічних експериментів з випадковим вибором варіантів відповіді.

Як видно з рисунка, успішно здати тести за випадкового вибору практично неможливо.

Загальна кількість спроб здачі тестів $n = 150$ і жодна з них не привела до позитивного результату (досягненню залікового значення коефіцієнта засвоєння). Це означає, що подія, коли випадковий вибір варіантів відповіді дозволяє здати тест, малоїмовірна. Тоді, використовуючи відоме розв'язання задачі про знаходження верхньої границі ймовірності події, що не відбулася в n випробуваннях [16], визначимо ймовірність події, яка полягає в тому, що випадковий вибір не дозволяє успішно здати тест, як

$$p_2 = \sqrt[n]{1-\beta}, \quad (2)$$

де $\beta = 0,99$ – довірна ймовірність події.

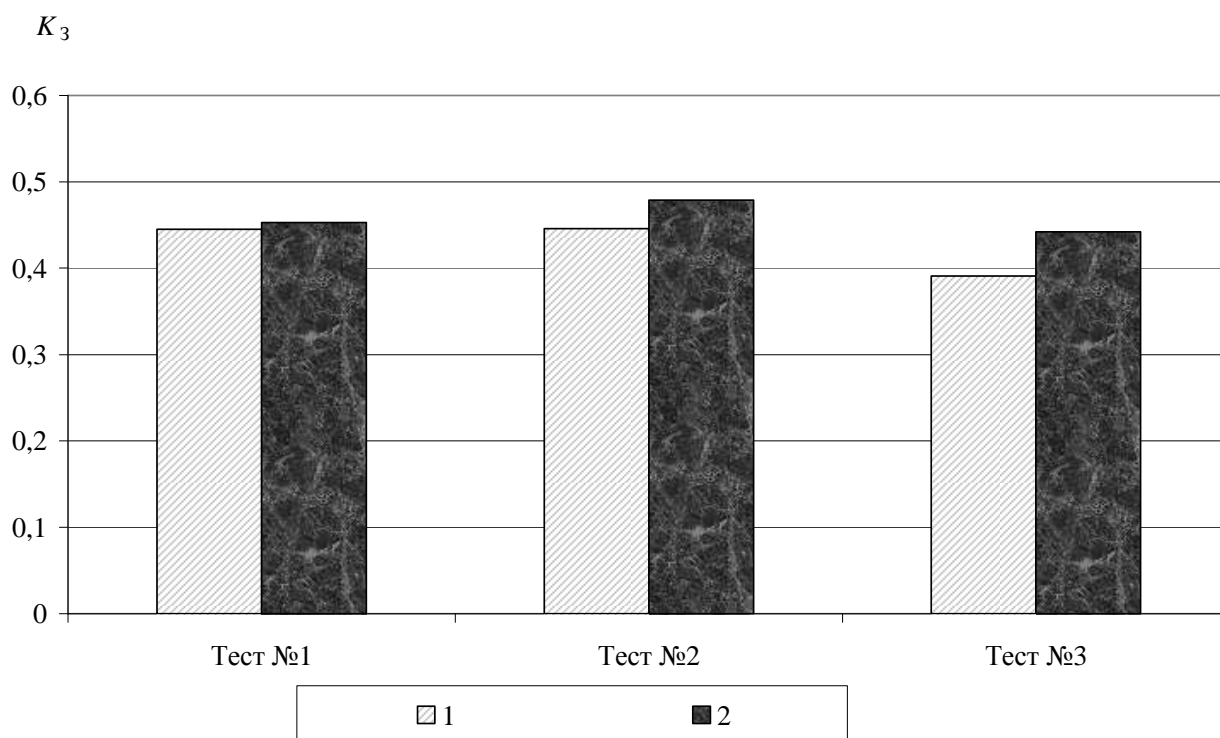


Рис. Коефіцієнти засвоєння при випадковому виборі варіантів відповіді:

1 – середнє значення коефіцієнта засвоєння;

2 – максимальне значення коефіцієнта засвоєння

Для наших умов розрахунків по (2) дає таке значення $p_2 = 0,97$. Отриманий результат показує, що у випадку повної відсутності контрольованих знань у групі з

100 респондентів, максимально тільки 3 з них зможуть досягти в процесі тестового контролю залікового значення коефіцієнта засвоєння.

У [10] на прикладі тесту №1 показано, що ТВФ неможливо здати успішно не тільки за випадкового вибору варіантів відповіді, але й спираючись на загальну ерудицію, логічні паралелі, інтуїцію, мимовільне запам'ятовування. Тому отримане значення технологічної надійності p_2 є її нижньою оцінкою.

Тоді нижня оцінка надійності ТВФ у цілому визначається як $p = p_1 \cdot p_2$. Для даних проведеного експерименту $p = 0,92$.

Отже, можна констатувати, що в разі проведення кваліфікованої розробки методичних матеріалів під час створення ТВФ, останні відповідають критеріям валідності й надійності.

Висновки.

1. Тести, розроблені на базі методу відновлюваних фрагментів, можуть застосовуватися для контролю засвоєння теоретичного матеріалу на рівні відтворення в будь-якій галузі знань.
2. Запропоновано методики перевірки надійності й валідності тестів для контролю засвоєння поняттєво-категоріального апарату.
3. Структуризація поняттєво-категоріального апарату контрольованої обмеженої галузі знань і достатня методична розробка контрольованих матеріалів забезпечують абсолютну валідність тестів з відновлюваними фрагментами.
4. Експериментально доведено, що тести, розроблені з використаннями методу відновлюваних фрагментів, відповідають критерію надійності.

Наведені в роботі результати можуть бути використані під час розроблення тестів у різних предметних галузях, а також служити базою для дослідження педагогічних можливостей і визначення галузі оптимального використання тестів з відновлюваними фрагментами.

Додаток

У наступних прикладах **напівжирно** виділені найменування аналізованих виразів, підкресленням виділені базові терміни, *курсивом* – конкретизуючих фрагментів, не виділені допоміжні фрагменти:

– у геометрії: **Опис формули для квадрата сторони довільного трикутника, який лежить проти гострого кута.** Квадрат сторони, що лежить *проти гострого кута*, дорівнює сумі квадратів двох інших сторін без подвоєного добутку однієї із цих сторін на проекцію іншої сторони на неї;

– у теоретичній механіці: **Аксиома паралелограма сил.** Система двох сил, прикладених до тіла в одній точці, має рівнодіючу, прикладену в *тій же точці* й рівну геометричній сумі сил;

– у термодинаміці: **Теорема Карно.** *Термічний* коефіцієнт корисної дії оборотного циклу Карно не залежить від *природи* робочого тіла і є функцією тільки абсолютних температур нагрівача й холодильника;

– в електростатиці: **Закон Кулона.** Сила електростатичної взаємодії між двома точковими електричними зарядами прямо пропорційна добутку величин зарядів і обернено пропорційна квадрату відстані між ними;

– у біології: **Правило необоротності еволюції.** *Окремі ознаки* далеких предків можуть з'являтися в нащадків, *зворотні мутації* можуть привести до *повторного виникнення даного аллеля*, але не *цілого генного комплексу* й відповідного *цілісного фенотипу*;

– у токарській справі: **Призначення задньої бабки.** *Задня бабка* – *основний вузол токарно-гвинторізного верстата*, який служить для *підтримки не закріпленого кінця оброблюваної деталі*, і *установлення інструменту*;

– у техніці високих напруг: **Характеристики резисторів генератора імпульсів напруги.** Для нормальної роботи генератора імпульсів напруги необхідно, щоб для *активних опорів еквівалентної схеми* встановлювалося таке співвідношення: опір зарядного резистора повинен бути *багато більше* опору формуючого резистора.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Яндекс словари. Энциклопедии [Электронный ресурс] / В. Г. Крысько
Этнопсихологический словарь. – М., 1999. – Режим доступа :

<http://slovari.yandex.ru/>

ПОНЯТИЙНЫЙ%20и%20КАТЕГОРИАЛЬНЫЙ%20АППАРАТ/значение/. –

Назва з екрана.

2. Новая философская энциклопедия, 2003 г. [Электронный ресурс]. – Режим доступа : <http://www.term.ru/dictionary/879/word/kategori>.
3. Большой психологический словарь. 2004 [Электронный ресурс] / Сост. Мещеряков Б., Зинченко В. – Режим доступа : <http://www.term.ru/dictionary/879/word/kategori>.
4. *Анцупов А. Я.* Словарь конфликтолога [Электронный ресурс] / А. Я. Анцупов, А. И. Шипилов // Национальная психологическая энциклопедия. – Режим доступа : <http://vocabulary.ru/dictionary/887/word/validnost-testa>. – Назва з екрана.
5. *Кузьмінський А. І.* Тест навчальних досягнень особистості як засіб педагогічного вимірювання / А. І. Кузьмінський, В. І. Єфіменко. – Черкаси : Видавничий відділ Черкаського державного університету імені Богдана Хмельницького, 2002. – 64 с. – ISBN 966-7986-33-0.
6. Яндекс словари. Энциклопедии [Электронный ресурс] / Психология от А до Я. – Режим доступа : <http://slovari.yandex.ru/валидность%20теста/значение/>. – Назва з екрана.
7. *Виненко В. Г.* Общие основы педагогики : учеб. пособие / В. Г. Виненко. – М. : Дашков и К, 2010. – 300 с. – ISBN 978-5-8114-0082-9.
8. Отраслевой терминологический стандарт. Термины и определения. [Электронный ресурс]. – Режим доступа : <http://do.gendocs.ru/docs/index-197255.html>.
9. *Раков С. А.* Педагогічні можливості тестів з математики формату SA / С. А. Раков // Комп'ютер у школі та сім'ї. – 2004. – №1. – С. 14–18.
10. *Петков О. О.* Оцінка рівня засвоєння термінологічної системи обмеженої галузі знань [Электронный ресурс] / О. О. Петков // Інформаційні технології і засоби навчання. – 2012. – № 4 (30). – Режим доступа : <http://journal.iitta.gov.ua/index.php/itlt/article/view/705/526>.

11. *Васильев И. Б.* Диагностика обученности и уровни усвоения знаний / И. Б. Васильев // *Професійна освіта: теорія і практика. Науково-методичний бюлетень.* – 2002. – №1–2 (15–16). – С. 51–69.
12. *Петков А. А.* Компьютерная программа "Компьютерная программа контроля уровня усвоения теоретического материала учебного курса "Теорема". Свидетельство о регистрации авторского права на произведение № 28537 от 24.04.2009. Государственный департамент интеллектуальной собственности Министерства образования и науки Украины.
13. *Гмурман В. Е.* Теория вероятностей и математическая статистика : учеб. пособие для вузов / В. Е. Гмурман. – М. : Высш. шк., 2003. – 479 с. – ISBN 5-06-004214-6.
14. *Сидоренко Е. В.* Методы математической обработки в психологии / Е. В. Сидоренко. – СПб. : ООО "Речь", 2003. – 350 с. – ISBN 5-9268-0010-2.
15. *Петков А. А.* Совершенствование контроля знаний при подготовке специалистов электротехнического профиля / А. А. Петков // *Вісник Національного технічного університету "Харківський політехнічний інститут". Збірник наукових праць. Проблеми удосконалення електричних машин і апаратів. Теорія і практика.* – Х. : НТУ "ХПІ". – 2010. – №36. – С. 94–98.
16. *Вентцель Е. С.* Теория вероятностей / Е. С. Вентцель – М. : Высш. шк., 2002. – 575 с.

Матеріал надійшов до редакції 03.01.2013 р.

ДИАГНОСТИЧЕСКИЕ ХАРАКТЕРИСТИКИ КОМПЬЮТЕРНЫХ ТЕСТОВ, ФОРМИРУЕМЫХ ПО МЕТОДУ ВОССТАНАВЛИВАЕМЫХ ФРАГМЕНТОВ

Петков Александр Александрович, кандидат технических наук, старший научный сотрудник, ведущий научный сотрудник Научно-исследовательского и проектно-конструкторского института "Молния" Национального технического университета "Харьковский политехнический институт", доцент кафедры "Инженерная электрофизика" Национального технического университета "Харьковский политехнический институт", г. Харьков, e-mail: apetkov@yandex.ru.

Аннотация

В статье рассмотрено определение валидности и надежности тестов, которые формируются по методу восстанавливаемых фрагментов. Проанализирована структура контролируемого теоретического материала ограниченной области знаний, языковых выражений, которые описывают предметы контроля, и надежности теста. Приведена методика определения наиболее важных составляющих надежности рассмотренных тестов: надежности количественного определения коэффициента усвоения и технологической надежности. Результаты проведенных педагогических экспериментов доказали, что тесты данного класса позволяют с высокой надежностью проводить контроль усвоения теоретического материала на уровне воспроизведения в любой области знаний. Показано, что валидность тестов с возобновляемыми фрагментами в основном обуславливается степенью структурирования и методической проработкой контролируемого материала и может достигать наперед заданных показателей, вплоть до уровня абсолютной валидности.

Ключевые слова: компьютеризированный контроль знаний, диагностика, валидность теста, надежность теста.

DIAGNOSTIC CHARACTERISTICS OF THE COMPUTER TESTS FORMED BY METHOD OF RESTORED FRAGMENTS

Alexandr A. Petkov, PhD (technical sciences), senior researcher, chief scientific employee of Design -&- Research Institute "Molniya" of National technical university "Kharkov polytechnic institute", associate professor of the Department "Engineering electrophysics" of National technical university "Kharkov polytechnic institute", Kharkov, e-mail: apetkov@yandex.ru.

Resume

Definition of validity and reliability of tests which are formed by a method of restored fragments is considered in the article. The structure of the controlled theoretical material of limit field of knowledge, language expressions that describe the subject of control, and reliability of test, is analyzed. The technique of definition of the most important components of reliability of the considered tests is given: reliability of quantitative determination of coefficient of assimilation and technological reliability. Results of the lead pedagogical experiments have proved, that tests of the given class allow to make the control

of mastering of a theoretical material over a level of reproduction in any field of knowledge with high reliability. It is shown, that validity tests with restored fragments basically caused by a degree of structurization and methodical study of a controllable material and can achieve beforehand set parameters, down to a level of absolute validity.

Keywords: the computerized control of knowledge, diagnostics, validity of the test, reliability of the test.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Yandex dictionaries. Encyclopedias [online] / V. G. Krys'ko. Ethnopsychological dictionary. – M., 1999. – Available from : [http://slovari.yandex.ru/PONJaTIJNYJ%20i%20KATEGORIAL"NYJ%20APPARAT/znachenie/](http://slovari.yandex.ru/PONJaTIJNYJ%20i%20KATEGORIAL). (in Russian)
2. New philosophical encyclopedia, 2003 [online]. – Available from : <http://www.term.ru/dictionary/879/word/kategori>. (in Russian)
3. Large psychological dictionary. 2004 [online] / Sost. Meshherjakov B., Zinchenko V. – Available from : <http://www.term.ru/dictionary/879/word/kategori>. (in Russian)
4. *Ancupov A. Ja.* Dictionary of conflictist [online] / A. Ja. Ancupov, A. I. Shipilov // Nacional'naja psihologicheskaja jenciklopedija. – Available from: <http://vocabulary.ru/dictionary/887/word/validnost-testa>. (in Russian)
5. *Kuzminskyi A. I.* Test of educational achievements of personality as means of the pedagogical measuring / A. I. Kuzminskyi, V. I. Yefimenko. – Cherkasy : Vydavnychiy viddil Cherkaskoho derzhavnoho universytetu imeni Bohdana Khmelnytskoho, 2002. – 64 s.– ISBN 966-7986-33-0. (in Ukrainian)
6. Yandex dictionaries. Encyclopedias [online] / Psihologija ot A do Ja. – Available from: <http://slovari.yandex.ru/validnost'%20testa/znachenie/>. (in Russian)
7. *Vinenko V. G.* General bases of pedagogic : ucheb. posobie / V. G. Vinenko. – M. : Dashkov i K, 2010. – 300 s. – ISBN 978-5-8114-0082-9. (in Russian)
8. Industry terminological standard. Terms and definition [online]. – Available from: <http://do.gendocs.ru/docs/index-197255.html>. (in Russian)
9. *Rakov S. A.* Pedagogical possibilities of tests are from mathematics of format SA / S. A. Rakov // Kompiuter u shkoli ta simi. – 2004. – №1. – S. 14–18. (in Ukrainian)

10. *Pietkov O. O.* Estimation of mastering level of terminological system of the knowledge limited field [online] / O. O. Pietkov // Informatsiini tekhnolohii i zasoby navchannia. – 2012. – № 4 (30). – Available from: <http://journal.iitta.gov.ua/index.php/itlt/article/view/705/526>. (in Ukrainian)
11. *Vasil'ev I. B.* Diagnostics of train and levels of mastering digestion of knowledge / I. B. Vasil'ev // Profesiina osvita: teoriia i praktyka. Naukovo-metodychnyi biuleten. – 2002. – №1–2 (15–16). – S. 51–69. (in Russian)
12. *Petkov A. A.* Computer program: "Computer program of control of level of digestion of theoretical material of educational course "Teorema". Svidetel'stvo o registracii avtorskogo prava na proizvedenie # 28537 ot 24.04.2009. Gosudarstvennyj departament intelektual'noj sobstvennosti Ministerstva obrazovanija i nauki Ukrainy. (in Russian)
13. *Gmurman V. E.* Theory of chances and mathematical statistics: ucheb. posobie dlja vuzov / V. E. Gmurman. – M. : Vyssh. shk., 2003. – 479 s. – ISBN 5-06-004214-6. (in Russian)
14. *Sidorenko E. V.* Methods of mathematical treatment are in psychology / E. V. Sidorenko. – SPb. : OOO "Rech'", 2003. – 350 s. – ISBN 5-9268-0010-2. (in Russian)
15. *Petkov A. A.* Perfection of the knowledge control at preparation of electrotechnical profile specialists / A. A. Petkov // Visnyk Natsionalnoho tekhnichnoho universytetu "Kharkivskyi politekhnichniyi instytut". Zbirnyk naukovykh prats. Problemy udoskonalennia elektrychnykh mashyn i aparativ. Teoriia i praktyka. – Kh. : NTU "KhPI". – 2010. – №36. – S. 94–98. (in Russian)
16. *Ventcel' E. S.* Theory of chances / E. S. Ventcel' – M. : Vyssh. shk., 2002. – 575 s. (in Russian)