

УДК 37.013-047.36

старший викладач кафедри інформаційних технологій і математичних дисциплін Мазур Н. П.  
Київський університет імені Бориса Грінченка (м. Київ)

### ВИМОГИ ДО ТЕСТУ ЯК ІНСТРУМЕНТУ ОТРИМАННЯ ДАНИХ МОНІТОРИНГОВОГО ДОСЛІДЖЕННЯ

*Анотація.* У статті йдеться про поняття тесту, наводиться короткий огляд їх видів та звертається увага на вимоги до педагогічних тестів – надійність та валідність. Розглядаються поняття та види валідності, поняття надійності тесту та формули її розрахунку. Крім того, зазначається необхідність обчислення цих показників при проведенні моніторингових досліджень навчальних досягнень учнів/студентів.

*Ключові слова:* тест, надійність, розрахунок надійності, валідність, види валідності.

*Анотація.* В статті говориться о поняття теста, приводиться краткий обзор видов и обращается внимание на требования к педагогическим тестам – надежности и валидности. Рассматриваются понятие и виды валидности, понятие надежности теста и формулы ее расчета. Кроме того, отмечается необходимость вычисления этих показателей при проведении мониторинговых исследований учебных достижений учащихся/студентов.

*Ключевые слова:* тест, надежность, расчет надежности, валидность, виды валидности.

*Annotation.* In article it is spoken about concept of dough, to be brought the short review of types and the attention to requirements to pedagogical tests – reliability and a validity is paid. The concept and types of a validity, concept of reliability of dough and formula of its calculation are considered. Besides, need of calculation of these indicators is noted when carrying out monitoring researches of educational achievements of pupils/students.

*Key words:* test, reliability, calculation of reliability, validity, types of validity.

**Вступ.** Сучасна модель освіти орієнтує школу на індивідуальні потреби учнів і створення якомога кращих умов їх особистісного розвитку. Відповідно зростає контроль суспільства і над тим, як цей розвиток відбувається. Сучасний вчитель здебільшого має самостійно виокремлювати у своїй діяльності та ефективно вирішувати різні психолого-педагогічні задачі, має бути педагогічно компетентним, як до розв'язання задач психолого-педагогічного, так і профільного (суто предметного) спрямування. Діяльність сучасного вчителя повинна бути інноваційною по суті. До інновацій у сфері контролю і оцінки якості результатів освітнього процесу відноситься і моніторинг.

Основними етапами проведення моніторингового дослідження є: цілепокладання та планування дослідження, розробка інструментарію, проведення дослідження, збір та обробка результатів, аналіз та інтерпретація результатів дослідження. До базових інструментів моніторингу відносять тести та анкети.

Теоретичні та практичні проблеми реалізації тестування досліджувалися В. Аванесовим, А. Анасазі, Н. Гронлундом, В. Звонніковим, О. Майоровим, Н. Самилкіною, М. Челишковою.

**Формулювання мети статті та завдань.** Мета та завдання написання статті полягають у огляді тесту як інструменту отримання даних дослідження, аналізі вимог, що визначаються до тестів та способів їх розрахунків.

**Виклад основного матеріалу статті.** Одним з етапів проведення моніторингу є отримання даних за допомогою тестів, контрольних робіт, зрізів знань, анкетування, опитування, які корелюються зі змістом освіти. Для перевірки теоретичних знань віддають перевагу саме тестуванню, оскільки різні види тестів та тестових запитань дозволяють провести контроль різнобічно та об'єктивно.

За основу поняття «тест» візьмемо тлумачення О. Майорова. Під тестом він розуміє «інструмент, що складається з кваліметрично вивіреної системи тестових завдань, стандартизованої процедури проведення й завчасно спроектованої технології обробки і аналізу результатів, та призначений для вимірювання якостей та властивостей особистості» [4, с. 14].

Якщо розглядати тест як інструмент вимірювання рівня навчальних досягнень учнів, то науковці підходять до цього питання по-різному. Так, наприклад, Локшина О.І. [6] запропонувала виокремити два головні підходи, на яких вимірювання учнівських досягнень можуть ґрунтуватися. Перший – це коли тест конструюють з навчального матеріалу, який містить знання й навички, що мав засвоїти учень. На основі результатів тестування встановлюють так звані норми. Результати відповідей учнів, класів, шкіл або освітніх систем відображаються у показниках, які порівнюють з розподілом балів усієї кількості тих, хто складав тест.

В основі альтернативного підходу лежить визначення рівня знань, які має продемонструвати середній учень на певному рівні розвитку, наприклад у 9-му класі. Цей рівень можуть встановлювати вчителі, шкільна рада, управлінські структури (наприклад, міністерство) тощо. При такому підході результати отримують як співвідношення або завдань, на які конкретний учень відповів правильно, до всіх одержаних завдань або учнів у конкретному класі чи системі, які досягли визначеного критерію, до всіх учнів, що брали участь у замірах навчальних досягнень.

Зазначені підходи дають різну по-суті інформацію. Перший – нормативне тестування (вимірювання) – відображає рівень знань конкретних учнів або груп учнів порівняно з контрольною групою; другий – критеріальне тестування – про рівень знань конкретних учнів або груп учнів порівняно з визначеним обсягом знань (вмінь, навичок). Досить часто, більшість тестів, що ґрунтуються на таких підходах, проводяться у паперовому вигляді з вибором відповіді [6].

У літературі зустрічається досить багато різних класифікацій тестів і зрозуміло, що їх можна класифікувати за різними ознаками. Наприклад, за процедурою створення (стандартизовані, нестандартні), за способами представлення (бланкові, предметні, апаратні, практичні, комп'ютерні), за спрямованістю (тести інтелекту, тести особистості, тести здібностей, тести досягнень), за характером дій (вербальні, невербальні), по ступеню однорідності завдань (гомогенні, гетерогенні) та ін.

Проте кожен з перелічених тестів повинен володіти рядом ознак, недотримання яких може призвести до викривлення результатів. Ці вимоги до тесту як інструменту вимірювання містять вимоги до показників якості тестів. До таких характеристик відносять валідність, надійність, стандартизованість, простоту, стислість у часі та ін.

**Надійність** відображає точність тестових вимірювань та стійкість тестових результатів до дій випадкових факторів. Дана характеристика має два значення: точність вимірювання та стійкість. Точність вимірювання означає стійкість до дії перешкод (стану тих, хто проходить тестування, їх ставлення до самої процедури та ін.) та можливих випадковостей при проведенні процедури тестування. Тест вважають стійким, якщо отримані близькі результати при повторному тестуванні в інших схожих умовах [7].

Н. Гронлунд зазначає, що тести з оцінки результатів мають бути надійними і у зв'язку з цим їх розробка має здійснюватися дуже ретельно. Якщо бал, отриманий учнем в результаті тесту буде відповідати тій оцінці, яку він отримав би при повторному проходженні того ж тесту або ідентичного до нього по формі, то дана оцінка вважається високо надійною [3].

Тест називається надійним, якщо він є внутрішньо узгодженим.

Надійність тесту, як правило, визначається після закінчення аналізу задач й складання остаточної форми тесту та для її визначення проводиться спеціальна апробація тесту.

На практиці використовуються три основні методи оцінки надійності тестів:

- 1) повторне тестування (ретестування) – визначення завадостійкості;
- 2) розщеплення групи – визначення завадостійкості;
- 3) розщеплення тесту – визначення внутрішньої узгодженості [4, с. 172].

Для розрахунку надійності за внутрішньою узгодженістю використовують, зазвичай, статистичні формули. Одним з найкращих показників для розрахунку надійності тесту, на думку фахівців, є коефіцієнт  $\alpha$

$$\alpha = \frac{k}{k-1} \left( 1 - \frac{\sum \delta_i^2}{\delta_y^2} \right) \quad (1),$$

де  $k$  – кількість завдань,  $\sum \delta_i^2$  – сума квадратів стандартних відхилень для завдань,  $\delta_y^2$  – квадрат стандартного відхилення для всього тесту.

Ще одним способом розрахунку коефіцієнта надійності є формула К'юдера-Річардсона (скорочена назва KR-20), яка являє собою частковий випадок коефіцієнту  $\alpha$  для завдань дихотомічного типу, а отже, використання якої для тестів навчальних досягнень є допустимою:

$$r_{kr} = \frac{k}{k-1} \left( 1 - \frac{\sum PQ}{\delta_y^2} \right) \quad (2),$$

де  $P$  – частка учнів, що вірно виконали завдання,  $Q = 1 - P$

Порівнюючи формули (1) та (2), можна бачити, що  $\sum \delta_i^2$  – сума квадратів стандартних відхилень для завдань замінена на суму добуток часток учнів, які правильно виконали завдання, та частки учнів, які з завданням не впоралися. Частки учнів, які правильно та неправильно виконали завдання, розглядалися як різноманітні показники складності тестових завдань [4].

З аналізу робіт по організації та проведенню тестування, можна зазначити рекомендовані шляхи підвищення надійності тестів:

- ✓ збільшення довжини тестів – чим довший тест, тим він надійніший. При наявності малої кількості завдань присутня велика ймовірність вгадування відповіді. Що стосується питання визначення достатньої кількості запитань у тесті, то цей показник може бути визначений лише після проведення відповідних розрахунків надійності;
- ✓ перевірка об'єктивності схеми оцінки – результати виконання завдань повинні оцінюватися однаково різними експертами;
- ✓ зниження ймовірності вгадування правильних відповідей;
- ✓ посилення інструкцій для учнів та тих, хто проводить тестування, інструкцій з перевірки результатів тестування, інших інструкцій, шляхом підвищення якості інструктажу та підготовки тих, хто буде проводити тестування.

Однієї надійності для обґрунтування якості тесту недостатньо. Ще однією з важливих характеристик тесту є його валідність. Вона особливо важлива для тестів, орієнтованих на критерій, оскільки визначення надійності для цих тестів ускладнене.

Найвдалішим на нашу думку означення валідності надано А. Анастасі. Валідність тесту – це поняття, яке вказує, що тест вимірює та наскільки добре він це робить [2]. Питання валідності постає в процесі розробки та практичного застосування тесту, коли ставиться завдання встановити відповідність між ступенем вираження певної властивості особистості та методом вимірювання цієї властивості. Чим краще в тесті відображена та властивість, заради вимірювання якої він створений, тим валіднішим є такий тест.

О. Майоров [4] пропонує виділяти три основні види валідності, які мають істотний відмінний зміст:

- 1) критеріальна валідність, яка має підвиди:
  - поточна (діагностична) валідність;
  - прогностична валідність;
- 2) конструктивна валідність, в рамках якої виділяють:
  - диференціальну валідність;
  - валідність за віковою диференціацією;
- 3) змістовна валідність.

Коротко охарактеризуємо у чому вони полягають.

Зміст *критеріальної валідності* полягає у визначенні здатності тесту бути індикатором або передбаченням строго визначеної психічної особливості, форми поведінки людини та ін.

Валідація тесту по критерію полягає у порівнянні результатів, отриманих тим, хто проходить тестування, у процесі виконання завдань з даними по критерію, та обчислення коефіцієнту кореляції тестового результату

з зовнішнім критерієм. У якості критерію може виступати будь-який показник, що вимірює ту ж характеристику, що й тест, який валідизується.

**Конструктивна валідність** визначається у тих випадках, коли уявлення про фермент (конструкт), який вимірюється, існує лише у свідомості дослідника. Розробник тесту може лише будувати гіпотезу про існування даного конструкту, про його форми та характер прояву. Встановлюється така валідність шляхом доведення правильності теоретичних концепцій, що покладені в основу тесту.

**Змістовна валідність** – основний вид валідності для тестів навчальних досягнень.

Визначення змістовної валідності використовується для тестів досягнень та тестів професійної успішності, коли повинен бути точно визначений матеріал, що використовується для тестування, та коли існує достатня зрозумілість змісту параметру, що вимірюється.

Даний вид валідності визначається на основі експертних оцінок.

Для визначення змістовної валідності тесту навчальних досягнень мінімально необхідно:

1) у відповідності з цільовими установками розробити технологічну матрицю;

2) провести її ретельну експертизу;

3) у відповідності до технологічної матриці розробити тестові завдання;

4) дати експертні висновки на відповідність тестових завдань технологічній матриці;

5) повідомити про виконані роботи у специфікації тесту або інструкції для того, хто проводитиме тестування [4].

Для всіх тестів навчальних досягнень має бути оцінена змістовна валідність. Для тестів, що використовуються для атестації учнів, змістовна та бажано критеріальна валідність.

**Висновки.** Підсумовуючи вищесказане можемо сказати, що тест як інструмент отримання даних дослідження увійшов практично у всі сфери педагогіки та психології. Проте використання ненадійних тестів, допуск великої кількості помилок у такій важливій справі, яким є тестування учнів/студентів, може стати причиною педагогічних та адміністративних помилок, наслідки яких важко виправити. Саме тому, при розробці та апробації тестів необхідно проводити ряд процедур їх перевірки на надійність та валідність. Надійність відображає точність тестових вимірювань та стійкість тестових результатів до дій випадкових факторів, а валідність вказує, що тест вимірює та наскільки добре він це робить. Особливо важливо акцентувати на цьому увагу студентів при формуванні їх готовності до організації та проведення моніторингових досліджень навчальних досягнень учнів.

#### Література:

1. Аванесов В.С. *Композиция тестовых заданий* / В.С. Аванесов. – М.: Ассоциация инженеров-педагогов, 1996. – 191 с.

2. Анастаси А., Урбина С. *Психологическое тестирование*. – 7-е изд. / А. Анастаси, С. Урбина – СПб.: Питер, 2007. – 688 с.

3. Гронлунд Н. *Тесты достижений в конструировании* / Н. Гронлунд. – Лондон, Прентис-Холл, 1982.

4. Майоров А.Н. Теория и практика создания тестов для системы образования. (Как выбирать, создавать и использовать тесты для целей образования) / А.Н. Майоров. – М., «Интеллект-центр», 2001. – 296 с.

5. Михайлычев Е.А. Дидактическая тестология / Е.А. Михайлычев. – М.: Народное образование, 2001. – 432 с.

6. Мониторинг якості освіти: становлення та розвиток в Україні: Рекомендації з освітньої політики / під заг. ред. О.І. Локишиної. К.: «К.І.С.», 2004. – 160 с.

7. Самылкина Н.Н. Современные средства оценивания результатов обучения / Н.Н. Самылкина. – М.: БИНОМ. Лаборатория знаний, 2007. – 172 с.

**References:**

1. Avanesov B.C. Kompozicija testovyh zadaniy / V.S. Avanesov. – М.: Associacija inzhenerov-pedagogov, 1996. – 191 s.

2. Anastazi A., Urbina S. Psihologicheskoe testirovanie. – 7-e izd. / A. Anastazi, S. Urbina – SPb.: Piter, 2007. – 688 s.

3. Gronlund N. Testy dostizhenij v konstruirovanii / N. Gronlund. – London, Prentis-Holl, 1982.

4. Majorov A.N. Teorija i praktika sozdaniya testov dlja sistemy obrazovaniya. (Kak vybirat', sozdavat' i ispol'zovat' testy dlja celej obrazovaniya) / A.N. Majorov. – М., «Интеллект-центр», 2001. – 296 с.

5. Mihajlychev E.A. Didakticheskaja testologija / E.A. Mihajlychev. – М.: Narodnoe obrazovanie, 2001. – 432 s.

6. Monitoryng h jakosti osvity: stanovlennja ta rozvytok v Ukrajinі: Rekomendaciji z osvitynoji polityky / pid zagh. red. O.I. Lokshynoji. К.: «К.І.С.», 2004. – 160 s.

7. Samylkina N.N. Sovremennye sredstva ocenivaniya rezul'tatov obuchenija / N.N. Samylkina. – М.: БИНОМ. Laboratorija znanij, 2007. – 172 s.

**Педагогіка**

**УДК: 371.383:78.085**

**кандидат педагогічних наук, доцент Мартиненко О. В.**  
Бердянський державний педагогічний університет

**ОСОБЛИВОСТІ ЗАСТОСУВАННЯ ПАРТЕРНИХ КОМПЛЕКСІВ У РОБОТІ ДИТЯЧИХ ХОРЕОГРАФІЧНИХ КОЛЕКТИВІВ**

*Анотація.* Стаття присвячена значенню партерної гімнастики в системі позашкільної хореографічної освіти дітей початкового рівня навчання. Автором виділені та охарактеризовані види партерних комплексів, визначені принципи їх побудови та вимоги до їх розробки, розкрито методику застосування партерних комплексів на уроках хореографії, зроблено акцент на впровадження інноваційних підходів в зміст партерної гімнастики (танцювальна терапія, контактна хореографія). На основі багаторічного досвіду роботи в дитячому хореографічному колективі доведено, що урізноманітнення змісту партерних комплексів може значно підвищити рівень хореографічної підготовки учнів.

*Ключові слова:* партерна гімнастика, позашкільне хореографічне навчання, види партерних комплексів.