

ОЦІНКА ЯКОСТІ ПЕДАГОГІЧНОГО ТЕСТУ ЯК ЗАГАЛЬНОНАУКОВОГО ДІАГНОСТИЧНОГО МЕТОДУ

У статті висвітлено та проаналізовано основні вимоги до педагогічних тестів як загальнонаукового діагностичного методу. Розглянуто основні критерії якості тестів навчальних досягнень, які висувають відомі науковці-тестологи.

Впровадження новітніх технологій навчання, що ґрунтується на нових підходах щодо подання та засвоєння знань, потребує сучасних методів їх вимірювання та оцінювання. Пошук досконалих методів вимірювання рівня професійної підготовки студентів на нинішньому етапі розвитку інформаційних технологій набуває надзвичайної актуальності, оскільки об'єктивізація процесу вимірювання, забезпечуючи зворотний зв'язок, дає можливість координувати цей розвиток. Отже, об'єктивні і точні методи вимірювання та оцінювання знань стають однією з рушійних сил наукового прогресу.

Порівняльний аналіз методів вимірювання рівня знань, які використовуються у сучасній педагогіці (усного опитування, письмової роботи, інтерв'ю та тестування), свідчить, що останній метод найбільше відповідає критеріям якості при визначенні рівня теоретичних знань, умінь і навичок студентів у сфері їх майбутньої професійної діяльності.

Використання тестів у навчальному процесі надійно увійшло у світову педагогічну практику, в тому числі вищої школи. В Україні цей процес також розвивається: збагачуються наші уявлення про призначення та педагогічні можливості тестів, їх форми, формати запитань, методи обробки результатів тестування та їх інтерпретації, питання специфікації тестів, їх підготовки, апробації, нормування, технології проведення тестування у ВНЗ.

Ситуація з тестуванням в Україні сьогодні нагадує 20-і роки ХХ ст., коли під керівництвом Блонського створювались перші вітчизняні шкільні стандарти і тести. До нині живе міф, що для складання тестового інструменту не потрібні спеціальні знання. Серед людей, які професійно займаються тестуванням, є такі, котрі прагнуть швидше створити тести. А доказом цього є величезна кількість книг з назвами: «Тести з ...», більшість яких не витримує критики.

Питань не стає менше, а значно більше. Широке використання тестових технологій – це дійсно інша культура освітнього процесу, що передбачає специфічну грамотність у відповідних питаннях психології, педагогіки, методики.

Існує проблема з об'єктивною інформацією, яку може дати тестування. На перший погляд, у такій інформації є проблема як у викладачів — для самооцінки професійної діяльності і оцінки досягнень студентів з метою корекції педагогічних впливів, так і в керівників навчального закладу — для визначення стратегії розвитку, оцінки якості роботи педагогів і корекції різноманітних управлінських рішень. Однак поява і використання в роботі об'єктивної інформації породжує чимало проблем, до вирішення яких педагоги не готові.

Метою статті є висвітлення й аналіз вимог до педагогічних тестів і тестування загалом, основних критеріїв, які висувають науковці-тестологи щодо конструювання якісних тестів.

Як відомо, лише правильно складений тест в змозі повністю відповідати сучасним цілям навчання та освіти. Багато викладачів мають певні утруднення у розробці та правильному методичному застосуванні таких тестів. Результати опитування 156 вчителів і викладачів навчальних закладів Запорізької області, які підтверджуються і результатами нашого опитування, засвідчили, що лише 32% з них постійно використовують у навчальному процесі тести, 68% — епізодично. Майже 80% викладачів, що використовують тестовий контроль, працюють за авторськими, власноруч розробленими тестами, при цьому маючи про тестові методики дуже поверхові знання [8, 7–8].

Таким чином, можна констатувати, що тут необхідні певні рекомендації. При здійсненні аналізу ми спиралися на фундаментальні дослідження тестування як методу педагогічної діагностики, висвітлені в роботах С. Аванесова, В. Беспалька, К. Інгенкампа, П. Клайна, А. Майорова, Л. Долінеро та ін.

Слово «тест» англійського походження. Класичним у педагогіці є визначення К. Інгекампа: «Тестування — це метод педагогічної діагностики, за допомогою якого вибір поведінки, що презентує передумови чи результати навчального процесу, повинен максимально відповідати принципам зіставлення, об'єктивності, надійності та валідності вимірів, повинен пройти обробку й інтерпретацію й бути прийнятним для застосування в педагогічній практиці» [11, 255].

Деяко інше означення поняття «тест» подають Л. Бурлачук та С. Морозов. Вони вважають, що тест — це фіксоване в часі випробування, призначене для встановлення кількісних та якісних індивідуально-психологічних якостей [3, 325–326].

Схоже означення тесту вводить психолог П. Клайн. Психологічний тест (а педагогічні тести він розглядає як підвид психологічних) — це стандартизоване, часто обмежене в часі випробування, призначене для встановлення кількісних та якісних індивідуально-психологічних особливостей [6, 15].

В. Аванесов вважає, що тест є науково обґрунтованим методом вимірювання досліджуваних особливостей особистості. Попри узагальнене поняття тесту, часто вводиться поняття педагогічного тесту. Так, той же вчений вважає, що педагогічний тест — це сукупність взаємопов'язаних завдань зростаючої складності і специфічної форми, що дозволяє якісно оцінити структуру і виміряти рівень знань [13].

Кожне з наведених визначень має свої недоліки та переваги. Тому, проаналізувавши ці означення і врахувавши мету статті, ми при дослідженні означеної проблеми опирались на наступне розуміння сутності понять «тест» і «тестування».

Тестування — це науково обґрунтований процес вимірювання (за допомогою тестів) якості властивостей особистості.

Тест — це інструмент, що складається з кваліметрично вивіреної системи тестових завдань, стандартизованої процедури проведення, із заздалегідь спроектованої технології обробки та аналізу результатів.

Як видно з поданих означень, важливими критеріями діагностичних тестів навченості є об'єктивність, дієвість (валідність), ефективність та надійність їх результатів. Цими критеріями користуються більшість вчених-тестологів.

Тест успішності, як і будь-який інший метод контролю, повинен бути насамперед об'єктивним. Тест успішності вважається об'єктивним тоді, коли вдається максимально виключити інтерсуб'єктивний вплив дослідників на його результати, коли останні не залежать від особистості людини, котра проводить вимірювання, опрацьовує результати та робить підсумки. Об'єктивність проведення вимірювання вимагає, щоб всі студенти проходили тестування в однакових умовах [4].

Варто зазначити, що об'єктивність не повинна бути самоціллю, вона лише забезпечує валідність і надійність вимірювань. Необ'єктивне вимірювання не може бути валідним та надійним.

Валідність тесту — це найважливіший критерій його якості і визначається як його дієвість і адекватність.

Як зазначають автори «Словника-довідника з психодіагностики», валідність (від англійського *valid* — дійсний; придатний; той, що має силу) — це комплексна характеристика методики (тесту), що включає відомості про сферу досліджуваних явищ і репрезентативність діагностичної процедури щодо них [3, 34].

Отже, валідність методу — це відповідність того, що вимірюється цим методом, тому що він повинен вимірювати. Цей критерій встановлює сферу дійсності, для якої метод дає статистично ймовірні результати. Якщо мова йде про тестування діяльності певного рівня, то кваліфікаційні завдання, які пропонуються у тесті, повинні відповідати саме такому рівню складності, тобто їх не можна виконати засобами діяльності нижчого рівня. При цьому говорять про функціональну валідність тесту.

А. Анастасі пропонує більш просте і загальне визначення валідності. На її думку, валідність тесту — це поняття, що вказує на те, що вимірює тест і наскільки добре він це робить [1, 133]. Отже, тест вважають валідним, коли він вимірює саме те, для чого призначений.

Важливо відзначити, що валідність пов'язана за багатьма параметрами зі сферою дослідження і завданнями діагностики. Тест не можна назвати валідним чи невалідним без вказівки сфери його застосування. Розглядаючи поняття «валідність», варто звернути увагу на суттєвий момент — відносний характер валідності. На думку А. Майорова, в загальному вигляді поняття валідності «розкриває, наскільки отримані результати тестування відповідають об'єктивній реальності, а оскільки не існує методу точного визначення об'єктивної реальності, то валідність — поняття відносне, оскільки тест, як і будь-який вимірювальний інструмент, ніколи не буває абсолютно еквівалентним реальності. Отже, валідність встановлюється відповідно характеристикі, ознаки, величини, що приймається за об'єктивну» [7, 212].

Валідність методу при вимірюванні успішності можна диференціювати за такими критеріями: валідність змісту, валідність відповідності і валідність прогнозу [9, 18–19].

Л. Долінер розрізняє очевидну (внутрішню), змістовну та функціональну валідність [5, 49–51].

а) *Очевидна (внутрішня) валідність*. Вважають, що тест є валідним, коли про нього складається враження у випробуваних, що він вимірює необхідне. Очевидна валідність не має ніякого відношення до справжньої валідності і лише допомагає створити атмосферу співпраці з випробуваним.

б) *Змістовна валідність* тесту. Цей термін визначає, чи завдання тесту відповідає всім аспектам досліджуваного питання (тема, розділ та ін.). Тобто, чи завдання тесту побудовано на основі тільки технічних навчальних елементів, які вивчалися учнями, і на тому рівні абстракції, на якій вони висловлювалися.

в) *Функціональна валідність* тесту може бути розкрита як відповідність контрольного завдання тій пізнавальній дії, яка перевіряється.

На думку відомого психолога-тестолога П. Клайна, тест можна називати валідним, якщо він вимірює те, для вимірювання чого він призначений. Але таке визначення не пояснює значення валідності. В цьому випадку виникає нове запитання: як ми дізнаємось, що тест вимірює те, для чого він призначений? На думку вченого, існує багато різних способів доведення валідності тестів, і кожен з них відповідає різним аспектам цього значення:

- тест є очевидно валідним, якщо про нього у опитуваних складається враження, що він вимірює те, що необхідно;
- конкурентна валідність оцінюється за кореляцію результатів певного тесту з результатами інших тестів;
- для встановлення прогностичної валідності тесту вивчаються кореляції між показниками тесту і деяким критерієм, що характеризує вимірювану властивість, але через певний час;
- якщо можна показати, що завдання відображають всі аспекти досліджуваної сфери, то тест є змістовно валідним, за умови, що чітко сформульована інструкція до нього;
- щоб продемонструвати конструктивну валідність тесту, необхідно якнайповніше описати зміну, для вимірювання якої призначений тест. Конструктивна валідність включає в себе всі підходи до визначення валідності, перераховані вище [6, 25–29].

У стандартних вимогах до педагогічних тестів валідність тесту визначають як комплекс інформації про характеристики тесту, процедури тестування та методики оцінювання. Тому автори методичного посібника зі складання тестів для вчителів [10, 85–86] загальне поняття валідності диференціюють за категоріальними ознаками на:

Валідність методу, що складається з валідності змісту та валідності відповідності. Валідність змісту визначає ступінь репрезентативності обсягу знань, визначеному в кількості тестових завдань, які містить тест. Вона пов'язана із загальним обсягом знань, які повинні засвоїти студенти, та з обсягом інформації, винесеної на тестування. Валідність відповідності — це «повна відповідність тесту рівня діяльності і змісту навчального матеріалу».

Валідність інструментарію — це відповідність *характеристик вимірювання певним критеріям, що забезпечують об'єктивність результатів. У її межах розрізняють валідність тесту та валідність тестових завдань.*

Валідність процедури тестування досягається чітким визначенням вимог щодо умов тестування, дій персоналу та поведінки тих, кого тестують.

Валідність процедури оцінювання.

Наступним критерієм якості тесту є його надійність, що показує, чи забезпечує тест як вимірювальний інструмент об'єктивність результатів. Використання ненадійних тестів може стати причиною педагогічних помилок.

Надійність методу вимірювання — це міра стійкості результатів, що впливає на точність, з якою можна виміряти ту чи іншу конкретну ознаку. Перевірка надійності методу стосується насамперед відновлення результатів при повторних вимірах.

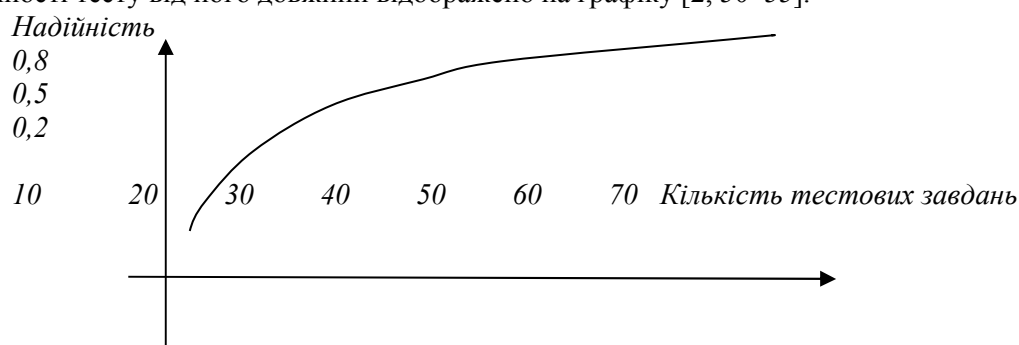
Ступінь надійності методу залежить від: об'єктивності методу; параметрів засобу вимірювання; стабільності характеристики, яку вимірюють [9, 17–18].

Під надійністю контрольного завдання розуміють ступінь точності, з яким може бути визначена та чи інша ознака, тобто встановлено, наскільки можна довіряти результатам певного тесту. Про надійність тестів іноді роблять висновки за такою ознакою: якщо в усіх випадках перевірки тесту чи його варіантів виявиться, що учні в розподілі за показниками успішності займуть ті ж місця, то такий тест можна вважати надійним.

М. Челишкова трактує термін «надійність» у двох значеннях: по-перше, якщо він забезпечує високу точність вимірювань, і, по-друге, тест вважається надійним, якщо він дає при повторному виконанні схожі результати за умови, що підготовка учнів не змінилась за час повторного виконання тесту [12, 29].

Надійність тесту залежить від його довжини. Поняття «довжина тесту» введено на початку ХХ ст. У. Спірманом та означає кількість завдань тесту [13].

В. Безпалько вважає, що для достатньої надійності підсумкового контролю вистачає великих розділів курсу навчання, тест має містити не менше як 40 завдань. Залежність надійності тесту від його довжини відображено на графіку [2, 50–55].



П. Клайн вважає, що в психометрії термін «надійність» має два значення. Тест називається надійним, якщо є внутрішньо узгодженим. Він також називається надійним, якщо дає ті ж показники для кожного випробуваного (при умові, що випробуваний не змінився) при повторному тестуванні. Надійність при повторному тестуванні через певний час називають ретестовою надійністю [6, 22–25].

Оцінка надійності проводиться різними методами. Зупинимось на деяких з них. Методи оцінки надійності тестів можна умовно поділити на дві групи: перша базується на двократному тестуванні з використанням одного і того ж тесту або за допомогою двох паралельних форм тесту; друга група передбачає однократне тестування. Двократне тестування з метою перевірки надійності тесту ще називають ретестовим методом оцінки надійності тесту.

Степінь надійності тесту визначається за допомогою коефіцієнта надійності (коефіцієнт кореляції), який характеризує стійкість показників при повторних дослідженнях за допомогою того ж тесту або його рівноцінного варіанта.

Метод повторного тестування має свої переваги та недоліки. До переваг можна віднести природність і простоту визначення коефіцієнта надійності. Основним недоліком цього методу є невизначеність у виборі інтервалу між тестуваннями. Думки вчених щодо цього питання розходяться. Крім того, при повторному тестуванні дуже важко досягнути тих же умов.

В зв'язку з цим існує ще один спосіб визначення надійності тестів, який не має недоліків попереднього методу, — це надійність паралельних форм. В цьому випадку одні і ті ж досліджувані проходять тестування за еквівалентними або паралельними наборами питань. Коефіцієнт кореляції між результатами обох тестувань — основний показник цієї

характеристики надійності. Як бачимо, досліджувані виконують два тести за однакових умов, але складність цього методу полягає в доведенні еквівалентності даних тестів.

Ще одним способом визначення надійності тесту є визначення надійності частин тесту. Суть цього методу зводиться до виконання досліджуваним двох рівноцінних частин тесту. Він виник як результат розвитку методу паралельного тестування і базується на твердженні про паралельність не лише форм тесту, а й окремих завдань всередині тесту. Метод розщеплення використовується, як правило, у тих випадках, коли повторне тестування є небажаним або існує лише одна форма тесту. Проблема, що виникає при використанні надійності частин тесту, пов'язана з тим, як розділити тест на дві частини. Оскільки завдання у тесті повинні розміщуватись у порядку зростання складності, то розбиттям тесту пополам ми отримуємо нееквівалентні форми тесту. Тому часто тест розщеплюють за парними і непарними номерами тестових завдань. Отримані показники по двох частинах тесту корелюються звичайними методами [1, 114–116], [3, 162].

Варто зазначити, що хоча надійність тесту — це важлива його характеристика, але вона сама по собі цінності не становить. Надійність потрібна для досягнення валідності. Тест може бути високонадійним, але не валідним, тому неефективним.

В. Аванесов вважає, що зі становленням нової теорії тестів починає відчуватись необхідність розширення кількості критеріїв, хоча б до трьох. Третім педагог пропонує критерій ефективності.

Ефективність — це порівняльний критерій, що дозволяє порівняти тести. Ефективним можна назвати тест, який краще, ніж інші тести, вимірює знання учнів потрібного рівня підготовки, з меншою кількістю завдань, якісніше, швидше, дешевше, і все це — за можливості одночасно [13].

Якщо порівняти поняття «ефективність» з поняттями «надійність» і «валідність», то найістотніша відмінність нового поняття від двох традиційних полягає у переході від середнього показника до диференційованого. Достатньо пригадати, що надійність стосується тесту, який складається із зафіксованого числа завдань, що пред'являються всім випробовуваним; тільки тоді можна знайти коефіцієнт надійності тесту як середню міру точності вимірювання. Аналогічно знаходиться і валідність тесту. Ефективний же тест, навпаки, припускає відхід від усереднювання і фіксованого для всіх випробовуваних числа завдань.

Психологи Л. Ф. Бурлачук та С. М. Морозов до ефективних тестів висувають ще одну суттєву вимогу — складність завдання тесту. Складність завдання тесту — це характеристика завдання тесту, що відображає статистичний рівень розв'язуваності у певній вибірці стандартизації. Показником складності тестового завдання є частка випробуваних вибірки, що розв'язали або не розв'язали дане завдання. Наприклад, якщо лише 20% випробуваних виконали завдання, його можна вважати складним для даної вибірки, якщо 80% — легким [3, 349–351].

Підбір завдань за показниками складності важливий для успішного використання тестів. Так, при підборі надто складних тестових завдань валідність і надійність тесту різко зменшується. Надто прості тестові завдання призведуть до одноманітності тесту і його неефективності.

Ще однією особливістю ефективних тестів, на думку П. Клайна, є дискримінативність. Досягнення задовільного розподілу показників є однією з цілей автора тестів. Немає необхідності відзначати те, що є очевидним: яка цінність тесту, за яким всі випробувані показали однаковий результат? За допомогою ретельного конструювання тесту можна забезпечити відповідний рівень дискримінативності, а це те, в чому тести виграють порівняно з іншими формами випробувань [6, 29–30].

Як бачимо, вимоги, які ставлять тестологи до ефективних тестів, між собою дещо відрізняються. В педагогічній діагностиці намагаються визначити якість результатів вимірювання (в нашому випадку — вимірювання успішності). Варто зазначити, що всі вищезазначені критерії спрямовані на єдину загальну вимогу: тест повинен давати якнайточнішу інформацію про досліджуване явище. Вони лише конкретизують цю вимогу і характеризують її з різних сторін.

Отже, дидактичний потенціал тестових методик може бути реалізованим лише при дотриманні певних умов та вимог, серед яких визначальною є професійна готовність педагога до використання тестових технологій як засобу контролю професійної підготовки майбутніх спеціалістів.

ЛІТЕРАТУРА

1. Анастаси А., Урбина С. Психологическое тестирование. — 7-е изд. — СПб.: Питер, 2005. — 688 с.
2. Безпалько В. П. Образование и обучение с участием компьютеров (педагогика III тысячелетия). — М.: Изд-во Москов. Психол.-социал. Ин-та, 2002. — 352 с.
3. Бурлачук Л. Ф., Морозов С. М. Словарь-справочник по психодиагностике. — СПб.: Питер, 1999. — 528 с.
4. Валеев Г. Х. Методология и методы психолого-педагогических исследований: Учеб. пособие для студентов 3–5-х курс пед. вузов по специальности «031000 — Педагогика и психология». — Стерлитамак: Стерлитамак. гос. пед. ин-т, 2002. — 134 с.
5. Долинер Л. И. Компьютерные тесты успеваемости как средство оптимизации учебного процесса // Вестник Московского университета. — 2004. — № 1.
6. Клайн П. Справочное Руководство по конструированию тестов / Перевод Е. П. Савченко. — М.: ПАН Лтд, 1994. — 283 с.
7. Майоров А. Н. Теория и практика создания тестов для системы образования. — М.: Народное образование, 2000. — 352 с.
8. Малихін А. Тести у навчальному процесі сучасної школи // Рідна школа. — 2001. — № 8.
9. Методичні рекомендації з організації тестового контролю освітньо-професійної підготовки вчителя. — Тернопіль: Вид-во ТНПУ ім. В. Гнатюка, 2004. — 100 с.
10. Парашенко Л., Леонський В., Леонська Г., Богачков Ю., Радченко В. Система тестування учнів у середній школі: Методичний посібник // Завуч. — 2007. — № 16–18.
11. Педагогическая диагностика / Под ред. К. Ингенкампа; Перевод с нем. — М., 1991. — 525 с.
12. Чельшкова М. Б. Применение математических моделей для разработки педагогических тестов: Учеб. пособие. — М., 1995. — 48 с.
13. <http://testolog.narod.ru/index.html>

Надія ГИШКА

ФОРМУВАННЯ ЗАГАЛЬНОНАВЧАЛЬНИХ УМІНЬ І НАВИЧОК В УЧНІВ ПОЧАТКОВИХ КЛАСІВ

У статті розкриваються закономірності і сутність процесу формування загальних навчальних умінь і навичок в учнів початкових класів. Подається загальна характеристика цих умінь і навичок школярів. Показана роль самоконтролю і контролю в програмованому навчанні учнів.

Швидкість соціальних і технологічних змін у сучасному світі, прискорене нагромадження інформаційних ресурсів і засобів навчання, що стають доступними більшості людей, мобільність населення зумовлюють переосмислення функцій і результатів загальної середньої освіти. Сучасне інформаційне суспільство ставить перед школою завдання формувати в учнів бажання і вміння вчитися, використовувати різні джерела інформації, застосовувати одержані знання для творчого вирішення проблеми [1].

О. Я. Савченко розглядає вміння вчитись як ключову компетентність загальноосвітньої школи. На її думку, «вміє вчитися той учень, який:

- сам визначає мету діяльності або приймає поставлену вчителем;
- виявляє зацікавленість у навчанні, докладає вольових зусиль;
- організовує свою працю для досягнення результату;
- відбирає або знаходить відповідні знання;
- усвідомлює свою діяльність і прагне її вдосконалювати;
- має вміння і навички самоконтролю та самооцінки» [7, 34].

Навчання учнів умінню вчитися є однією з кардинальних проблем дидактики, що на різних етапах її розвитку набула рис або певної дидактичної теорії. До вирішення проблеми «вчити вчитися» ведуть різні шляхи, одним з яких є озброєння учнів у процесі навчання певним комплексом умінь та навичок, що мають міжпредметне значення. У нових умовах розбудови