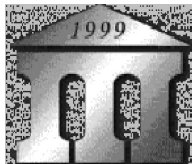


ББК 74.04

П 24



ЦЕНТР ТЕСТУВАННЯ ПРИ МОЗ УКРАЇНИ

РЕКОМЕНДОВАНО

Вченою радою Центрального інституту післядипломної педагогічної освіти АПН України

Рецензенти:

В. В. Олійник — доктор педагогічних наук, професор, ректор Центрального інституту післядипломної педагогічної освіти АПН України, член-кореспондент АПН України

Н. Г. Протасова — доктор педагогічних наук, професор, завідувач кафедри управління освітою Національної академії державного управління

П 24 **Педагогічне оцінювання і тестування.** Правила, стандарти, відповідальність. Наукове видання / Я. Я. Болюбаш, І. Є. Булах, М. Р. Мруга, І. В. Філончук. — К.: Майстер-клас, 2007.— 272 с.

ISBN 978-966-471-041-8

У пропонованому посібнику подано аналіз сучасного стану процедур із забезпечення якості, якими необхідно керуватися під час зовнішнього оцінювання. Розглянуто також відповідні стандарти, якими визначаються правила, вимоги та критерії до систем освітнього і психологічного тестування та оцінювання. Проаналізовано права і обов'язки учасників процесу оцінювання.

Узагальнено інформацію щодо сучасних підходів до стандартизації педагогічного тестування, ролі стандартизації у забезпеченні справедливого оцінювання. Детально розглядаються стандарти, спрямовані на визначення основних технічних та технологічних вимог до тестів, міжнародні рекомендації з використання тестів і кодекси належної практики та професійної відповідальності у сфері педагогічних вимірювань.

Наведено офіційні переклади основних міжнародних документів зі стандартизації вимог, процедур та відповідальності учасників процесу педагогічного оцінювання і тестування. Книга призначена для вчителів та викладачів вищої школи, а також широкого кола працівників освітянських організацій і установ, які причетні до оцінювання в освіті. Вона буде корисною також особам, які оцінюються, учням, студентам та їхнім батькам.

ББК 74.04

ISBN 978-966-471-041-8

© Центр тестування при МОЗ України, 2007

© ТОВ «Видавництво «Майстер-клас», 2007

© М. К. Линник, дизайн обкладинки, 2007

ЗМІСТ

ВСТУП	4
РОЗДІЛ 1. ОЦІНЮВАННЯ ТА РЕГУЛЮВАННЯ В ОСВІТІ	
Мета регулювання та її досягнення	8
Регулювання у сфері оцінювання	9
Регулювання та уряд	10
Процес регулювання	12
Висновки	13
РОЗДІЛ 2. СТАНДАРТИЗАЦІЯ ТА СПРАВЕДЛИВЕ ОЦІНЮВАННЯ	
Місце стандартизації у забезпеченні справедливого оцінювання	17
Політика з оцінювання та тестування	17
Учасники процесу оцінювання (тестування)	18
Документи із стандартизації	20
Базові національні стандарти тестування США та Канади	20
Базовий національний кодекс тестування США	26
Міжнародні документи з використання тестів	28
Вимоги до професійної відповідальності при оцінюванні	37
Висновки	39
РОЗДІЛ 3. ЛІЦЕНЗУВАННЯ, СЕРТИФІКАЦІЯ, РЕЄСТРАЦІЯ	
Нова філософія на ринку освітніх послуг	44
Система ліцензування і сертифікації компетентності	46
Принципи функціонування Центрів сертифікації	47
Висновки	55
РОЗДІЛ 4. ВИМОГИ ДО ПРОФЕСІЙНОЇ КОМПЕТЕНТНОСТІ ВИКЛАДАЧІВ	
Професійні вимоги до вчителя	58
Консорціум з підтримки та оцінювання вчителів-початківців	61
Національна Рада з акредитації програм підготовки викладачів	72
Національна Рада із стандартів у галузі професійного викладання	75
Особливості ліцензування викладацької діяльності у США	79
Ліцензійні педагогічні тести, що використовуються у США	83
Висновки	87
ДОДАТКИ	
A. Міжнародні документи з регулювання педагогічного оцінювання	90
Стандарти освітнього та психологічного тестування	90
ETS Стандарти якості та справедливості	95
Рекомендації з освітнього та психологічного тестування	97
Кодекс справедливого тестування в освіті	98
Рекомендації Міжнародної комісії з тестування	107
EFPA. Модель аналізу та оцінювання психологічних тестів	140
B. Інші регуляторні документи	142
Кодекс професійної відповідальності при педагогічних вимірюваннях	142
Стандарти компетентності вчителів з оцінювання учнів	157
Стандарти компетентності адміністраторів у галузі освіти з оцінювання учнів	166
C. Англomовні варіанти стандартів	174
ITC Guidelines on Adapting Tests	174
ITC International Guidelines for Test Use	177
ITC International Guidelines on Computer-Based and Internet Delivered Testing	204
Code of Professional Responsibilities in Educational Measurement	241
Standards for Teacher Competence in Educational Assessment of Students	253
Standards for Teacher Competence in Educational Assessment of Students	257
Competency Standards in Student Assessment for Educational Administrators	262
ЛІТЕРАТУРА	269

ВСТУП

Належні правила — належний результат

Складання тестового іспиту для отримання посвідчення водія чи з метою професійної сертифікації, тестування для вступу до вищого навчального закладу в країнах з розвинутою економікою є частиною суспільного життя. Складаючи тестовий іспит, усі учасники цього процесу хочуть бути впевнені у тому, що він проводитиметься справедливо, а його результати будуть чесними. Цього прагнуть і ті, хто тестується, і ті, хто використовує результати тестування.

Таким чином, усі учасники процесу тестування та споживачі його результатів, а це передусім: розробники тесту, користувачі тесту та екзаменовані мають приділяти увагу політиці тестування та забезпеченню рівних можливостей і прав екзаменованих.

Політика тестування розробляється для того, щоб:

- гарантувати досягнення мети особистої та корпоративної;
- гарантувати уникнення потенційного використання не за призначенням;
- продемонструвати прихильність до належної практики;
- гарантувати використання тесту відповідно до його мети;
- гарантувати, що тест не використовується для несправедливої дискримінації;
- гарантувати оцінювання на основі повної та доцільної інформації;

- гарантувати використання тестів лише кваліфікованим персоналом.

Політика тестування має забезпечити:

- належне використання тесту;
- охорону та безпеку тестових матеріалів та результатів;
- вимоги до персоналу, який може адмініструвати тести, розраховувати тестові бали та інтерпретувати результати тестування;
- конфіденційність доступу до результатів тестування та розрахунку тестових балів;
- кваліфікаційні вимоги для користувачів тестів та їхній тренінг;
- підготовку екзаменованих до тестування та зворотний зв'язок щодо повідомлення результатів;
- відповідальність перед екзаменованими перед тестуванням, під час тестування та після нього;
- відповідальність та підзвітність кожного окремого користувача.

Кожна політика має регулярно переглядатися та оновлюватися відповідно до прогресу в тестуванні чи змін у практиці. Зацікавлені сторони повинні мати доступ до політики тестування та бути поінформованими про неї.

Відповідальність за політику тестування несуть усі його учасники.

В освітньому та психологічному тестуванні задіяні, а отже, зазнають впливу окремі особи, заклади та суспільство в цілому. Особи, які зазнають впливу, це — студенти, батьки, викладачі, адміністратори у сфері освіти, кандидати на працевлаштування, працівники, пацієнти, керівники всіх рангів та ті, хто проводить оцінювання (evaluators). Заклади, які зазнають впливу, це — школи, коледжі, бізнес-структури, виробництва та державні установи.

Окремі особи та заклади отримують користь тоді, коли тестування служить для досягнення їхньої мети. Суспільство отримує користь тоді, коли досягнення цілей окремими особами та закладами спрямоване на загальне благо. Інтереси різних сторін у процесі тестування загалом, але не завжди, збігаються. Наприклад, якщо тестування проводиться з метою консультативної допомоги чи працевлаштування, інтереси окремих осіб та закладів часто збігаються. Проте,

коли тест використовується для відбору або для визначення рівня досягнень при сертифікації чи ліцензуванні, інтереси задіяних осіб можуть не співпадати.

Критерієм має бути відповідальність кожного учасника тестування за дотримання правил та стандартів педагогічних вимірювань, одержані результати та прийняті оціночні рішення.

Наведені у книзі документи є, по суті, спробою винесення на обговорення освітян України етико-правової проблеми, яка стосується вимог фактично до всіх педагогічних працівників, їхньої компетентності та відповідальності за виконання одного з етапів навчального процесу — педагогічного оцінювання.

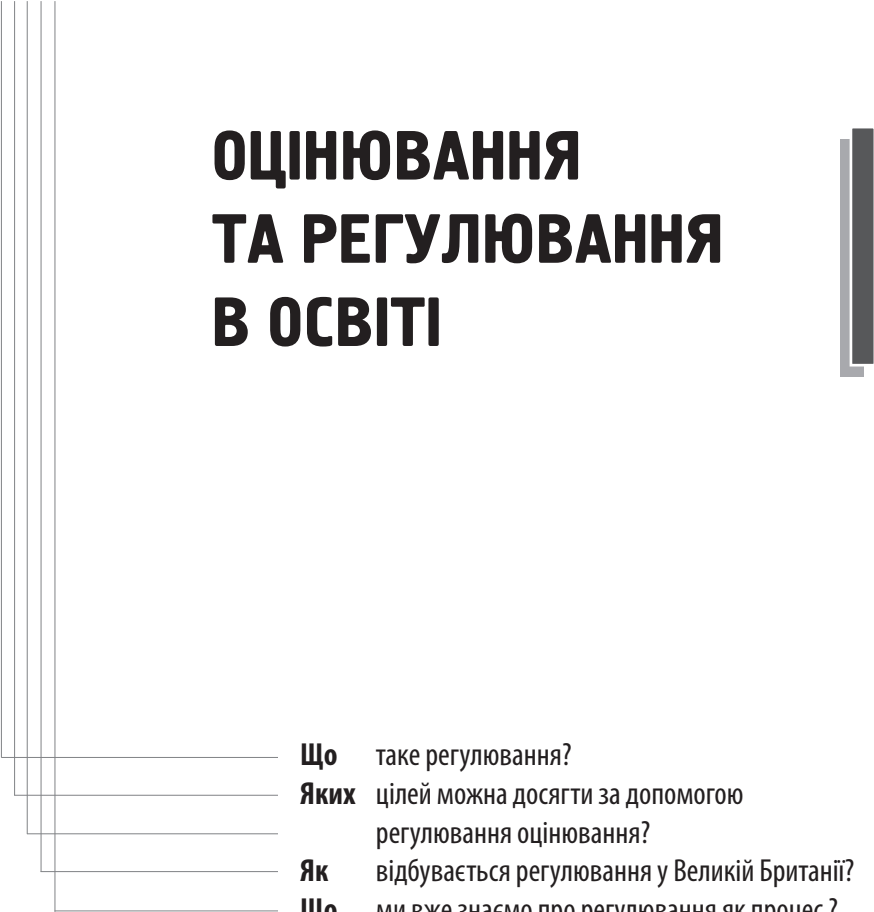
Сподіваємося, що публікація міжнародних регуляторних документів стимулюватиме до дискусії про складові професійно відповідальної практики оцінювання на всіх освітніх рівнях та для всіх її учасників.

Автори



1.

РОЗДІЛ



ОЦІНЮВАННЯ ТА РЕГУЛЮВАННЯ В ОСВІТІ

- Що таке регулювання?
- Яких цілей можна досягти за допомогою регулювання оцінювання?
- Як відбувається регулювання у Великій Британії?
- Що ми вже знаємо про регулювання як процес?

Розглянемо для прикладу функціонування регуляторного органу Великої Британії — Управління з питань кваліфікації, навчальних планів і програм¹ (QCA). Але, по-перше, дамо відповідь на запитання: що таке регулювання?

Наведені у книзі матеріали базуються на доповіді Дж. Вілмута, що представлена на експертному семінарі-наradі зі стандартизації оцінювання, який проводився в межах проекту Європейської Комісії TEMPUS/TACIS «Справедливе оцінювання» (Ноттінгем, 2006).

МЕТА РЕГУЛЮВАННЯ ТА ЇЇ ДОСЯГНЕННЯ

Регулювання — це система контролю за діяльністю інших осіб, установ чи організацій з боку регуляторних органів. Як правило, цей процес здійснюється в інтересах суспільства та з метою захисту інтересів користувачів послуг.

Регуляторною діяльністю займаються відповідні органи, які фактично виступають як третя сторона між тими, хто надає, та тими, хто споживає певні послуги або товари.

Цілями регулюючих органів є:

- збільшення довіри до послуги, що регулюється;
- поліпшення публічного розуміння функціонування об'єктів, які регулюються;
- забезпечення адекватності, ефективності та виправдання витрачених коштів через контроль над цінами та якістю;
- здійснення заходів для досягнення та підтримання високих стандартів;
- розгляд спірних питань.

Регуляторні органи, крім контролюючої, можуть виконувати розви-

¹ Управління з питань кваліфікації, навчальних планів і програм — *Qualifications and Curriculum Authority (QCA)*.

ваючу функцію, безпосередньо впливаючи не лише на якість кінцевого продукту (послуги), а й на процес її створення.

Існує три види регулювання: мінімальне, співрегулювання та саморегулювання.

Мінімальне регулювання — це найменше зовнішнє втручання, необхідне для правильного оцінювання.

Співрегулювання — це ситуація, коли структури, що регулюються, можуть регулювати також і самих себе у певних видах діяльності та деяких функціях. При цьому мають місце певні обмежувальні фактори для регулятора.

Саморегулювання виконується самостійно органами, які підзвітні безпосередньо суспільству.

Регулювання розрізняється за трьома факторами:

- 1) за масштабом;
- 2) за сферою відповідальності;
- 3) за становищем стосовно уряду.

Регулювання у сфері оцінювання

У сфері оцінювання може регулюватися зовнішнє або кваліфікаційне оцінювання, усі види оцінювальної діяльності, оцінювання в межах навчального плану і програм. Також регулювання оцінювання може бути лише однією зі складових загального управління освітою та її якістю.

Необхідно чітко розуміти різницю між тими, хто проводить оцінювання, та тими, хто його регулює.

Існують такі сфери відповідальності регуляторних органів:

- 1) регуляторні органи мають можливість визначати зв'язок між оцінюванням та навчальним планом і програмою;
- 2) регулювати процедури, пов'язані з оцінюванням;
- 3) стандартизувати систему оцінювання;
- 4) розглядати скарги щодо оцінювання;
- 5) надавати поради уряду з питань оцінювання.

Регулювання та уряд

Регуляторний орган може бути незалежним від уряду або близьким до політико-творчого процесу, але не бути частиною його, або входити до складу Міністерства (освіти), галузевого міністерства.

Незалежному регуляторному органу (той, що працює незалежно від уряду) легше радитися з широким колом експертів та надавати поради уряду. Процес роботи такого регулятора є безперервним, навіть тоді, коли політика уряду змінюється.

Для прикладу розглянемо систему регулювання у Великій Британії.

Регулювання оцінювання у Великій Британії

Управління з питань кваліфікації, навчальних планів і програм регулює ці питання в Англії, Уельсі та Північній Ірландії і несе відповідальність за створення навчальних програм в Англії.

Шотландське Управління з питань кваліфікації² (SQA) відповідає за розробку навчальних програм та регулює порядок отримання кваліфікації в Шотландії.

В Уельсі аналогічне Управління³ (ACCAC) відповідає за розробку навчальних програм та деякі аспекти оцінювання.

У Північній Ірландії діє аналогічна Рада⁴ (CCEA), що відповідає за розробку навчальних програм та деякі аспекти оцінювання.

Сфера діяльності регуляторного органу

Можна вирізнити три рівні у сфері діяльності регуляторного органу.

I рівень — регулювання окремих складових оцінювання

До цього рівня входять перевірка та спостереження за індивідуальною кваліфікацією, за окремими екзаменами або тестами.

II рівень — регулювання організації

На цьому рівні регулювання надає гарантію того, що організація, яка здійснює оцінювання, виконує поставлені перед нею завдання, сис-

² Шотландське Управління з питань кваліфікації — *Scottish Qualifications Authority (SQA)*.

³ Управління з питань кваліфікації, навчальних планів і програм та оцінювання Уельсу — *Qualifications, Curriculum & Assessment Authority for Wales (ACCAC)*.

⁴ Рада з навчальних планів і програм, екзаменів та оцінювання Північної Ірландії — *Council for the Curriculum, Examinations and Assessment in Northern Ireland (CCEA)*.

теми контролю — на своєму місці, а також того, що регулювання забезпечує визначення ступеню ризику при оцінюванні.

III рівень — регулювання системи в цілому

На третьому рівні регулювання сприяє перегляду функціонування та послідовності системи кваліфікації, визначенню ступеня ризику появи проблем у процесі розвитку освіти зокрема і суспільства взагалі.

Регулювання в освіті, яке здійснює уряд Великої Британії, реалізовано на III рівні та впливає на розвиток усієї системи освіти в країні. Воно базується на таких п'яти принципах.

Пропорційність. Ступінь втручання регуляторних органів має залежати від ступеня ризику.

Підзвітність.

Послідовність. Цей принцип реалізується у використанні критеріїв, в інформаційних запитах, у винесенні суджень.

Прозорість. Процеси та діяльність є відкритими та прозорими.

Досягнення мети. Заходи, що здійснюються, мають відповідати поставленій меті.

Для прикладу наведемо мету Управління з питань кваліфікації, навчальних планів і програм як регуляторного органу: «Ефективно регулювати відповідні органи, кваліфікації, екзамени та національні тести за навчальним планом з метою забезпечення їхньої чесності, надійної стандартизації та діяльності Комісії як публічного захисника учня».

Процеси та висновки Управління мають бути відкритими та підзвітними громадськості. Виходячи з головної мети цього Управління як регуляторного органу, до програмних завдань його діяльності входять:

- регулювання більшості кваліфікацій (у тому числі всіх, які фінансуються державою);
- заохочення осіб віком від 14 до 19 років до отримання кваліфікації та набуття відповідних освітніх ступеней;
- забезпечення дотримання стандартів та інформування населення;
- ефективне регулювання витрат;
- зменшення кількості відповідних органів, що присуджують дипломи, та збільшення довіри до їхніх внутрішніх підходів до забезпечення якості;

- спрощення системи кваліфікацій, яка має бути зрозумілою;
- забезпечення діяльності Управління як якісного регуляторного органу.

Регулювання і якість

Регуляторний орган визначає фактори, які показують, що оцінювання є задовільним. Здійснює контроль за якістю. А також має системи, які забезпечують відповідність стандартам якості, тобто системи забезпечення якості.

ПРОЦЕС РЕГУЛЮВАННЯ

На схемі зображено основні об'єкти, суб'єкти та процеси регулювання на прикладі регулювання кваліфікацій.

Підсумовуючи, можна назвати також додаткову користь, яка отримується за рахунок регулювання в системі оцінювання. Її зумовлено такими чинниками.

1. Існує громадська впевненість, що зміст кваліфікацій та ступенів відповідає реальності.
2. Регулювання контролюється урядом.
3. Суспільство розуміє, у чому полягає система оцінювання.
4. Регулятор є професійним радником та експертом уряду.
5. Регулювання відбувається послідовно та адекватно.
6. Ті, хто навчаються, є захищеними під час змін та реформ.
7. Існує контроль над витратами та прибутками.
8. Регулювання зменшує ризик можливих невдач.
9. Існує професійна реакція на можливу невдачу.
10. Регулятор підтримує або координує загальну дослідницьку діяльність та діяльність, пов'язану з розвитком оцінювання зокрема та освіти взагалі.

Регулювання сприяє підвищенню рівня бюрократії та обсягів роботи з паперами. Стосунки між регулятором та тими, чия діяльність він регулює, можуть бути складними.

Водночас, регулювання призвело до більшої прозорості та послідовності в оцінюванні. Ймовірно, існує вищий рівень громадського розуміння та довіри до отримання кваліфікації.

Регулятори Великої Британії провели значну дослідницьку роботу. Вони стали центром дискусій експертів з питань, пов'язаних з оцінюванням. Регулятори встановили зв'язки з великою кількістю зацікавлених сторін та створили консультаційний форум.

Деякі запитання з регулювання

Чи є процедура оцінювання, що застосовується вами, та її результати відкритими для громадських спостерігачів і зрозумілими для тих, хто навчається?

Як ви визначаєте рівень громадської довіри до систем оцінювання? Він збільшується чи, навпаки, зменшується?

Чи знаєте ви, скільки коштує процес установлення рівня кваліфікації?

В якому напрямку ви розвиваєте плани і цілі для зменшення витрат на оцінювання?

Який зв'язок існує між відповідними органами, регуляторами та урядом? Чи є регулювання незалежним від уряду?

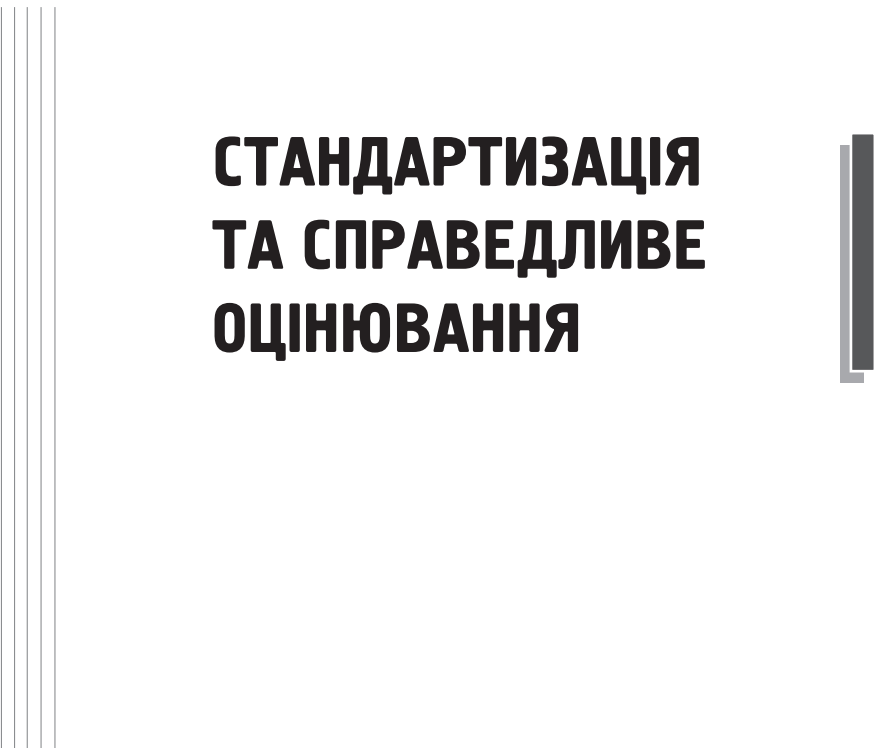
Яка організація має високу репутацію як джерело досвіду: в освітній спільноті, в уряді?

Хто оцінює ступінь ризику та відповідає за його зменшення? Чи існують системи, які забезпечують прийняття неупереджених професійних рішень при виникненні проблем?



2.

РОЗДІЛ



СТАНДАРТИЗАЦІЯ ТА СПРАВЕДЛИВЕ ОЦІНЮВАННЯ

- 
- Роль** стандартизації у забезпеченні справедливого оцінювання.
 - Учасники** процесу оцінювання та вимоги до їхньої компетентності.
 - Міжнародні** документи зі стандартизації.

Мета розділу: узагальнити інформацію про сучасні підходи до стандартизації освітнього тестування та вимог до учасників цього процесу.

Освітнє та психологічне тестування являє собою один із найвагоміших внесків біхевіоризму в розвиток суспільства західних країн. Воно забезпечило суттєве поліпшення порівняно з попередніми практиками на виробництві, в уряді та освіті. Воно дало інструмент для ширшого та більш справедливого доступу до освіти та зайнятості. Хоча не всі тести добре розроблені, не всі підходи до тестування раціональні та корисні, наявні докази підтверджують, що належне використання добре сконструйованих та валідних тестів дає кращу базу для прийняття певних важливих рішень щодо окремих осіб та програм, ніж інші методи педагогічної діагностики.

Освітнє та психологічне тестування у західних країнах завжди було об'єктом прискіпливого розгляду, критики та дебатів як поза межами, так і в колі професійної тестової громади. Найпоширеніші критичні зауваження полягають у тому, що тести відіграють занадто велику роль у житті студентів та працівників (students and employees). Будучи необ'єктивними, вони створюють зайві перешкоди. З урахуванням цих та інших критичних зауважень у книзі наводяться регуляторні документи, мета яких — створити засади для оцінки якості тестових практик в аспекті впливу на різних учасників цього процесу.

Необхідно підкреслити вживання терміна «тест» у регуляторних документах, наведених у книзі.

Термін «тест», як правило, вживається стосовно вимірів різних проявів діяльності (constructed performance sample type) або структурованої поведінки (structured behavior sample type), під час яких екзаменовані повинні (за власним розумінням чи інструкцією) намагатися показати свій найкращий результат. Тоді як інструменти для визначення

інтересів та характеристик особистості шляхом самооцінювання типово і правильно називати формами опитування чи перевірки («inventories», «questionnaires» або «checklists»). Але у наведених документах усі інструменти вимірювання задля простоти названо тестами.

Тест (як інструмент вимірювання) — це сукупність тестових завдань (item), які підібрані за певними правилами для вимірювання певного кількісного показника.

МІСЦЕ СТАНДАРТИЗАЦІЇ У ЗАБЕЗПЕЧЕННІ СПРАВЕДЛИВОГО ОЦІНЮВАННЯ

Метою стандартизації, та, як ми зазначали у попередньому розділі — регулювання, є встановлення критеріїв для оцінювання якості тестів, систем оцінювання та наслідків використання тестів.

Політика з оцінювання та тестування

Усі суб'єкти процесу оцінювання мають дотримуватися вимог, норм та правил тестування й оцінювання. Як засвідчив проведений аналіз основних міжнародних регуляторних документів, ці вимоги, норми, правила, стандарти утворюють так звану **«Політику з тестування та оцінювання»**, яка окреслює межі, шлях та горизонти існування та розвитку не лише освіти, а й суспільства загалом і відображають фактично стан його розвитку.

Будь-яка політика потребує регулярного перегляду та оновлення у відповідь на розвиток тестових технологій, методів оцінювання або зміни у практиці. Усі зацікавлені сторони, які беруть безпосередню участь, а також які мають опосередковане відношення, повинні отримувати інформацію щодо політики з тестування та оцінювання.

РОЗДІЛ. 2

Політика з тестування та оцінювання визначається для того, щоб:

- забезпечити досягнення цілей індивідуумів та організацій;
- запобігти неналежному використанню тестування та оцінюванню;
- продемонструвати дотримання належних підходів;
- забезпечити відповідність результату оцінювання його меті;
- тестування та оцінювання не призводило до дискримінації учасників оцінювання;
- висновки базувалися на всебічній відповідній та належній інформації;
- методи тестування (оцінювання) застосовувались, а результати використовувались лише кваліфікованим та обізнаним персоналом.

Політика з тестування та оцінювання **має охоплювати** такі сфери:

- належне використання тестів та методів оцінювання;
- безпека та захист матеріалів і результатів;
- кваліфікаційні вимоги до осіб, які використовують тестування і оцінювання;
- підготовка користувачів тестів та осіб, що тестуються;
- доступ до тестових результатів та питання конфіденційності тестових балів;
- зворотний зв'язок з екзаменованими;
- відповідальність перед екзаменованими до, під час і після тестування;
- відповідальність та підзвітність кожного, хто використовує методи тестування і оцінювання.

Учасники процесу оцінювання (тестування)

Освітнє та психологічне тестування та оцінювання впливає на **індивідуумів, установи та суспільство** в цілому.

Індивідуумами можуть бути учні, студенти, батьки, учителі, адміністратори в галузі освіти, претенденти на роботу, працівники, пацієнти, інспектори, посадові особи та оцінювачі. Отже, усі суб'єкти,

які безпосередньо або опосередковано мають відношення до оцінювання та його результатів.

Установи — це передусім школи, коледжі, університети, підприємства, промисловість, урядові заклади, тестові організації тощо. Тобто навчальні заклади, що надають освітні послуги і здійснюють підготовку учнів, студентів, фахівців; роботодавці, які наймають фахівців; урядові заклади, які визначають політику та контролюють якість, а отже, відповідають за розвиток суспільства і країни; атестаційні та тестові організації, професійну діяльність яких спрямовано на забезпечення якісного вимірювання та оцінювання.

Індивідууми та установи отримують користь, якщо оцінювання і тестування допомагає їм досягти своєї мети. **Суспільство отримує користь**, коли досягнення індивідуумів та установ сприяють загальному розвитку.

Інтереси *сторін у процесі* тестування є гармонійними, *але не завжди*.

Отже, підсумовуючи можна сказати, що у процесі оцінювання і тестування, як правило, беруть участь **три основних його учасники**:

- розробники тесту (оцінювання), які розробляють, видають, а також адмініструють та обробляють тестові (оціночні) матеріали;
- користувачі тестів (оцінювання), яким потрібні результати тестування (оцінювання) для прийняття певних рішень;
- особи, які оцінюються, — за своїм вибором, за розпорядженням чи за необхідністю.

Також можуть бути ще такі учасники, як:

- замовники (спонсори тесту, оцінювання) розробки певного інструменту або певної послуги, наприклад, організації чи урядові установи;
- адміністратори тесту (оцінювання), які керують реальним проведенням тестування (оцінювання);
- експерти-рецензенти тесту, які здійснюють академічний розгляд і аналіз, оцінюючи відповідність тесту очікуваній меті.

Функції та ролі цих учасників можуть поєднуватися, а інколи — поділятися залежно від завдань, які виконуються у процесі оцінювання.

ДОКУМЕНТИ ІЗ СТАНДАРТИЗАЦІЇ

Ми розглянемо основні міжнародні документи із стандартизації процесу та результату оцінювання і тестування. Ці документи розроблено найвідомішими світовими професійними організаціями, які займаються педагогічними вимірюваннями та оцінюваннями.

Документи є фактично **Кодексом** правил, які повинні бути основою діяльності будь-якої тестової організації при підготовці та проведенні тестувань; урядових та навчальних установ при прийнятті рішень та використанні результатів; екзаменованих та громадськості як активних учасників поширення у суспільстві процесів, що сприяють установленню справедливості.

Представлені документи згруповані у три класи: базові національні стандарти, міжнародні стандарти та стандарти компетентності тих, хто займається педагогічним оцінюванням.

Базові національні стандарти тестування США та Канади

Стисло зупинимося на основних регуляторних нормативних документах, розроблених провідними атестаційними організаціями Сполучених Штатів Америки та Канади. Нижче буде розглянуто такі документи:

1. *Стандарти освітнього та психологічного тестування (1999).*
2. *ETS Стандарти для забезпечення якості та справедливості (2002).*
3. *Рекомендації з освітнього та психологічного тестування (CPA, 1996).*

Детальніше ці стандарти наведено у додатках.

Стандарти освітнього та психологічного тестування

Розробники Стандартів:

Американська асоціація педагогічних досліджень⁵ (AERA)

Американська психологічна асоціація⁶ (APA)

Національна рада з вимірювань в освіті⁷ (NCME)

Документи, використані у розробці «АРА Стандартів»:

- Технічні рекомендації для психологічних тестів і діагностичних методик (APA, 1954).
- Технічні рекомендації для тестів з оцінювання успішності (AERA і NCME, 1955).
- Стандарти для освітніх і психологічних тестів та підручників (APA, AERA і NCME, 1966).
- Стандарти для освітніх та психологічних тестів (APA, AERA і NCME, 1974).
- Стандарти освітнього та психологічного тестування (APA, AERA і NCME, 1985).

Отже, Стандарти освітнього та психологічного тестування⁸ (США, 1999) [1] (далі — АРА Стандарти) є логічним продовженням удосконалення технічних вимог до тестів та тестування, що здійснювалося впродовж останньої половини ХХ сторіччя у Сполучених Штатах Америки.

⁵ Американська асоціація педагогічних досліджень — *The American Educational Research Association (AERA), founded in 1916* – <http://www.aera.net>.

⁶ Американська психологічна асоціація — *American Psychological Association (APA)*, – <http://www.apa.org>.

⁷ Національна рада з вимірювань в освіті — *National Council on Measurement in Education (NCME)* – <http://www.ncme.org/>.

⁸ Стандарти освітнього та психологічного тестування (США, 1999) — *The Standards for Educational and Psychological Testing*.

Відповідно до мети створення АРА Стандарти призначені:

1. Регулювати користування тестами (методами оцінювання) залежно від умов їх застосування.
2. Формулювати технічні стандарти для досконалої професійної практики, а не директиви щодо дій в соціумі.
3. Надавати можливість установлювати: технічну адекватність тесту (методу оцінювання); відповідність та специфічність для конкретних умов застосування; коректність висновків, зроблених на підставі результатів тестування (оцінювання).
4. Вимагати від розробників та користувачів тестів збору та оприлюднення інформації в обсязі, який дозволить експерту визначити, чи досягнуто мети.
5. Виступати **етичним еталоном**, але не містити механізмів примусу.
6. Визнавати, що стандарти не можуть застосовуватись однаково до різних інструментів оцінювання та в різних умовах.
7. Бути представленими на рівні, що дозволяє користуватися Стандартами широкому колу людей, котрі працюють з тестами та результатами тестування.
8. Не гальмувати розвиток методів оцінювання.
9. Відображати сучасний рівень консенсусу між експертами.

Цей документ складається з трьох частин і має такий зміст.

Частина I: Конструювання, оцінювання і документування тесту

1. Валідність.
2. Надійність та помилки вимірювання.
3. Розробка та перегляд тесту.
4. Шкали, норми та порівнюваність балів.
5. Адміністрування тесту, підрахунок балів та звітність.
6. Допоміжна документація для тестів.

Частина II: Справедливість у тестуванні (оцінюванні)

7. Справедливість у тестуванні (оцінюванні) та користуванні тестами.
8. Права та обов'язки осіб, що тестуються (оцінюються).
9. Тестування (оцінювання) осіб різного мовного походження.
10. Тестування осіб із вадами та інвалідністю.

Частина III: Умови застосування тестування (оцінювання)

11. *Обов'язки та відповідальність користувачів тестів (оцінювання).*
12. *Психологічне тестування та оцінювання.*
13. *Освітнє тестування та оцінювання.*
14. *Тестування для працевлаштування та підтвердження дипломів.*
15. *Тестування з метою оцінювання програм та вироблення державної політики.*

Таким чином, цей регулюючий документ визначає вимоги до створення якісного тесту, умов і процедур його використання та правил його застосування для досягнення відповідної мети оцінювання.

ETS Стандарти якості та справедливості**ETS Стандарти розроблені**

Службою освітнього тестування США⁹ (ETS)

Документи, застосовані у розробці цього «Стандарту»:

- ETS Стандарти якості та справедливості (покладені в основу корпоративної політики Наглядовою радою ETS у 1981 р.)
- ETS Стандарти, переглянуті та оновлені у 1987 р.
- ETS Стандарти, переглянуті та оновлені у 2000 році після перегляду Стандартів освітнього та психологічного тестування, опуб-лікованих організаціями AERA, APA та NCME у 1999 р.

ETS Стандарти (США, 2002)¹⁰ [2] близькі до APA Стандартів.

Ці Стандарти відрізняються розділами, що відображають специфіку продукції та послуг ETS, яка є провідною у світі тестовою організацією, що спеціалізується у сфері педагогічного оцінювання.

⁹ Служба освітнього тестування — *Educational Testing Service (ETS)* — <http://www.ets.org>.

¹⁰ ETS Стандарти якості та справедливості (США, 2002) — *ETS Standards for Quality and Fairness*

Нижче наведено структуру змісту ETS Стандартів.

Зміст документа

- Глава 1: Процедури розробки.
- Глава 2: Придатність для використання.
- Глава 3: Послуги для замовників.
- Глава 4: Справедливість.
- Глава 5: Використання та захист інформації.
- Глава 6: Валідність.
- Глава 7: Розробка інструменту оцінювання.
- Глава 8: Надійність.
- Глава 9: Критерій «склав», шкалювання та вирівнювання.
- Глава 10: Адміністрування (проведення) оцінювання.
- Глава 11: Видача результатів оцінювання, звітність.
- Глава 12: Використання оцінювання.
- Глава 13: Права та обов'язки осіб, що тестуються та оцінюються.

Як впливає зі змісту, ці Стандарти визначають також вимоги до створення якісного педагогічного тесту, умов і процедур його використання та правил застосування.

Рекомендації з освітнього та психологічного тестування

Рекомендації розроблені

Канадською психологічною асоціацією¹¹

Документи, застосовані у розробці даних «Рекомендацій» — це документи різних організацій, зокрема APA Стандарти.

Метою Рекомендацій є визначення критеріїв для оцінювання тестів, підходів до тестування і оцінювання та наслідків використання тестів.

Структура Рекомендацій близька до структури APA Стандартів.

Три частини стосуються *технічних рекомендацій* та взаємозв'язку між ними:

- технічними питаннями (особливо Частина I)
- професійною практикою (особливо Частина II)
- професійною етикою (особливо Частина III).

Ці Рекомендації (CPA, 1996)¹² [3] відображають сучасний рівень консенсусу канадських фахівців та адаптують АРА Стандарти у Канаді.

Зміст документа, який, як і для АРА Стандарту, складається з трьох частин, наведено нижче.

Зміст документа

Частина I: Тестовий інструмент (загальні рекомендації)

Валідність.

Надійність та помилки вимірювання.

Розробка та перегляд тесту.

Шкалювання, нормування, порівнюваність та вирівнювання балів.

Видання тестів, технічні інструкції та інструкції для користувачів.

Частина II: Використання тестів (оцінювання)

Загальні принципи використання тестів (оцінювання).

Тестування (оцінювання) осіб з інвалідністю та вадами.

Тестування (оцінювання) і дві офіційні мови Канади.

Освітнє тестування і психологічне тестування у школах.

Використання тестів під час клінічного оцінювання та консультування.

Тестування для працевлаштування.

Професійне та фахове ліцензування і сертифікація.

Оцінювання програм.

Частина III: Адміністративні процедури

Адміністрування (проведення) тесту, підрахунок балів, звітність.

¹¹ Канадська психологічна асоціація — *The Canadian Psychological Association (CPA)*, заснована в 1939 році, <http://www.cpa.ca>.

¹² Рекомендації з освітнього та психологічного тестування (CPA, 1996) — *Guidelines for Educational and Psychological Testing*

Базовий національний кодекс тестування США

Кодекс справедливого тестування в освіті

ЖСТР Кодекс розроблено

Об'єднаним комітетом з практики тестування¹³ (ЖСТР)

Документи, застосовані у розробці цього «Кодексу»:

- Кодекс справедливих тестових практик в освіті (1988)
- Стандарти освітнього та психологічного тестування (APA, AERA і NCME, 1999)

Цей документ за метою та призначенням належить до іншого типу регулюючих документів порівняно з розглянутими раніше стандартами. І називається він уже не Стандарт, а має назву Кодекс справедливого тестування в освіті (США, 2004)¹⁴, далі — ЖСТР Кодекс.

Мета ЖСТР Кодексу [4] — представити сутність APA Стандартів [1] так, щоб вони були значимі та зрозумілі для всіх учасників процесу тестування та оцінювання.

Кодекс побудовано як довідник — це допоможе фахівцям використовувати тести в освіті таким чином, щоб і процес, і його результати були справедливими до всіх тестованих осіб. **Справедливість (Fairness)** — первинна категорія в усіх аспектах тестування, і саме на її забезпеченні акцентується увага у Кодексі.

Кодекс можна широко застосовувати для тестування і оцінювання в освіті. Це перш за все вступний відбір, педагогічне оцінювання, педагогічна діагностика, професійна орієнтація учнів та студентів, діагностування професійної компетентності фахівців.

Кодекс призначений переважно **для професійно розроблених тестів: (інструментів оцінювання)**, що використовуються в офіційних про-

¹³ Об'єднаний комітет з практики тестування — *Joint Committee on Testing Practices (JCTP)* — <http://www.apa.org/science/jctpweb.html>, засновано у 1985 р. організаціями AERA, APA і NCME.

¹⁴ Кодекс справедливого тестування в освіті (JCTP, США, 2004) — *Code of Fair Testing Practices in Education*.

грамах оцінювання. Кодекс не призначений **для тестів (інструментів оцінювання), створених учителями** для власного використання у класі.

Положення Кодексу можуть застосовуватися як при використанні існуючих тестів, так і при розробці нових, або коли змінюються умови проведення тестування. Кодекс дає окремі рекомендації як для **розробників тестів**, так і для **користувачів тестів** у таких чотирьох критичних сферах:

- A. Розробка та вибір належних тестів (інструментів оцінювання).
- B. Адміністрування тестів (оцінювання) і розрахунок балів.
- C. Звітність та інтерпретація результатів.
- D. Інформування осіб, які оцінюються.

Нижче наведено фрагмент змісту Кодексу із збереженням його структури (табл. 2). Повний текст Кодексу наведено у додатку.

Таблиця 2. Фрагмент розділу A. Розробка та вибір належних тестів

РОЗРОБНИКИ ТЕСТІВ	КОРИСТУВАЧІ ТЕСТІВ
1	2
Розробники тестів повинні надавати інформацію і підтверджуючі докази, які користувачі потребують для вибору належних тестів.	Користувачі тестів повинні вибирати такі тести, що відповідають запланованій меті та які підходять до цільових екзаменованих.
A-1. Надати докази того, що вимірює тест, цільових сфер використання, цільових екзаменованих, та сильні сторони і обмеження тестів, у т.ч. точність тестових балів.	A-1. Визначити мету тестування, зміст і вміння, що оцінюються, цільових екзаменованих. Вибір і використання найприйнятнішого тесту на основі глибокого аналізу інформації.
A-2. Описати, яким чином відбирались зміст і вміння, що оцінюються, і яким чином розроблялись тести.	A-2. Переглянути і вибрати тест на основі прийнятності змісту і вмінь, що оцінюються, відповідності змісту визначеній меті.
A-3. Повідомляти інформацію щодо характеристик тесту з деталізацією, що відповідає цільовим користувачам тестів.	A-3. Переглядати матеріали, надаті розробниками тестів, і вибирати тести, щодо яких є чітка, точна та повна інформація.

Міжнародні документи з використання тестів

У цьому розділі ми стисло розглянемо документи Міжнародної комісії з тестування¹⁵ (ITC). ITC — це «Асоціація національних психологічних асоціацій, тестових комісій, видавців та інших організацій, діяльність яких спрямована на впровадження ефективної політики тестування та оцінювання і на належний розвиток, оцінку та використання освітніх та психологічних інструментів.» (Директива ITC, 2001). Асоціація була формально заснована 1978 року Жаном Кардіне (Jean Cardinet) і зараз нараховує 24 повних члени (національні професійні психологічні асоціації), 51 асоційований член (інші тестові комісії, видавці та дослідницькі організації, що займаються тестуванням) та понад 50 індивідуальних членів (особи, які працюють у сфері тестів та тестування чи зацікавлені у цьому). До неї на сьогодні входять представники більшості країн Західної та Східної Європи і Північної Америки, а також представники деяких країн Середнього та Далекого Сходу, Південної Америки та Африки. Вона зареєстрована як неприбуткова організація при IAAP та IUPsyS. Міжнародна Комісія з Тестування сприяє обміну інформацією серед її членів та стимулює їх співпрацю з проблем, пов'язаних з побудовою, розповсюдженням та використанням психологічних тестів та інших психодіагностичних засобів. Для досягнення цих цілей ITC організовує міжнародні зустрічі та дискусії, сприяє публікації відповідної інформації своїми власними силами та за рахунок інших наукових осередків, стимулює міжнародну співпрацю у дослідницьких проектах з належного використання тестів з точки зору науки та етики.

Проекти ITC є прямим утіленням цілей ITC. Результатом трьох нещодавніх проектів стало затвердження Рекомендацій, які отримали широке міжнародне визнання. Крім того, ми стисло зупинимося на міжнародному регуляторному документі Європейської федерації психологів — Модель Аналізу та Оцінювання Психологічних Тестів.

Нижче будуть розглянуті такі документи:

- *Міжнародні Рекомендації ITC з використання тестів (2000);*
- *Міжнародні Рекомендації ITC з адаптації тестів (1996);*
- *Міжнародні Рекомендації ITC з комп'ютерного та інтернет тестування (2005);*

¹⁵ Міжнародна комісія з тестування — *The International Test Commission (ITC)* — <http://www.intestcom.org>. ITC є неакціонерною компанією, зареєстрованою в США у 1978 р.

— EFPA¹⁶ *Модель аналізу та оцінювання психологічних тестів* (2005).
Більш детально ці документи подано в Додатках.

Міжнародні рекомендації ІТС з використання тестів

Розроблено Міжнародною комісією з тестування — ІТС

Документи, застосовані у розробці цього документа:

- Стандарти освітнього та психологічного тестування (APA, AERA та NCME, 1985).
- Рекомендації з освітнього та психологічного тестування (CPA, 1987);
- Кодекс справедливого тестування в освіті (JCTP, 2000);
- Обов'язки користувачів стандартизованих тестів (Американська асоціація для консультування та розвитку — AACD, 1992);
- BPS Стандарти використання професійно-технічних тестів рівня А і рівня В (Британська психологічна спілка — BPS, 1995 та 1996);
- Міжнародні рекомендації ІТС з використання тестів (ІТС, 1995);
- Додаток до Рекомендацій з використання психологічних тестів (Австралійська психологічна спілка — APS, 1997).

У практиці психологічного тестування існують дві ключові проблеми. **По-перше**, необхідно переконатися, що наявні тести відповідають стандартним необхідним мінімальним технічним вимогам. **По-друге**, необхідно знати, що люди, які їх використовують, компетентні та знають як це робити.

¹⁶ Європейська федерація асоціацій професійних психологів — *The European Federation of Professional Psychologists Associations (EFPA)* – <http://www.efpa.be/>. Зустрічається також під аббревіатурою EFPPA – <http://www.cop.es/efppa>. Була заснована в Німеччині в 1981 році.

¹⁷ Британська психологічна спілка — *British Psychological Society (BPS)* – <http://www.bps.org.uk/>. Була заснована в Британії в 1901 році.

Рекомендації¹⁸ [5] є спільною працею фахівців з психологічного та освітнього тестування різних країн: Бельгії, Данії, Англії, Франції, Німеччини, Нідерландів, Норвегії, Словенії, Південної Африки, Іспанії, Швейцарії, США. Це проект Міжнародної Комісії з тестування, який має на меті належне використання тестів та розповсюдження якісних практик у психологічному та освітньому тестуванні. Робота, проведена ІТС для популяризації належної практики адаптації тестів була важливим кроком уперед у забезпеченні стандартизації якості тестів, адаптованих для використання в різних культурних та мовних середовищах.

Проект з розробки Рекомендацій з використання тестів почався з пропозиції до Ради ІТС у 1995 році. Метою було визначити загальні міжнародні рамки, у межах яких можна розробляти певні місцеві (національні) стандарти, кодекси практик, кваліфікації, критерії реєстрації користувачів тощо, для задоволення місцевих потреб. Отже, Рекомендації мали сприяти здійсненню таких завдань:

1. Розробити набір рекомендацій стосовно компетентності, що вимагаються від користувачів тестів.
2. Охопити всі сфери використання тестів.
3. Провести широкі консультації, залучаючи всіх зацікавлених осіб.
4. Сформулювати Рекомендації у формі оціночних тверджень щодо результату.

Проект «Міжнародні Рекомендації ІТС з використання тестів» отримав підтримку від BPS, APA, NCME, EAPA¹⁹, EFPPA²⁰ і від значної кількості європейських та американських видавців тестів. Після офіційного схвалення Радою ІТС у Гаазі у 1999 році та Постійним Комітетом EFPPA з тестів та тестування²¹ (EFPPA Standing Committee on Tests and Testing) у Римі у 1999 році, Рекомендації були введені в дію в Стокгольмі, на Конгресі ІТС у липні 2000 року.

¹⁸ Міжнародні Рекомендації ІТС з використання тестів (2000) — *ITC International Guidelines for Test Use*.

¹⁹ Європейська Асоціація з психологічного оцінювання — *European Association of Psychological Assessment (EAPA)* — <http://www.uam.es/centros/psicologia/paginas/eapa/>. Була заснована в Іспанії в 1990 році.

²⁰ Європейська Федерація асоціацій професійних психологів — *European Federation of Professional Psychologists Associations (EFPPA)* — <http://www.cop.es/efppa>. Зустрічається також під аббревіатурою EFPA — <http://www.efpa.be/>. Була заснована в Німеччині в 1981 році.

На сьогодні вже здійснено офіційні переклади з англійської такими мовами: голландською/фламандською, хорватською, французькою, німецькою, литовською, норвезькою, португальською, словенською, іспанською та португальською Південної Америки, іспанською, шведською, датською, китайською.

Метою документа є не «винахід» нових Рекомендацій, а зведення та узагальнення ключових принципів, що містяться в існуючих рекомендаціях, кодексах практики, стандартах та інших відповідних документах, і створення гармонійної структури, у межах якої Рекомендації будуть зрозумілими та використаними.

Основний акцент у Рекомендаціях зроблено на належному використанні тестів та сприянні належній практиці в оцінюванні.

Рекомендації, які фактично є міжнародним корпоративним документом, на думку розробників, мають виконувати декілька завдань.

По-перше, у цьому документі наголос зроблено на необхідності враховувати вимоги та потреби різних зацікавлених сторін, які поділяються на три великі категорії:

- *ті, хто задіяний у **підготовці тестів** (наприклад, автори тестів, видавці та адміністратори);*
- ***споживачі тестів** (наприклад, користувачі тестів, екзамновані та інші треті сторони — батьки, опікуни тощо);*
- *ті, хто **впроваджує правила тестування** (наприклад, професійні тестові організації, урядові структури та законодавці).*

Отже, Рекомендації спрямовано переважно на:

- *тих, хто відповідає за вибір та використання тестів;*
- *тих, хто адмініструє, підраховує бали та інтерпретує тести;*
- *тих, хто задіяний у процесі вироблення звітів і встановлення зворотного зв'язку;*
- *розробників тестів;*

²¹ Постійний Комітет EFPPA з тестів та тестування — EFPPA Standing Committee on Tests and Testing

РОЗДІЛ. 2

- тих, хто задіяний у підготовці користувачів тестів;
- тих, хто тестується, та осіб, пов'язаних з ними (батьки, роботодавці, партнери);
- професійні тестові організації;
- політиків та законодавців.

По-друге, Рекомендації мають на меті також допомогти створенню національних Стандартів з використання тестування у країнах, які починають запроваджувати об'єктивні методи оцінювання в освіті.

Рекомендації повинні використовуватися для розробки **національних Стандартів** шляхом їх контекстуалізації.

Фактори, що мають ураховуватися при застосуванні Рекомендацій у розробці національних Стандартів, включають:

- соціальні, політичні, лінгвістичні та культурні розбіжності;
- існуючі національні рекомендації та практично-орієнтовані стандарти;
- розбіжності, пов'язані з умовами тестування (освітні, клінічні, на роботі та інші);
- різні підходи у споживачів результатів тестування (екзаменовані чи їх батьки, розробники тестів, роботодавці або інша третя сторона);
- розбіжності, пов'язані з використанням результатів тестування для прийняття рішень;
- різний ступінь можливостей (традиції, досвід) перевірки достовірності результату та практики щодо застосування методик вирівнювання.

По-третє, одним із завдань проекту ІТС є створення рекомендацій, які визначають **компетентності, необхідні користувачам тестів** (знання, уміння, здібності та інші особисті характеристики). Компетентності, визначені Рекомендаціями, повинні бути представлені у вигляді вимірюваних критеріїв успішності. Ці критерії складають основу для розробки специфікацій ознак наявності компетентностей, очікуваних від особи, що претендує на статус користувача тестів.

Такі компетентності мають охоплювати такі питання, як:

- професійні та етичні стандарти у тестуванні;
- права екзаменованих та інших сторін, задіяних у процесі тестування;
- вибір тестів та оцінювання їх якості;
- адміністрування тестів, підрахунок балів та інтерпретація;
- написання звітів та зворотний зв'язок.

Стратегія розробки документа передбачала вирішення питань якості використання тестів та особливостей оцінювання в першу чергу через визначення компетентності користувачів тестів. Це не скасовує потреби вирішення питань якості тестів. Проте підхід, який наголошує на компетентності користувачів, дає більше гарантії, що користувачі:

- не будуть використовувати погані тести;
- знають, коли тестувати, а коли ні;
- знають, які тести доцільні, а які ні;
- знають, як застосовувати тести;
- знають, як ставитися до екзаменованих.

Зміст документа

Структура Рекомендацій складається з 24 основних секцій та розрізняє основні аспекти компетентності:

- знання, розуміння та вміння, котрі належать до процесу тестування: що користувач тестів повинен уміти зробити;
- знання та розуміння, які потрібні для підтримки процесу тестування.

Рекомендації містять:

- ключову мету та сферу дії тверджень;
- детальний опис компетентності користувача тестів, пов'язаної з етичним застосуванням тестів;
- детальний опис компетентності користувача тестів, пов'язаної з **належною практикою** у використанні тестів.

Наведемо стисло перелік основних вимог до знань, розумінь та вмінь, які включають різні аспекти компетентності, визначені у Рекомендаціях.

Знання, розуміння та вміння, що лежать в основі компетентності користувача тестів.

Відповідні декларативні знання, зокрема:

- знання базових принципів та методик психометрії та технічних вимог щодо тестів (надійність, валідність, стандартизація);
- знання тестування (оцінювання) та вимірювання у достатньому обсязі, який дозволяє правильно розуміти результати тестування;
- знання та розуміння відповідних теорій і моделей здібностей особистості та інших психологічних конструктів;
- перелік тестів та їх розробників для відповідної професійної сфери.

Відповідні інструментальні знання та вміння, зокрема:

- знання конкретних методів та інструментів оцінювання, у тому числі використання методів комп'ютерного тестування і оцінювання;
- спеціалізовані знання та практичні вміння, пов'язані з використанням спеціалізованих тестів;
- знання та розуміння конструкту або конструктів, що лежать в основі тестових балів.

Відповідні загальні особистісні задачно-орієнтовані вміння, зокрема:

- виконання заходів з адміністрування тесту, підготовки звітів і забезпечення зворотного зв'язку з екзаменованими та іншими клієнтами;
- навички усного та письмового спілкування, достатні для забезпечення належної підготовки екзаменованих та взаємодії з відповідними сторонами (батьки чи особи, які визначають стратегію організації).

Відповідні контекстні знання та вміння, зокрема:

- знання того, коли варто застосовувати тест, а коли ні;
- розуміння, як треба інтегрувати тестування з іншими, менш формалізованими, складовими оцінювання (біографічні дані, неструктуроване інтерв'ю);

- знання сучасних професійних, нормативних та етичних проблем, пов'язаних із застосуванням тестів, та їх практичний вплив на застосування.

Відповідні вміння з управління завданнями, зокрема:

- знання кодексів поведінки та належної практики, що належать до застосування тестів, тестових даних, виготовлення та збереження звітних форм, збереження та відповідальність за тестові матеріали та тестові дані;
- знання соціального, культурного та політичного контексту і умов, в яких застосовуватиметься тест, і шляхів потенційного впливу цих факторів на результати, їхню інтерпретацію та способи їх застосування.

Відповідні вміння з управління нестандартними та надзвичайними ситуаціями, зокрема:

- як діяти в ситуаціях, де є ризик виникнення **неналежного використання тестів** або **неналежного розуміння, як інтерпретувати тестові бали**.

На завершення огляду міжнародних документів стисло зупинимося ще на трьох. Це Міжнародні Рекомендації ІТС з адаптації тестів²² [6], Міжнародні Рекомендації ІТС з комп'ютерного та Інтернет-тестування²³ [7] та EFPA Модель аналізу та оцінювання психологічних тестів²⁴ [8].

Міжнародні Рекомендації ІТС з адаптації тестів

Метою розробки цього документа є створення комплексу рекомендацій для адаптації психологічних та педагогічних тестів, що використовуються в різних лінгвістичних та культурних контекстах.

Рекомендації поділяються на чотири основні категорії, що стосуються: культурного контексту; технічних аспектів розробки і адаптації інструменту; адміністрування тесту; документації та інтерпретації.

²² Міжнародні Рекомендації ІТС з адаптації тестів — *ITC Guidelines on Adapting Tests*.

²³ Міжнародні Рекомендації ІТС з комп'ютерного та Інтернет-тестування — *ITC International Guidelines on Computer-Based and Internet-Delivered Testing*.

²⁴ EFPA Модель аналізу та оцінювання психологічних тестів — *EFPA Review model for the description and evaluation of psychological tests*.

Міжнародні Рекомендації ІТС з комп'ютерного та Інтернет-тестування

Основними цілями проекту є:

- розробка комплекта **визнаних міжнародною спільнотою рекомендацій**, що сприятимуть належній практиці у комп'ютерному тестуванні²⁵ (CBT) та тестуванні через Інтернет;
- підвищення рівня зайнятих у процесі тестування сторін.

EFPA Модель аналізу та оцінювання психологічних тестів

Розроблено

Європейською федерацією асоціації психологів — EFPA

Документи, що застосовані у розробці цієї «Моделі»:

- Критерії аналізу тестів Британської психологічної спілки (BPS)
- Критерії Комітету із справ тестування (COTAN) Асоціації психологів Нідерландів (NIP).

Модель EFPA являє собою опитувальник для аналізу та оцінювання якості психологічних тестів. Певні підходи можуть бути використані для педагогічного оцінювання.

²⁵ Комп'ютерне тестування — *computer-based testing (CBT)*.

Вимоги до професійної відповідальності при оцінюванні

Кодекс професійної відповідальності при педагогічних вимірюваннях

Розроблено

Національною радою з вимірювань в освіті — NCME

Документи, застосовані у розробці «Кодексу»:

- Стандарти освітнього та психологічного тестування (AERA, APA, NCME). 1985.
- Положення щодо принципів належної практики. Національна асоціація радників з прийому до коледжів (NACAC). 1988.
- Кодекс справедливого тестування в освіті. JСТР. 1988.
- Принципи справедливого оцінювання в освіті Канади. Спільний дорадчий комітет (JAC). 1993.
- Етичні стандарти. Американська асоціація з консультування (ACA). 1988.
- Етичні стандарти AERA. 1992.
- Етичні принципи психологів та кодекс поведінки. APA. 1992.
- Відповідальність користувачів стандартизованих тестів: переглянуте положення. Американська асоціація з консультування (ACA) та Асоціація оцінювання у консультуванні (AAC). 1989.
- Стандарти компетентності вчителів і викладачів з педагогічного оцінювання учнів і студентів. Американська федерація вчителів (AFT), NCME, Національна педагогічна асоціація (NEA). 1990.
- Стандарти оцінювання персоналу. Спільний комітет по стандартах освітнього оцінювання (JCSEE). 1988.
- Стандарти оцінювання програм. (JCSEE). 1994.

Метою Кодексу²⁶ [9] є опис професійних обов'язків і відповідальності осіб, які задіяні в оцінюванні та тестуванні.

²⁶ Кодекс професійної відповідальності при педагогічних вимірюваннях (NCME, США, 1995) — *Code of Professional Responsibilities in Educational Measurement*.

Оскільки існує тісний зв'язок між професійно відповідальною практикою та обґрунтованим результатом, цей Кодекс узгоджений з відповідними розділами таких базових документів:

1. *Стандарти з освітнього та психологічного тестування [1].*
2. *Кодекс справедливого тестування в освіті [4].*

Кодекс стосується будь-якого типу оцінювання, що здійснюється як складова педагогічного процесу, у тому числі й формалізованих/неформальних, традиційних та альтернативних методик, а також — процедур збору інформації, що використовується для прийняття освітніх рішень на різних рівнях.

Кодексом професійної відповідальності керуються у своїй діяльності такі групи осіб:

- *освітяни (учителі, викладачі, керівники, адміністратори шкіл, коледжів та університетів);*
- *фахівці (шкільні психологи, тестологи, консультанти, радники);*
- *технічний, законодавчий та політичний персонал в галузі освіти на рівні регіонів або країни;*
- *персонал дослідних, оцінювальних та тестових організацій;*
- *тих, хто надає послуги з підготовки до тестування.*

Загальні види діяльності з оцінювання та відповідальність за Кодексом

Кодекс надає перелік основних професійних обов'язків і описує відповідальність усіх суб'єктів процесу оцінювання та тестування — тих, хто:

1. *Розробляє продукти оцінювання (програми, інструменти, методи).*
2. *Рекламує та продає інструменти оцінювання.*
3. *Вибирає методи та інструменти оцінювання.*
4. *Адмініструє оцінювання.*
5. *Підраховує бал у результаті оцінювання.*
6. *Інтерпретує, застосовує та повідомляє результати оцінювання.*
7. *Навчає з питань оцінювання.*
8. *Оцінює програми та здійснює дослідження поведінки у сфері оцінювання.*

Професійна відповідальність, викладена у Кодексі в межах восьми, названих вище, основних видів діяльності, передбачає, що особи, задіяні у педагогічному оцінюванні, будуть:

1. *Захищати здоров'я та безпеку екзаменованих.*
2. *Знати зміст державних і федеральних нормативних документів стосовно професійної поведінки та діяти згідно з ними.*
3. *Підтримувати і вдосконалювати професійну компетентність з оцінювання.*
4. *Надавати послуги з оцінювання лише в межах своєї компетентності, досвіду та професійної кваліфікації.*
5. *Сприяти розумінню обґрунтованих практик з оцінювання в освіті.*
6. *Дотримуватися найвищих стандартів поведінки та сприяти професійно відповідальній поведінці в межах освітніх інституцій та агенцій, що надають освітні послуги.*
7. *Виконувати всі професійні обов'язки чесно, відкрито, з належною справедливістю.*

Повний текст цього надзвичайно важливого регуляторного документа наведено у додатку.

ВИСНОВКИ

На завершення цього огляду наведемо структурну схему (див. табл. 3), де зроблена спроба згрупувати основні міжнародні регуляторні документи різних типів (стандарти, рекомендації, правила, кодекси), спрямовані на різні суб'єкти процесу оцінювання і тестування.

Представлені документи чітко виокремлюють вимоги до професійних, моральних та етичних якостей усіх суб'єктів процесу оцінювання і тестування, що фактично є умовою запровадження об'єктивних методів оцінювання та гарантією, що їхнє застосування буде справедливим. Можна підсумувати, що існує чотири групи регуляторних документів.

Технічні стандарти

До них належать Стандарти, спрямовані на визначення основних технічних та технологічних вимог до тестів. Це:

1. *Стандарти освітнього та психологічного тестування (США, 1999).*
2. *ETS Стандарти для забезпечення якості та справедливості (США, 2002).*
3. *Рекомендації з освітнього та психологічного тестування (CPA, 1996).*

Ця група Стандартів має на меті:

1. *Регулювати користування тестами залежно від умов їхнього застосування.*
2. *Формулювати технічні вимоги для досконалої професійної практики.*
3. *Надавати можливість установлювати валідність тесту та коректність висновків за результатами тестування.*

Кодекс належної практики

Мета Кодексу справедливого тестування і оцінювання в освіті (США, 1996) — представити сутність технічних Стандартів (APA, ETS, CPA), зробити їх зрозумілими для всіх учасників процесу.

Кодекс — це довідник для фахівців при виконанні їх обов'язків з використання тестів, які мають бути справедливими до всіх тестованих осіб. Він надає окремі рекомендації для розробників тестів та їхніх користувачів у чотирьох критичних сферах:

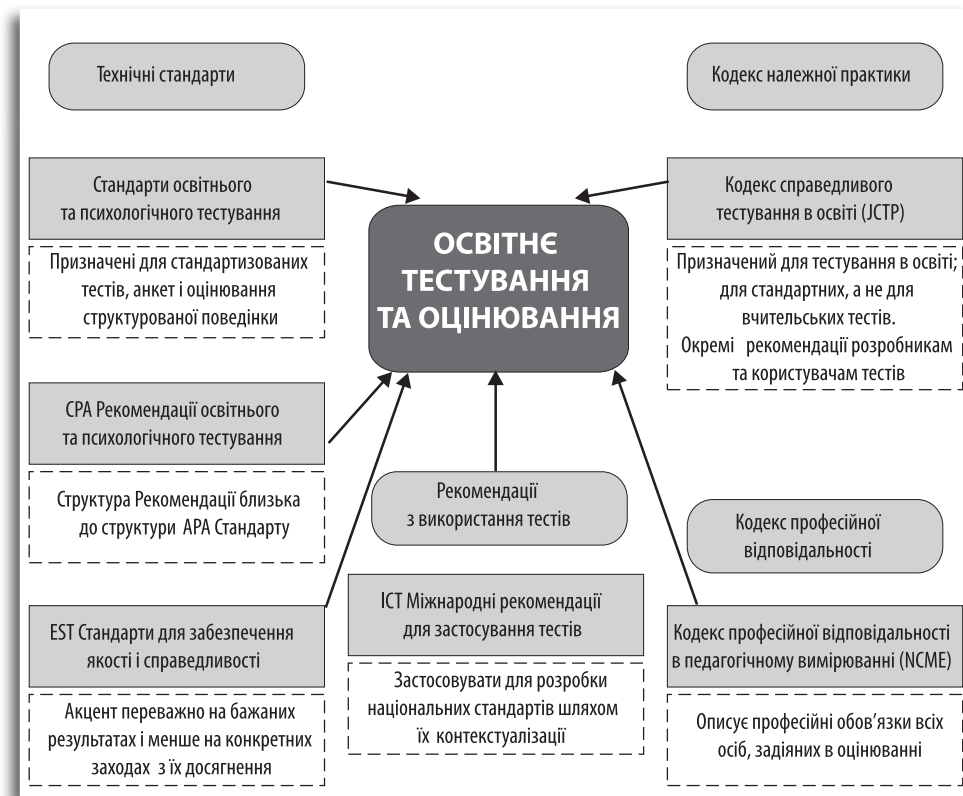
- A. *Розробка та вибір належних тестів.*
- B. *Адміністрування тестів і розрахунок балів.*
- B. *Звітність та інтерпретація результатів.*
- G. *Інформування осіб, які оцінюються.*

Рекомендації з використання тестів

Міжнародні Рекомендації ІТС з використання тестів (2000) розроблені фахівцями з різних країн: Бельгії, Данії, Англії, Франції, Німеччини, Нідерландів, Норвегії, Словенії, Південної Африки, Іспанії, Швейцарії, США.

Здійснені офіційні переклади такими мовами: голландською/фламандською, хорватською, французькою, німецькою, литовською,

Таблиця 3. Стандарти справедливого освітнього тестування та оцінювання



норвезькою, португальською, словенською, іспанською та португальською Південної Америки, іспанською, шведською, датською, китайською.

Метою Рекомендацій є узагальнення ключових принципів існуючих стандартів та створення Рекомендацій для їх адаптації на національному рівні різних країн.

Кодекс професійної відповідальності

Кодекс професійної відповідальності у сфері педагогічних вимірювань (NCME, США, 1995) містить регуляторні документи, опуб-

ліковані Buross Інститутом психічних вимірювань²⁷, які стосуються унормування багатьох аспектів у галузі педагогічних вимірювань і оцінювання та вимог до професійної компетентності педагогічних працівників.

Ураховуючи значний інтерес до питань, пов'язаних із компетентністю осіб, які проводять педагогічне оцінювання, у книзі представлено також такі три основні регуляторні документи:

1. *Кодекс професійної відповідальності при педагогічних вимірюваннях (Code of Professional Responsibilities in Educational Measurement [9]).*
2. *Стандарти компетентності вчителів з оцінювання учнів (Standards for Teacher Competence in Educational Assessment of Students [10]).*
3. *Стандарти компетентності освітніх адміністраторів з оцінювання учнів (Competency Standards in Student Assessment for Educational Administrators [11]).*

Кодекс професійної відповідальності також стосується фахівців з тестування, технічного, законодавчого та політичного персоналу в галузі освіти, персоналу дослідних, оцінювальних та тестових організацій.

У Додатках А і В представлено переклади цих документів українською мовою, які здійснено Центром тестування при МОЗ України.

Переклади носять інформаційний характер окрім **офіційного перекладу** Міжнародних Рекомендацій ІТС з використання тестів, що здійснено Центром тестування за дозволом Президента ІТС prof. Jacques Gregoire (Копії дозволів представлено у додатку С).

²⁷ Buross Інститут психічних вимірювань (*Buross Institute of Mental Measurement* — <http://www.unl.edu/buross>) засновано у 1938 році. Одним із напрямів діяльності закладу є задоволення суспільного інтересу та сприяння впровадженню педагогічних вимірювань шляхом, зокрема публікації тестів і стандартів з тестування.

3.

РОЗДІЛ

ЛІЦЕНЗУВАННЯ, СЕРТИФІКАЦІЯ, РЕЄСТРАЦІЯ

- Освітньо-** сертифікаційне тестування — засіб діагностування якості освіти.
- Система** ліцензування і сертифікації компетентності.
- Принципи** функціонування Центрів сертифікації.

НОВА ФІЛОСОФІЯ НА РИНКУ ОСВІТНІХ ПОСЛУГ

Інтеграція України у міжнародне співтовариство передбачає можливість співставлення рівнів та узгодження принципів підготовки фахівців. Індекс людського розвитку²⁸, який визначається за трьома показниками (валовий національний продукт, середня тривалість життя та рівень професійних знань фахівців), є однією з основних ознак, за якими оцінюється економічний потенціал країни. А враховуючи те, що і валовий національний продукт, і середня тривалість життя тією чи іншою мірою залежать від рівня професійних знань, на перший план виходить проблема запровадження сучасних технологій педагогічного оцінювання професійної компетентності фахівців.

Гострота цієї проблеми зумовлена, з одного боку, існуванням приватних закладів освіти та можливістю навчання за кордоном, з іншого — застарілістю й не розробленістю наукових основ дидактики контролю та відсталістю системи державного оцінювання якості освіти та рівня кваліфікації фахівців. Як правило, вітчизняна педагогічна наука і практика застосовує такі діагностичні технології, які унеможливають порівняння підготовки учнів, студентів, спеціалістів. Такі висновки базуються на основі детального аналізу, який дає можливість з'ясувати головні причини кризового стану вітчизняної педагогічної діагностики. Поза сумнівом, причина цього лежить у історичних традиціях.

Вибір засобів педагогічної діагностики базується на відповідних дидактичних принципах вимірювання результатів навчання. Як відомо, контроль результатів навчання забезпечує зворотний зв'язок у системі навчального процесу: *«Вимірювання — права рука навчання. Без вимірювання не буде оцінювання. Без оцінювання не зможе виникнути зворотний зв'язок. Без зворотного зв'язку не може з'явитися інформація про досягнуті результати. Без інформації про досягнуті результати не буде успішного навчання»* [12].

На сучасному етапі розвитку країни освіта як галузь економічної діяльності постала перед вибором: або чекати на кошти від держави, або формувати ринок освітянських послуг і таким чином вижити. Початок 90-х років минулого століття для України характеризувався виникненням недержавних навчальних закладів, для яких навчання

²⁸ Індекс людського розвитку (Human development index). Дані про нього наводяться у Звітах ООН щодо людського розвитку.

є комерційною діяльністю, та нової функції державних навчальних закладів — надання освітніх послуг на контрактній основі. Сьогодні, освоюючи певні спеціальності, на умовах контрактної оплати навчається більше ніж половина всіх студентів. Здійснюється спроба держави зменшити цю частку, але, аналізуючи досвід різних країн, можна з упевненістю сказати, що комерційна діяльність із надання освітніх послуг і надалі залишатиметься однією з найважливіших ланок діяльності навчальних закладів. Отже, освіта в Україні вже реально живе за новими законами, які поступово наближаються до ринкових. Ще одна характерна ознака — поступова відмова від державного розподілу, запровадження конкурсу при працевлаштуванні на престижні професії, підготовка фахівців відповідно до потреб ринку праці.

Таким чином, з одного боку, знайдені шляхи виходу з кризи середини 90-х років минулого століття допомогли зберегти викладацькі кадри, що врятувало вітчизняну освіту від руйнування. А з іншого — ще гостріше постала проблема щодо якості освіти взагалі та підготовки фахівців зокрема. Особливо це стосується напрямів підготовки, де вочевидь критичність помилки від некваліфікованих дій фахівця.

Як же ці проблеми вирішуються в розвинених країнах світу з історичними традиціями захисту людини? Найхарактернішою ілюстрацією цього є США.

Здоров'я, безпеку і добробут населення у країні захищає існуюча система ліцензування та сертифікації професійної діяльності фахівців. Цією системою встановлено перелік професій, займатися якими фахівець не може без отримання відповідного сертифіката. Динаміка поширення обов'язкової сертифікації професійної діяльності у США така: у 1952 р. сертифікація здійснювалася для 80 професій, у 1986-му — уже для 800, а у 1990 р. — понад 1000 професій потребувала проходження ліцензування. Понад 60 професій потребують обов'язкового ліцензування в усіх штатах країни. Це перш за все лікарі, медичні сестри, льотчики, інженери, архітектори, будівельники, фармацевти, податкові інспектори, юристи, державні службовці, викладачі тощо. Хоча б в одному із штатів країни для здійснення практичної діяльності потрібен сертифікат для продавців збанкрутілих підприємств, постачальників лимонів, продавців яєць, ловців устриць, садівників, фермерів, які вирощують коноплі, метеорологів тощо [13].

СИСТЕМА ЛІЦЕНЗУВАННЯ І СЕРТИФІКАЦІЇ КОМПЕТЕНТНОСТІ

Система ліцензування та сертифікації професійної діяльності забезпечує: по-перше, захист населення від некваліфікованих дій некомпетентного фахівця; по-друге, захист прав самого фахівця в умовах ринкових відносин як гарантія від дискримінації за ознаками статі, раси, походження, наклепу тощо. Оцінювання професійної компетентності залежно від ступеня обов'язковості та використовуваних процедур поділяється на ліцензування, сертифікацію та реєстрацію:

Таблиця 4.

	Стандартизоване оцінювання	Без оцінювання
Обов'язковість	<i>Ліцензування</i> (Licensure)	<i>Реєстрація</i> (Registration)
Добровільність	<i>Сертифікація</i> (Certification)	

Ліцензування та сертифікація професійної діяльності фахівців здійснюється спеціальними організаціями — центрами сертифікації та ліцензування (Licensure Boards), які є відокремленими від навчальних закладів і роботодавців та функціонують, як правило, в статусі неприбуткових організацій. Основною функцією цих організацій є діагностування рівня професійної компетентності з видачею сертифіката, наявність якого є визнанням того, що рівень професійних умінь та навичок фахівця є не нижчим за мінімально необхідний.

Така форма незалежної фахової експертизи спеціаліста, яка здійснюється третьою стороною, запускає в дію елемент ринкових відносин — конкуренцію навчальних закладів, різних за статусом і формою самоврядування, формами і методами навчання, а отже, сприяє запровадженню управління, що забезпечує якість підготовки. Незалежна професійна педагогічна експертиза результатів навчання змінює систему взаємовідносин **викладач — викладач, викладач — студент (учень), студент — студент**. У такій системі допомога

«знайомому» студенту буде спрямована не на фальсифікацію його екзаменаційної оцінки, а на додаткові заняття з ним, отже, сприятиме формуванню нормальних професійних відносин. Крім того, за такої системи візуалізується результат професійної діяльності викладача (це один із чинників протидії зовнішньому оцінюванню з боку викладацьких кіл) та збільшується мотиваційний елемент навчання студентів, оскільки з'являється можливість бути об'єктивно оціненим за здібностями і посісти відповідне рейтингове місце.

Принципи функціонування Центрів сертифікації

На ринку освітніх послуг залишаються лише ті Центри сертифікації, які забезпечують високу якість педагогічної діагностики.

Як приклад найбільш авторитетних можна навести ETS і NBME.

Служба освітнього тестування

ETS (<http://www.ets.org>) є найбільшою у світі приватною некомерційною організацією, що займається тестуванням і педагогічними дослідженнями. Вона пропонує широкий спектр продуктів і послуг, що легко пристосовуються до вимог замовників, має великий досвід у сфері педагогічних вимірювань. Ця унікальна комбінація забезпечує достатню гнучкість у задоволенні різноманітних потреб у навчанні й оцінюванні.

Завданням ETS є підвищення якості й забезпечення рівного доступу до освіти для людей у усьому світі. Вона допомагає вчителям навчати, студентам — учитись, а батькам — вимірювати освітні й інтелектуальні успіхи своїх дітей. Це робиться шляхом:

- *урахування думок педагогів, батьків, сторонніх критиків;*
- *вивчення потреб студентів і навчальних закладів;*
- *лідерства в розробці нових інноваційних продуктів і послуг.*

Місія: підвищення якості й забезпечення рівного доступу до освіти шляхом справедливого і правильного оцінювання, проведення досліджень і надання супутніх послуг. Продукти й послуги ETS забезпечують вимір знань і умінь, підвищують успішність і сприяють освітньому і професійному росту.

Мета: зайняти провідні у світі позиції в забезпеченні справедливого й об'єктивного оцінювання, у проведенні досліджень і наданні супут-

ніх послуг, що допомагали б індивідуумам, батькам, учителям, освітнім установам, представникам бізнесу, уряду, державі, шкільним округам, а також фахівцям у сфері вимірювань і дослідникам.

Цінності: соціальна відповідальність, рівність, доступність і якість. ETS підтримує ці цінності шляхом урахування думок педагогів, батьків і критиків.

3 історії виникнення

ETS була заснована в 1947 р., коли Американська рада з освіти (American Council on Education, ACE — <http://www.educause.edu>), Фонд Карнегі зі сприяння прогресу у викладанні (Carnegie Foundation for the Advancement of Teaching — <http://www.carnegiefoundation.org>) і Рада зі вступних іспитів до вузів (College Entrance Examination Board) надали свої програми тестування, частину коштів і провідних спеціалістів для створення Служби освітнього тестування під керівництвом Henry Chauncey, який був переконаний, що життя можна поліпшити шляхом рівного доступу до якісної освіти. Колишній помічник декана Гарвардського університету, бачив упровадження стандартизованого тестування однією з необхідних умов для створення суспільства, що базується на досягненнях, а не класовому становищі і привілеях. Зі створенням Служби засновники ETS упровадили в життя концепцію розвитку організації у сфері освіти, яка могла б зробити істотний внесок у прогрес освіти в Сполучених Штатах Америки.

Протягом декількох років після 1947 р. ETS перейшла кордони країни, в якій вона народилася, щоб зробити якісну освіту доступною людям різних країн світу. Сьогодні ETS характеризується прогресивним напрямом діяльності в галузі освіти, забезпечуючи найкращі освітні і професійні можливості в усіх країнах. Вона стала світовим лідером у сфері освітніх досліджень і вимірювань.

60 років тому ця незалежна некомерційна організація розширилася настільки, що охопила своєю діяльністю освітян усього світу. Сьогодні ETS відома своїми дослідженнями у сфері освіти і наданням міжнародних послуг. Досягнення ETS ґрунтуються на концепціях і принципах, закладених її засновниками. Цих понять і принципів вона дотримується і сьогодні.

Структура

Головний офіс ETS розташований у Принстоні, Нью-Джерсі (США), проте її представництва є в усьому світі. Сьогодні вона розробляє, адмініструє тестові іспити та обробляє результати більш ніж

24 мільйонів тестованих щорічно у понад 180 країнах та у більш ніж 9000 підрозділах в усьому світі.

Понад 2500 професіоналів працюють в офісах ETS у Сполучених Штатах і по всьому світі. З них більш ніж 1100 мають підготовку і досвід в освіті, психології, статистиці, психометрії, інформатиці, соціології і гуманітарних науках; 600 мають магістерські ступені, а 250 — докторські. Співробітники ETS мають високу кваліфікацію у сфері педагогічних досліджень і педагогічного оцінювання. Своїми розробками вони допомагають освітянам усього світу шукати нові шляхи підвищення якості освіти. В ETS працюють розробники тестів, статистики і психометристи, провідні спеціалісти з розробки методів оцінювання, експерти з великомасштабного оцінювання, дослідники, фахівці в галузі освітньої політики.

За підготовку і здійснення стратегічного керівництва міжнародним розвитком відповідальним є ETS Global Division (раніше — Міжнародний відділ ETS, ETS International Division). Цей відділ пропонує рішення, спрямовані на задоволення різноманітних потреб клієнтів. Він тісно співпрацює із замовниками на всіх етапах розробки того або іншого продукту і сприяє поширенню цих продуктів у всьому світі.

ETS має дві філії:

- *Capstar — провідна організація у сфері надання послуг із підготовки, онлайнового дистанційного навчання й оцінювання, розробки і проведення сертифікаційних і ліцензійних іспитів для асоціацій, корпорацій, уряду й академічних установ;*
- *ETS Global BV — філія, міжнародний підрозділ ETS, що впроваджує розробки ETS в освітні і ділові співтовариства в усьому світі (штаб-квартира знаходиться в Амстердамі).*

ETS постійно розширюється, відкриваючи нові регіональні офіси в Європі, Азії, на Близькому Сході, в Північній Африці, Південній Америці, Канаді, Китаї, Кореї, Індії і Латинській Америці. Кожен офіс є підрозділом філії ETS Global BV під керівництвом співробітників ETS. Регіональні офіси відповідальні за забезпечення зв'язків із споживачами/партнерськими організаціями в регіоні і управління подальшим розвитком організації.

У 2004 р. було створено новий регіональний офіс — ETS EMEA. Він відповідає за управління діяльністю ETS у Європі, на Близькому Сході й в Африці та покликаний полегшити доступ до продуктів і послуг ETS у цих регіонах. Детальніша інформація міститься на сайті www.etseurope.org.

Послуги

Регіональні офіси пропонують широкий спектр продуктів, послуг і освітніх рішень, включаючи:

- *програми і послуги з вивчення англійської мови;*
- *підготовку і технічну допомогу;*
- *розробку і проведення великомасштабного оцінювання;*
- *розробку і проведення тестування;*
- *консультативні послуги і нестандартні рішення.*

Центральний офіс ETS у Європі знаходиться в Утрехті (Utrecht, the Netherlands), крім того є офіс у Лондоні (London, the UK).

Клієнтами ETS є керівники органів освіти і члени шкільних рад, студенти і установи освіти, центри професійної підготовки.

Продукти

ETS надає широкий спектр продуктів і послуг, що підтримують процеси викладання, навчання й оцінювання. Серед них можна відзначити такі: методики вивчення англійської мови й інструменти оцінювання у вищій освіті; сучасні технологічні рішення, придатні як для окремого класу, так і для комп'ютеризованого середовища. Це, наприклад, програмний засіб для оцінювання в реальному масштабі часу ступеня засвоєння матеріалу кожним студентом — Discourse; мережовий продукт для швидкого оцінювання здібностей із написання есе — Criterion.

До продуктів ETS належать такі поширені тестові екзамени, як SAT, ACT, TOEFL та інші.

ETS надає також послуги з підготовки і консультування, застосовуючи великий досвід в педагогічних вимірюваннях. Протягом багатьох років ETS розробляла продукти і надавала послуги з педагогічного оцінювання, допомагаючи вимірювати продуктивність і успіхи в навчанні окремих індивідуумів. Тепер, використовуючи свій досвід і лідируючі позиції, ETS Global Division поширює подібні послуги на освітні установи, бізнесові структури і уряди в усьому світі. Шляхом формулювання умінь і впровадження тестування в широку практику ETS надає набір програм, що стимулює прогрес у навчанні й оцінюванні.

Партнерство. ETS залучає велику кількість фахівців і дослідників у галузі освіти в усьому світі. Вона використовує їхній досвід для забез-

печення користувачів новими продуктами і послугами. При створенні цих продуктів вона також співпрацює з академічними установами, мовними школами й іншими тестуючими організаціями.

Зв'язки. ETS випускає журнал Global Study Magazine, що розповсюджується в більш ніж 100 країнах і має близько 100 000 читачів у всіх країнах.

NBME

Національна Рада медичних екзаменаторів США (National Board of Medical Examiners / NBME — www.nbme.org) — це незалежна неприбуткова організація, яка покликана забезпечувати проведення високоякісної експертизи професіоналів у сфері охорони здоров'я. Основними завданнями, що стоять перед NBME, є захист населення від некваліфікованих дій працівників охорони здоров'я та проведення досліджень у сфері педагогічних вимірів і розробки інструментів оцінювання.

Рада була утворена у 1915 р. у відповідь на потребу в проведенні добровільних загальнонаціональних іспитів, результати яких могли б використовуватися при видачі ліцензій на проведення самостійної медичної практики відповідно уповноваженими на це органами.

З моменту свого заснування Рада займається розробкою і проведенням медичних ліцензійних іспитів, що стали всесвітньо визнаним етапом тестування й оцінювання в медичній освіті. На сьогодні вона також забезпечує медичні школи тестами з окремих базових і клінічних дисциплін.

Штат NBME складається з 250 осіб і включає лікарів, фахівців у сфері оцінювання знань і технічний персонал. Приблизно 80 членів Національної Ради входять до правління, що складається з кваліфікованих фахівців у сфері медицини, медичної освіти й оцінювання, медиків-практиків, представників Комітету з тестування Національної Ради, представників професійних організацій і громадськості. Чверть членів Національної Ради представляють інші організації, включаючи окремих представників від ВВС США, армії, флоту, служби організації охорони здоров'я, ветеранських організацій, Федерації медичних рад Штатів, Асоціації американських медичних коледжів, Американської Ради медичних фахівців, Американської медичної асоціації, Ради суспільств медичних фахівців, Американської медичної студентської асоціації, Національної студентської медичної асоціації і Секції лікарів-резидентів Американської медичної асоціації.

USMLE

У 1992 р. NBME і Федерація медичних рад Штатів (Federation of State Medical Boards, FSMB — <http://www.fsmb.org>) взялися до впровадження спільної програми — Національних медичних ліцензійних іспитів США (United States Medical Licensing Examinations, USMLE), що стала спадкоємицею сертифікаційних іспитів NBME і федеральних ліцензійних іспитів (FLEX).

USMLE — це іспит, що складається з трьох етапів (Степ 1, Степ 2 і Степ 3), які забезпечують загальноприйнятну процедуру оцінки для здобувачів ліцензії на медичну практику. Для одержання ліцензії медичної ради штату студенти-медики повинні здати всі три іспити USMLE. Ці іспити оцінюють здатність лікаря застосовувати знання основних концепцій і принципів, а також демонструвати спрямовані на пацієнта базові уміння, що є важливими з точки зору профілактики і лікування і які складають основу безпечної та ефективної медичної допомоги.

Результати USMLE повідомляються медичним відомствам, які відповідають за видачу початкової ліцензії на здійснення медичної практики.

Починаючи з 2004 р. екзамени Степ 1, Степ 2 — **Клінічні знання** (Step 2 CK) можна здавати цілорічно на комп'ютері. Іспит Степ 2 містить також окремий екзамен з перевірки **клінічних умінь** (Step 2 CS). Степ 3 також складається з двох частин: перша — багатовибіркові тестові завдання, а друга — комп'ютеризоване моделювання різних клінічних випадків. Складання Степ 1 і Степ 2 — Клінічні знання здійснюється в центрах тестування компанії Prometric по всьому світі, а складання Степ 3 здійснюється в центрах тестування компанії Prometric у Сполучених Штатах.

Здача 1-го або 2-го етапів є обов'язковою для одержання документа про закінчення медичної школи.

Степ 1 і Степ 2 являють собою одноденний тестовий іспит, тоді як Степ 3 є дводенним тестовим іспитом, що складається як із багатовибіркових питань, так і з розбору клінічних випадків.

NBME адмініструє також предметні іспити (shelf exams) в медичних школах. Це іспити як із базових предметів, так і з клінічних дисциплін. Багато медичних шкіл проводять ці іспити наприкінці навчання замість випускних. Тепер ці іспити проводяться в бланковій формі (папір-олівець).

Загальна кількість тестових запитань, що пропонуються кожному екзаменованому, досягає 600 — 750. Час і процедура іспиту строго регламентуються; питання в буклетах розміщені в довільній послідов-

ності. Іспит оцінює медичні знання й навички, що є важливими для надання безпечної й ефективної допомоги пацієнтам.

Степ 1 (<http://www.usmle.org/step1/default.htm>) звичайно складається після завершення курсу базових наук, як правило, наприкінці другого року навчання в медичній школі, оцінюючи знання і розуміння студентами ключових концепцій базових біомедичних наук і здатність застосовувати ці знання на практиці. Особливий акцент робиться на принципах і механізмах, що лежать в основі здорового способу життя, захворювання і способах терапії. Степ 1 гарантує досконале оволодіння не тільки дисциплінами, що є основою для безпечної та компетентної медичної практики, але й науковими принципами, необхідними для підтримання кваліфікації шляхом постійного відновлення знань.

Степ 2 (<http://www.usmle.org/step2/default.htm>) складається звичайно в останній рік навчання у медичній школі. Оцінює здатність студентів застосовувати медичні знання, навички і розуміння клінічних наук у тому обсязі, який необхідний для надання медичної допомоги під керівництвом старших колег. Акцент робиться на зміцненні здоров'я і профілактиці. Степ 2 гарантує освоєння клінічних наук і формування основних умінь, спрямованих на безпечну і компетентну медичну допомогу пацієнтам.

Степ 3 (<http://www.usmle.org/step3/default.htm>) оцінює здатність застосовувати медичні знання і розуміння біомедичних і клінічних наук у ступені, необхідному для самостійної медичної практики. Акцент робиться на вмінні вести пацієнта за амбулаторних умов. Степ 3 забезпечує кінцеве оцінювання лікарів, які несуть особисту відповідальність за надання загальної медичної допомоги і лікування. Ті, хто складають екзамен, зазвичай розпочинають підготовку до іспиту 3-го рівня по закінченні перших шести місяців резидентської програми.

Створення й адміністрування тесту, обробка результатів

Кількість балів і результат «склав/ не склав» звичайно повідомляється студентам та їхнім медичним школам, додатково ліцензійним органам штату з метою ухвалення рішення про видачу первинної ліцензії на медичну практику. Незважаючи на те що ліцензійні органи кожного штату можуть установлювати власні правила для ліцензування лікарів, усі вони на сьогодні розглядають прохідний бал USMLE як основну планку в ліцензуванні.

Тестові завдання до екзаменів кожного з етапів створюють досвідчені викладачі і клініцисти, які є визнаними авторитетами у своїх галу-

зях. Ці експерти відбираються з академічних кіл, лікарів-практиків і співробітників ліцензійних органів з усієї території США і Канади. Наприклад, для створення питань для екзамєну Степ 2 існують п'ять предметних комітетів (внутрішня медицина, акушерство і гінекологія, педіатрія, психіатрія і хірургія). Комітети складаються з восьми членів, які розробляють по 50 запитань у рік. Кожен комітет щороку збирається разом на три дні, щоб розглянути всі запитання, написані його членами. Кожне запитання зачитується вголос його автором, потім обговорюється і в разі необхідності переробляється комітетом. Запитання, що мають відповідний зміст, технічно правильні і добре написані, схвалюються для потенційного використання на іспиті. Потім голови всіх предметних комітетів працюють спільно протягом трьох днів. Для цієї зустрічі готується чорновий варіант іспиту відповідно до точного опису його змісту. Голови читають уголос запитання, написані їхніми комітетами, а голови інших комітетів повинні схвалити кожне запитання окремо як таке, що відповідає змісту певного іспиту. Голови зобов'язані розробити інші запитання замість відхилених колегами. Двоступінчастий процес розгляду запитань забезпечує точність при розгляді кожного запитання експертами, а міждисциплінарний комітет приймає рішення щодо його прийнятності. Участь усіх цих людей у процесі побудови іспиту допомагає забезпечити представлення в ньому поточних наукових концепцій і широко демонструє знання, що вважаються необхідними для початку медичної практики під керівництвом досвідченого колеги. Кожен початківець, розробляючи тестові запитання, бере участь у дводенному семінарі, що є частиною підготовки до написання високоякісних тестових матеріалів.

Національна Рада підтримує сувору засекреченість іспитів, що запобігає доступу до тестових матеріалів перед проведенням іспиту. Заходів безпеки дотримуються і при поверненні тестових буклетів у NBME відразу по завершенні іспиту. NBME має у своєму розпорядженні спеціальні засоби аналізу результатів іспиту на предмет «списування» відповідей або передчасного доступу до екзамєнаційних запитань. У випадку виявлення в екзамєнованих осіб матеріалів, що порушують екзамєнаційний процес, NBME починає відповідні законні дії.

Інші програми й послуги

USMLE — найбільший екзамєнаційний проект NBME. Проте Національною Радою надаються й інші послуги.

Для медичних шкіл — NBME надає допомогу й інформаційну підтримку педагогам-медикам, розробляючи високоякісні інструменти

оцінювання, відомі як предметні іспити, придатні для різних підрозділів медичних шкіл, що дозволяють вимірювати знання з кожної традиційної базової і клінічної дисципліни та порівнювати успіхи кожного студента-медика із загальнонаціональними показниками.

Через програми зв'язку з медичними школами, науковими товариствами й організаціями студентів-медиків NBME допомагає в підготовці до USMLE. Персонал NBME проводить семінари з написання тестових завдань.

Для **студентів-медиків** NBME надає можливість on-line реєстрації, використання програм самооцінки й інформацію про предметні іспити.

Для **професійних медичних організацій** NBME забезпечує тестування, надає послуги в галузі освіти, консультує і проводить наукові дослідження для безлічі медичних рад, товариств і наукових організацій, пов'язаних з охороною здоров'я. Послуги включають розробку, адміністрування й аналіз більш ніж 30 іспитів з метою сертифікації, повторної сертифікації, перепідготовки, самооцінки або оцінки спеціальної компетенції. NBME пропонує як консультативні, так і комплексні послуги, які гарантують, що програми оцінювання відповідають експериментальній моделі та їх можна оновити з використанням найновіших статистичних і технологічних розробок.

Самооцінювання: у 2003 р. NBME впровадив on-line програму для самооцінки змісту базових і клінічних дисциплін для студентів-медиків і випускників.

Розробки

NBME постійно підтримує проведення досліджень у сфері оцінювання клінічних навичок, розробки передових методів тестування і продовжує дослідження вірогідності й надійності своїх екзаменаційних програм. У травні 2001 р. був створений Інноваційний центр, призначений для розробки продуктів і послуг, спрямованих на розвиток нових напрямків, що відповідають стратегічній місії NBME.

ВИСНОВКИ

Якість сертифікації базується на трьох чинниках: об'єктивності оцінювання, сучасних технологіях тестування, перевірці практикою.

Об'єктивність оцінювання. Результат сертифікаційних тестувань має бути максимально захищеним від впливу суб'єктивних факторів і максимально відповідати реальним показникам особи, яка тестувалася. Це забезпечується дотриманням таких умов: об'єктивність проведення тестування, об'єктивність обробки даних, об'єктивність інтерпретації результатів.

Окремо слід відзначити, що сертифікаційні тестування, які належать до «high stakes» — тестувань високого рівня, проходять за закритими тестами, які щорічно відповідно до технології стандартизованих тестувань, оновлюються. У більшості країн несанкціоноване розповсюдження екзаменаційних матеріалів є приводом для порушення судової справи.

Сучасні технології тестування. Конкуренція освітньо-сертифікаційних центрів є одним із чинників, що обумовили високий науковий рівень розвитку методів педагогічної діагностики у США. З іншого боку, **перевірка практикою** стимулює центри використовувати найсучасніші технології оцінювання та тестування. У свою чергу, роботодавці зацікавлені в існуванні ринку спеціалістів, кваліфікація яких підтверджена сертифікаційними тестуваннями. Так, дані опитування серед фірм, які працюють у сфері інформаційних технологій, свідчать, що 78% керівників фірм вважають працю сертифікованих спеціалістів ефективнішою; 92% компаній набули фінансових переваг унаслідок сертифікації своїх співробітників; інвестиції компаній у навчання та сертифікацію персоналу окуповуються приблизно через 8 — 10 місяців.

Отже, проблемою, що постає перед українським суспільством сьогодні, є дефіцит об'єктивних критеріїв, які забезпечували б порівняння фахівців, закладів, послуг, програм тощо і створювали б основу для конкуренції, що в результаті призводило б до поліпшення економічної ситуації в країні. Остання певною мірою є наслідком того, що формування фахівців (початкова, середня, вища освіта) здійснюється за відсутності об'єктивного оцінювання їхніх особистих здібностей та професійної компетентності. Причина цього — відсутність викладацьких кадрів, здатних застосувати методи об'єктивного контролю під час виховання та навчання учнів, студентів, фахівців. Слід підкреслити, що суб'єктивізм при оцінюванні сприяє скоріше протекціонізму, ніж конкуренції. Тому формування нового покоління через педагогічний ланцюг: **«новий викладач — новий випускник — новий фахівець»** сприятиме підвищенню індексу людських ресурсів України, а отже, і зростанню економічного потенціалу нашої країни.

4.

РОЗДІЛ

ВИМОГИ ДО ПРОФЕСІЙНОЇ КОМПЕТЕНТНОСТІ ВИКЛАДАЧІВ

Професійні	вимоги до вчителя.
Ліцензійні	педагогічні тести, що використовуються у США.
Особливості	ліцензування викладацької діяльності у США.

Центральне місце у забезпеченні якості освіти (яку можна визначити як відповідність меті) посідає якість викладання і, отже, якість підготовки самих учителів. Але чи відповідає сучасним освітнім цілям та роль, яку вчитель відіграє сьогодні? Чи відповідає сучасним освітнім цілям методика і зміст підготовки вчителів? На жаль, можна констатувати, що ні, не відповідає і потребує істотних змін. У широкому аспекті такі зміни мають містити в собі:

- перегляд цілей і стандартів у сфері освіти і, як результат, перегляд навчальних програм (у тому числі методик викладання);
- реформу системи базової підготовки вчителів і системи підвищення кваліфікації, щоб і вчителі-початківці, і досвідчені педагоги мали можливість підвищувати свій професійний рівень;
- розробка і впровадження стандартів якості.

Підвищення якості педагогічної освіти, що безпосередньо впливає на розвиток суспільства і цивілізації, тісно пов'язане сьогодні із стандартизацією. Детально питання щодо якості вищої освіти і державних стандартів вищої освіти в Україні розглянуто в роботі [14].

Основні ідеї нових підходів до реформування системи педагогічної освіти полягають у забезпеченні неперервності і різноманітності освіти, гарантуванні її якості, орієнтації на широкий діапазон відмінностей у фаховій підготовці викладачів, чіткому визначенні складових професійної компетентності вчителів.

ПРОФЕСІЙНІ ВИМОГИ ДО ВЧИТЕЛЯ

Аналіз процесів, що відбуваються в системі освіти, указує на наявність великої кількості факторів (включаючи і фактори, причини яких лежать за межами навчального закладу), що впливають на якість

освіти. Найголовнішими з них є два. Першим фактором, що визначає якість навчання, є якість підготовки викладача, точніше — **якість навчання, що він забезпечує**. Другим важливим фактором є якість навчального закладу.

Зосередимо свою увагу на першому з цих факторів. Якість викладача — це не тільки питання кваліфікації або особистісних характеристик учителя. Це і похідне від того, що вчитель фактично робить, яким чином поводить у аудиторії. **Ніщо не може зрівнятися з гарним учителем, який добре навчає** [15]. Саме тому визначення стандартів якості викладача є основним для розуміння ролі ліцензійного тестування.

Однак, що ми маємо на увазі, коли говоримо про **гарне викладання**? Які критерії якості існують у цій сфері? На наш погляд, ці критерії визначаються тим, якою мірою навчальний процес відповідає цілям, що ставляться суспільством перед освітою в цілому. Такий підхід змушує змінити традиційне уявлення про гарного вчителя. Для вчителя уже недостатньо виконувати тільки завдання передачі своїм учням набору знань, що є важливими для їхньої подальшої діяльності. Ця мета залишається надзвичайно важливою, однак тепер гарне викладання має також спрямовуватися на розвиток і виховання таких навичок та здібностей, що мають вирішальне значення для сучасної і майбутньої діяльності молодих людей, сприятимуть розвитку їхніх особистісних якостей, дадуть змогу виконувати свою роль у суспільних інституціях.

У кожній країні існують різні критерії щодо якості викладання. На наш погляд, ці критерії визначаються тим, якою мірою навчальний процес відповідає цілям, що ставляться з боку суспільства перед освітою в цілому. Розглянемо підходи до оцінювання вчителів та уявлення у різних країнах про **кваліфікованого вчителя**.

Певний інтерес представляє собою система підготовки вчителів у **Литві**, що входила до недавнього часу до складу СРСР. У цій країні вважають, що процеси відновлення суспільства можуть бути успішно реалізовані лише в тому випадку, якщо ми будемо готувати [16]:

- *учителя-фахівця, який знає свій предмет і вміє працювати у цій сфері (вважається, що знання і навички можна отримати в процесі навчання і практичної діяльності);*
- *учителя-диакта-методиста, який не тільки знає свій предмет, але й здатний передати іншим свої знання, навички, досвід (отже, окрім знань, потрібна ще психологічна і методична підготовка, що набуває великого значення при підготовці до роботи в школі);*

- учителя-вихователя-педагога, який толерантно ставиться до інших, бажає і вміє допомагати їм, може впливати на них не тільки словом, але й ділом, своїм особистим прикладом.

Дещо інший підхід спостерігається у **Франції**. Там акцент було зроблено на **перегляді традиційних уявлень про професіоналізм і кваліфікаційні вимоги**. Так само, як і в інших країнах, у Франції були розроблені нові вимоги до вчителів, що отримали назву **«відкритого професіоналізму»**. Зміни в процесі викладання і навчання в школах супроводжувалися відходом від індивідуалізму у викладанні і науковому аналізі практичної діяльності вчителя [17]. За новими професійними вимогами викладач повинен:

- проводити навчання у тісній співпраці з членами шкільного колективу (підхід «групового викладання»);
- здійснювати індивідуальне і групове навчання учнів, приділяючи основну увагу загальному розвитку дитини;
- уміти визначати потреби учнів і проблеми у навчанні;
- уизначати освітні цілі шкільних навчальних програм, аналізувати їх і переглядати у міру необхідності;
- уміти розробляти і адаптовувати навчальні програми і методики викладання, проводити дослідницьку роботу;
- уміти проводити оцінювання досягнутих результатів;
- співпрацювати з батьками, представниками громадськості і особами, відповідальними за суспільну і культурну діяльність тощо;
- уміти вести постійний діалог з учнями для здійснення тіснішого педагогічного контакту, а також для можливої корекції плану викладання/навчання;
- брати участь у програмах підвищення кваліфікації як викладачі, так і тих, кого навчають;
- брати участь у плануванні і управлінні діяльністю школи, роботі на основі свого досвіду пропозицій щодо розвитку системи освіти в цілому.

Загальні уявлення про **кваліфікованого вчителя**, професійні вимоги до викладачів, запроваджені в різних країнах, було розглянуто у роботі [18]. Певною мірою вони схожі між собою, але найбільш розвинутою є система стандартів у сфері професійного викладання, яка застосовується у Сполучених Штатах Америки, де підготовці вчителів приділяється винятково важлива роль. Президент США Джордж

Буш через два тижні після інавгурації виступив з новою програмою реформування системи освіти, однією з головних цілей якої є вдосконалення якості підготовки вчителів.

У США діють три системи стандартів, що тісно пов'язані між собою і описують послідовні етапи викладацької кар'єри. INTASC²⁹ стандарти описують знання, навички, здібності та нахили викладачів-початківців. NCATE³⁰ стандарти націлені на вимоги до програм підготовки викладачів і описують знання і навички кандидатів у викладачі. NBPTS³¹ стандарти встановлюють вимоги до викладачів, які вже займаються професійною діяльністю.

Консорціум з підтримки та оцінювання вчителів-початківців

Консорціум з підтримки та оцінювання вчителів-початківців був створений у 1987 р. Це об'єднання державних освітніх агентств, вищих навчальних закладів і національних освітніх організацій, пов'язаних з реформою освіти, ліцензуванням і підвищенням (професійної) кваліфікації викладачів. Основні засновники INTASC — державні освітні агентства, що відповідають за ліцензування викладачів і підвищення їхньої (професійної) кваліфікації. Головний принцип його діяльності можна сформулювати так: **кваліфікований викладач повинен поєднувати зміст навчання з педагогічним розумінням, аби бути впевненим, що навчання ведеться на високому рівні.**

Зразок базових стандартів INTASC для ліцензування викладачів містить норми, яким має відповідати викладання незалежно від предмета або рівня і які повинні бути основою для реформи системи підготовки і підвищення (професійної) кваліфікації викладачів [19]. Базові стандарти складаються з 10 норм, кожна з яких містить вимоги до компетентності, здібностей і майстерності викладача. Базові стандарти на сьогодні поширюються і на стандарти викладання певної дисципліни.

²⁹ Консорціум з підтримки та оцінювання вчителів-початківців — *Interstate New Teacher Assessment and Support Consortium*.

³⁰ Національна Рада з акредитації програм підготовки викладачів — *National Council for Accreditation of Teacher Education*.

³¹ Національна Рада із стандартів у галузі професійного викладання — *National Board for Professional Teaching Standard*.

Базові стандарти INTASC

Норма 1: Викладач розуміє основні концепції, методи оцінювання, структуру дисциплін(и), що викладається, і здатний моделювати навчальні ситуації, що роблять певні аспекти предмета зрозумілими для студентів³².

Компетентність. Викладач розуміє основні концепції, припущення, спірні питання, методи опитування і шляхи пізнання, які є найважливішими в дисципліні³³, що викладається.

Викладач розуміє, як концептуальна система поглядів студентів та їхні хибні уявлення про предметну область можуть впливати на їхнє навчання. Він може пов'язувати зміст своєї дисципліни з іншими предметними областями.

Здібності. Викладач розуміє, що предметні знання — це не фіксований набір фактів, а складна система, що розвивається. Він прагне не відставати від нових ідей і досягнень у цій сфері.

Викладач бере до уваги і враховує логіку розвитку змісту дисципліни, бачить його прояв у повсякденному житті.

Викладач постійно поповнює свої знання і бере участь у професійних бесідах про зміст і особливості викладання дисципліни.

Продуктивність. Викладач ефективно використовує різні способи пояснення основних положень дисципліни і пов'язує ці положення з вивченими раніше. Розуміє і використовує різні теорії пізнання і методи дослідження.

Викладач здатний оцінювати навчальні ресурси і методичні матеріали з точки зору їх повноти, точності і придатності для представлення окремих ідей та концепцій.

Викладач залучає студентів до моделювання і перевірки гіпотез, використовуючи властиві дисципліні методи дослідження і стандарти доказів. Сприяє розвитку аналітичних здібностей і критичного мислення студентів.

Викладач здатний створювати міжпредметні зв'язки, що дають змогу студентам поєднувати знання, навички і методи дослідження, які використовуються в різних предметних областях.

³² Під терміном «студент» розуміємо студентів або учнів.

³³ Під терміном «дисципліна» розуміємо навчальну дисципліну або навчальний предмет.

Норма 2: Викладач розуміє закономірності розвитку і навчання учень/студент, здатний організовувати навчальне середовище, що сприяє їхньому інтелектуальному, соціальному й індивідуальному розвитку.

Компетентність. Викладач розуміє, яким чином відбувається процес навчання: як студенти засвоюють знання, набувають навичок, розвивають розумові здібності; знає, як правильно організувати навчання для досягнення найкращого результату.

Викладач розуміє, що розвиток фізичних, соціальних, емоційних, моральних, і пізнавальних здібностей студентів впливає на навчання, і вміє використовувати ці фактори в навчальному процесі.

Викладач може визначити рівень підготовленості і розуміє, як прогрес у кожній окремій області може впливати на розвиток і вивчення інших.

Здібності. Викладач розуміє індивідуальні розбіжності в здібностях, виявляє повагу до різноманітних талантів усіх учнів і активно допомагає їм розвивати впевненість у собі. Він використовує сильні здібності студентів як основу для професійного зростання, а їхні помилки — як стимул для навчання.

Продуктивність. Викладач аналізує індивідуальну і групову роботу для вибору методики, що відповідає поточним потребам учнів у кожній області (пізнавальній, соціальній, емоційній, моральній і фізичній) і полегшує їх подальше навчання.

Викладач стимулює закріплення попередніх знань і пов'язує нові ідеї з уже відомими, установлює зв'язок із попереднім досвідом студентів, залучає їх до формування навчальних завдань.

Шляхом доброзичливої бесіди з групою викладач впливає на розвиток мислення студентів, аналізує їхні усні та письмові відповіді.

Норма 3: Викладач розуміє, що студенти відрізняються у своїх підходах до навчання, і створює навчальне середовище, індивідуально пристосоване до кожного з них.

Компетентність. Викладач розуміє і може ідентифікувати розбіжності в підходах до навчання і практичної діяльності, включаючи особливості сприйняття, різну здатність до освоєння матеріа-

лу, і може створити методику, що використовує сильні сторони студентів як основу для їхнього зростання.

Викладач знає особливості слухачів, включаючи інвалідність, специфічні фізичні або розумові проблеми. Викладач розуміє складнощі, що виникають при вивченні іншої мови, і знає, як допомогти студентам.

Викладач розуміє як індивідуальний досвід і попереднє навчання, а також мова, культура і родинне виховання впливають на навчальні досягнення студентів. Викладач ураховує культурну і соціальну розмаїтість і знає, як пристосувати методику навчання до життєвого досвіду і культурних особливостей студентів.

Здібності. Викладач виходить з того, що всі учні/студенти можуть добре вчитися, і постійно допомагає їм у досягненні успіху. Він шанобливо ставиться до талантів і здібностей студентів, дотримуючись принципу індивідуального підходу.

Викладач поважає студентів як особистості, делікатно ставиться до суспільних і культурних норм. Він допомагає студентам відчувати себе особистостями, учить їх цінувати одне одного.

Продуктивність. Викладач розробляє методику навчання відповідно до рівня розвитку, особливостей сприйняття, сильних сторін і потреб студентів. Він використовує методи навчання, що враховують різний досвід тих, кого навчають. Викладач створює навчальне середовище, в якому поважаються індивідуальні особливості.

Викладач створює відповідні умови (час і умови роботи, завдання, засоби спілкування і методи контролю) для окремих студентів, що мають специфічні особливості або потреби.

Викладач здатний ураховувати культурні і сімейні особливості студентів і використовує цю інформацію при розробці методики викладання (наприклад, пропонуючи завдання, пов'язані з культурою і попереднім досвідом студентів).

Норма 4: Викладач розуміє і використовує різноманітні методи навчання, що сприяють розвитку у студентів професійних навичок, критичного мислення та уміння вирішувати проблеми.

Компетентність. Викладач розуміє, які процеси пізнання пов'язані з певними видами навчання (критичне і творче мислення, структурування і вирішення проблем, винахідливість, запам'я-

товування і пригадування) і як ці індивідуально-типологічні властивості можуть стимулюватися.

Викладач розуміє переваги й обмеження, пов'язані з різними методами навчання (колективне навчання, прямий інструктаж, групове обговорення, незалежне навчання, міждисциплінарний метод).

Він знає, як поліпшити навчання шляхом використання різноманітних матеріалів, а також людських і технологічних ресурсів (комп'ютери, аудіовізуальні технології, відеозаписи, експерти, первинні документи й експонати, тексти, довідники й інші друковані видання).

Здібності. Викладач оцінює розвиток критичного мислення студентів, їхню здатність незалежно вирішувати проблеми, їхні індивідуальні можливості.

Викладач оцінює гнучкість і лабільність, необхідні для пристосування методики до особливостей аудиторії.

Продуктивність. Викладач ретельно продумує шляхи досягнення мети навчання відповідно до студентських потреб.

Для активізації навчального процесу викладач використовує різні стратегії навчання і вчення, що сприяють розвитку критичного мислення, умінню вирішувати проблеми та формуванню професійних якостей; допомагають студентам усвідомлювати відповідальність за вибір і використання навчальних ресурсів.

Викладач змінює свою роль у навчальному процесі (інструктор, помічник, слухач) залежно від змісту, методичних цілей і потреб студентів.

Використовуючи альтернативні підходи, викладач розробляє достатню кількість наочності і зрозумілих образів, що сприяють кращому сприйняттю і розвитку критичного мислення.

Норма 5: Викладач використовує розуміння мотивації і особливостей поведінки індивідуума і групи для створення навчального середовища, що сприяє позитивному спілкуванню й активній участі в процесах навчання і самомотивації.

Компетентність. Викладач використовує знання про мотивацію людської поведінки, отримані з психології, антропології і соціології, для розробки стратегії організації і підтримки індивідуальної і групової роботи.

Викладач розуміє, як функціонують соціальні групи, як вони впливають на людей і як люди впливають на ці групи. Викладач знає, як організувати спільну продуктивну роботу у складних соціальних умовах.

Викладач володіє принципами ефективного управління класом і здатний використовувати різні методи для створення клімату позитивного співробітництва і змістовного навчання.

Викладач ідентифіковує фактори і ситуації, що можуть підсилити або зменшувати внутрішню мотивацію, і знає, як підвищити самотивацію студентів.

Здібності. Викладач здатний брати відповідальність за створення позитивного клімату в класі і брати участь у підтримці такого клімату в усій школі.

Викладач усвідомлює, як його участь сприяє вираженню і розвитку демократичних засад у класі. Викладач оцінює взаємодопомогу студентів у навчанні і визнає важливість колегіальних стосунків у навчанні.

Викладач визнає значення внутрішньої мотивації для постійного розвитку і навчання студентів. Викладач використовує принцип безперервного розвитку індивідуальних здібностей студентів і враховує, як мотивація може сприяти такому розвитку.

Продуктивність. Викладач створює безконфліктне навчальне середовище, у якому студенти вчаться усвідомлювати свою відповідальність, брати участь у прийнятті рішень, працювати спільно і незалежно.

Пов'язуючи теми занять з особистими інтересами студентів, даючи їм можливість вибирати досліджуваний матеріал, спонукаючи їх ставити запитання і досліджувати важливі для них проблеми, викладач залучає студентів до індивідуальної і колективної навчальної діяльності, що допомагає розвитку мотивації успіху.

Викладач організовує, розподіляє різні навчальні ресурси для забезпечення активного і рівноправного залучення студентів до вирішення навчальних завдань.

Викладач намагається збільшити кількість аудиторного часу, що витрачається на навчання, створюючи можливість для комунікації відповідно до навчальних завдань.

Викладач допомагає групі розвивати здібності до взаємодії, навчальних дискусій, індивідуальної і групової відповідальності,

що створюють в аудиторії позитивний клімат відкритості, взаємної поваги і підтримки.

Викладач аналізує обстановку в аудиторії і приймає корегуючі рішення для посилення колективної взаємодії, мотивації і ефективності навчання.

Норма 6: Викладач використовує ефективні вербальні, невербальні й інформаційні технології для підтримки атмосфери активної взаємодії, співробітництва і толерантності в аудиторії.

Компетентність. Викладач розуміє закономірності теорії комунікації, розвитку мовлення і роль мови в навчанні.

Викладач розуміє, як культурні і гендерні розбіжності можуть позначатися на процесі спілкування в аудиторії.

Викладач визнає важливість вербального і невербального спілкування. Він знає і може ефективно використовувати вербальні, невербальні, а також інформаційні технології.

Здібності. Викладач усвідомлює, яку важливу роль відіграє мова у самовираженні, розвитку індивідуальності, зростанні освіченості.

Викладач оцінює і заохочує до різних способів спілкування в аудиторії. Він — вдумливий і уважний слухач.

Викладач оцінює культурні складові спілкування і прагне підсилити їх у всіх студентів.

Продуктивність. Викладач моделює ефективні стратегії спілкування при передачі ідей та інформації, постановці запитань (моніторинг впливу повідомлень; повторне формулювання ідей і побудова зв'язків; використання візуальних, слухових і кінестетичних реплік, чутливих до поданих і отриманих невербальних повідомлень).

Викладач знає, як у різних ситуаціях та для досягнення певної мети формулювати запитання і стимулювати обговорення, перевіряючи, наприклад, глибину знань учня, допомагаючи студентам чітко формулювати свої думки, правильно приймати ризиковані рішення і вирішувати проблеми, стимулюючи конвергентне і дивергентне мислення, цікавість та допомагаючи студентам правильно формулювати запитання.

У своєму спілкуванні викладач ураховує культурні і гендерні розбіжності (правильне використання візуального контакту,

інтерпретацію мови, рухів тіла та усних тверджень, відгук на різні способи спілкування).

Викладач знає, як використовувати різноманітні інформаційні інструменти, включаючи аудіовізуальні засоби і комп'ютери для навчальних цілей.

Норма 7: Викладач розробляє методику навчання на підставі знань предмета, студентів, суспільства і навчальних цілей.

Компетентність. Викладач розуміє теорію навчання, предмет, правила побудови навчального плану, закономірності розвитку студента і знає, як використовувати ці знання при розробці методики, що забезпечить досягнення навчальних цілей.

Викладач знає, як при розробці методики, що ефективно пов'язує навчальні цілі і знання студентів, використовувати навчальні матеріали, індивідуальні інтереси, потреби та здібності студентів, суспільні ресурси.

Викладач знає, коли і як корегувати навчальні плани, виходячи з рівня знань студентів та інших зовнішніх факторів.

Здібності. Викладач оцінює різні сторони довгострокового і короткострокового планування. Він упевнений, що плани повинні бути завжди відкриті для корегування і перегляду з урахуванням потреб студентів і обставин, що змінюються.

Викладач розглядає планування як колегіальну діяльність.

Продуктивність. Як індивідуум і член команди викладач розробляє корисні для учнів навчальні завдання, що відповідають навчальним цілям і засновані на принципах ефективного управління (активізують попередні знання студентів, знижують упередженість, спрямовані на дослідження і вирішення проблем, розвивають нові навички на підставі вже надбаних).

Викладач розробляє методики проведення занять для різних рівнів, що відповідають індивідуальним потребам розвитку різних слухачів і допомагають прогресу кожного з них.

Викладач створює коротко- і довгострокові плани, що враховують студентські потреби, і пристосовує їх для того, щоб підтримати і використати студентські успіхи і мотивацію.

Викладач реагує на непередбачені зміни, оцінює плани з погляду кінцевих і конкретних навчальних цілей і систематично корегує навчальні плани, щоб вони відповідали потребам студентів і підвищували якість навчання.

Норма 8: Викладач розуміє і використовує формальні і неформальні методи оцінювання для оцінки і гарантії безперервного інтелектуального, соціального і фізичного розвитку учнів.

Компетентність. Викладач розуміє переваги й обмеження різних типів оцінювання (критеріально– і нормоорієнтоване, традиційні стандартизовані і кваліфікаційні тести, есе) для визначення того, як студенти вчаться, що вони знають і здатні робити, які види практичної діяльності будуть сприяти їх подальшому зростанню і розвитку.

Викладач знає, як відбирати, розробляти і використовувати інструменти оцінювання навчальних досягнень, а також з іншими діагностичними цілями.

Викладач розуміє теорію педагогічних вимірів і категорії, що пов'язані з оцінюванням (валідність, надійність, точність, шкалювання).

Здібності. Викладач вважає поточне оцінювання основою освітнього процесу і визнає, що правильне і систематичне використання різних методів оцінювання є необхідним для моніторингу його якості.

Викладач прагне використовувати оцінювання швидше для визначення сильних сторін студентів і для допомоги, а не для перекриття доступу студентів до подальшого навчання.

Продуктивність. Викладач використовує різні формальні і неформальні методи оцінювання (спостереження, портфоліо, класне тестування, практичні завдання, проекти, студентська самооцінка, експертне оцінювання, стандартизоване тестування) для розширення знань про учнів, оцінювання прогресу студентів, корегування методів навчання.

Викладач використовує оцінювання для заохочення студентів до самооцінки, щоб допомогти їм визначити свої сильні та слабкі сторони, учить їх установлювати особисті навчальні цілі.

Викладач оцінює вплив класу на особистість, збираючи інформацію шляхом спостереження за взаєминами в класі, опитуючи і аналізуючи студентські роботи.

Викладач контролює вплив власних навчальних методик і своєї поведінки на успіхи тих, хто навчається.

Викладач отримує інформацію про рівень знань студентів, їхню поведінку, потреби і прогрес у батьків, колег, безпосередньо в студентів і правильно використовує цю інформацію.

Викладач складає звіти про роботу студентів і може дати студентам, батькам і колегам надійну, основу на відповідних індикаторах інформацію про їхній прогрес.

Норма 9: Викладач — мислячий практик, який постійно оцінює вплив своєї діяльності на інших (студенти, батьки, інші професіонали в галузії освіти) і активно шукає можливості професійного зростання.

Компетентність. Викладач усвідомлює суть методів опитування, що дає йому змогу враховувати у практичній діяльності різні стратегії самооцінки і вирішення проблем, їхній вплив на навчання і прогрес студентів, взаємодію між ними.

Викладач ознайомлений з основними напрямками педагогічних досліджень і ресурсами професійного зростання (професійна література, колеги, професійні асоціації).

Здібності. Викладач оцінює критичне мислення і здатність до самовдосконалення.

Викладач постійно міркує, оцінює і вивчає. Він бажає надавати й отримувати допомогу. Викладач постійно шукає, розвиває й оновлює методики, що відповідають індивідуальним потребам студентів.

Викладач визнає свою професійну відповідальність за участь у впровадженні відповідних професійних методів як самостійно, так і за допомогою колег.

Продуктивність. Викладач використовує класні спостереження, інформацію про студентів і результати досліджень для оцінювання результатів навчання і як основу для досліджень, впливу на розробку та перегляд методик навчання.

Викладач використовує професійну літературу, поради колег та інші ресурси для самовдосконалення.

Викладач використовує досвід колег по школі, колег з інших професійних об'єднань при вирішенні проблем, активному поширенні досвіду і встановленні зворотного зв'язку.

Норма 10: Викладач підтримує стосунки з колегами, батьками і громадськими організаціями для підтримки здоров'я студентів і поліпшення їхнього навчання.

Компетентність. Викладач розглядає школи як громадські організації і усвідомлює відповідні аспекти діяльності системи, у якій він працює.

Викладач розуміє, як зовнішні фактори (сімейні обставини, суспільне середовище, здоров'я й економічні умови) можуть впливати на навчання і життя студентів.

Викладач розуміє і дотримується законів, що регулюють права студентів і обов'язки викладачів (наприклад, щодо рівного доступу до освіти, прав студентів-інвалідів, конфіденційності, належного поведження зі студентами, повідомлення про ситуації, пов'язані з жорстоким поведженням із дітьми).

Здібності. Викладач розуміє важливість всебічного розвитку дитини. Він піклується про всі аспекти здоров'я дитини (фізичного, емоційного, соціального та фізичного) і попереджає про труднощі.

Викладач готовий консультуватися із спеціалістами щодо навчання і благополуччя його студентів. Викладач поважає особисте життя студентів і конфіденційність інформації.

Викладач готовий співпрацювати з іншими професіоналами з метою поліпшення навчального середовища.

Продуктивність. Викладач бере участь у колегіальних діях, спрямованих на створення в школі продуктивного навчального середовища.

Викладач контактує від імені студентів з їхнім оточенням, консультуючись з батьками, адвокатами, іншими викладачами школи і професіоналами з інших громадських організацій. Для створення сприятливого навчального середовища він здатний ідентифікувати і використовувати суспільні ресурси.

Викладач установлює достойні і продуктивні відносини з батьками й опікунами, сприяє спільним діям на підтримку навчання і здоров'я студента.

Викладач спілкується зі студентами, чуйний до їхніх бід, досліджує ситуації і шукає необхідну для вирішення проблеми допомогу, діє як захисник студентів.

Національна Рада з акредитації програм підготовки викладачів

Національна Рада з акредитації програм підготовки викладачів — це професійний орган, що сприяє здійсненню високоякісної підготовки викладачів. NCATE оцінює якість навчання і підготовки викладачів шляхом професійної акредитації шкіл, коледжів і департаментів освіти.

NCATE — це коаліція 33 професійних асоціацій учителів, викладачів педвузів, фахівців із змісту освіти, місцевих і регіональних керівників, що є прихильниками якісного навчання — усього понад три мільйони осіб.

Стандарти NCATE встановлюють вимоги до знань, умінь і здібностей кандидатів [20]. Кандидати^а, що готуються до роботи в школах як викладачі або інші шкільні професіонали, мають знати зміст і демонструвати педагогічні і професійні знання, уміння і здібності, необхідні для допомоги студентам^б у навчанні. Оцінювання повинне показувати, чи відповідають кандидати професійним, державним та інституційним стандартам^с.

Таблиця 5. Приклад стандарту продуктивності вчительської праці

Елементи стандарту	Ступінь прояву		
	неприйнятна	прийнятна	запланована
1	2	3	4
Знання змісту для кандидатів у викладачі (початкова підготовка викладачів і підготовка, що продовжується)	Кандидати у викладачі мають неадекватне знання предмета, який вони планують викладати (нездатні дати приклади важливих принципів або концепцій, що містяться в професійному, державному й інституційному стандартах)	Кандидати у викладачі знають предмет, який вони планують викладати (здатні пояснити важливі принципи і концепції, що містяться в професійному, державному й інституційному стандартах)	Кандидати у викладачі показують глибоке знання предмета, який вони планують викладати (демонструють своє знання через дослідження, критичний аналіз і синтез змісту предмета відповідно до професійних, державних і інституційних стандартів)

1	2	3	4
Знання педагогіки для кандидатів у викладачі (початкова і триваюча підготовка викладачів)	Кандидати у викладачі не розуміють зв'язків змісту предмета і педагогіки (як це вимагається в професійному, державному й інституційному стандартах), що не дає їм змоги розробити методику навчання, яка б поєднувала технологію і враховувала б культурне оточення студентів і знання змісту	Кандидати у викладачі демонструють добре знання інституціональних методик навчання, що використовують зміст та педагогічні знання і навички, що містяться в професійному, державному й інституційному стандартах. Допомогають студенту у вивченні предмета шляхом представлення його змісту наочним і виразним чином та об'єднуючи різні методики	Кандидати у викладачі показують повне розуміння педагогіки (як це вимагається професійним, державним і інституційним стандартами). Вони демонструють усебічне розуміння предмета, який вони планують викладати, що дозволяє їм застосовувати різні методики навчання. Вони використовують відповідне поєднання методик і представляють зміст предмета привабливим, виразним і наочним
Професійні і педагогічні знання й уміння для кандидатів у викладачі (початкова і триваюча підготовка викладачів)	Кандидати не засвоїли професійні і педагогічні знання й уміння, що містяться в професійному, державному й інституційному стандартах; демонструють недостатні знання про школу, родину і суспільство або нездатність розробити навчальні завдання, що використовують попередній досвід студентів	Кандидати використовують свої професійні і педагогічні знання й уміння, що містяться в професійному, державному й інституційному стандартах, для полегшення навчання. При розробці змістовних навчальних завдань вони враховують попередній досвід студентів, особливості школи, родини і соціального середовища, у яких вони працюють	Кандидати показують повне розуміння професійних і педагогічних знань і умінь, що містяться в професійному, державному й інституційному стандартах. Для полегшення навчання вони розробляють змістовні навчальні завдання. Вони аналізують свою практичну діяльність і для поліпшення навчального процесу, вносять необхідні корективи. Вони знають, як студенти навчаються і як зробити навчальний

РОЗДІЛ. 4

1	2	3	4
			матеріал більш доступним для них. Вони розглядають особливості школи, родини і суспільства у зв'язку з попереднім досвідом студентів і пов'язують навчальний матеріал з навколишньою дійсністю
Здібності для всіх кандидатів	Кандидати не знайомі з професійними здібностями, що містяться в професійному, державному й інституційному стандартах. Вони не використовують ці здібності в роботі зі студентами, родинами, і групами	Кандидати знайомі зі здібностями, що очікуються відповідно до професійного, державного й інституційного стандартів від професіоналів, і це позначається на їхній роботі зі студентами, родинами і групами	Робота кандидатів зі студентами, родинами і групами відображає здібності, що очікуються від професійних педагогів відповідно до професійного, державного й інституційного стандартів. Кандидати усвідомлюють, коли їхні власні здібності повинні змінюватись, і здатні планувати дії, спрямовані на це
Навчання студентів кандидатами у викладачі (початкова і триваюча підготовка викладачів)	Кандидати у викладачі не можуть правильно оцінити знання студента або удосконалити навчальні завдання на підставі рівня розвитку студентів або їх попереднього досвіду	Кандидати у викладачі володіють методами навчання студентів. Це видно з того, як вони оцінюють результати навчання, використовують оцінювання в навчальному процесі і розробляють наочні навчальні завдання для студентів, що враховують рівень розвитку студентів і їх попередній досвід	Кандидати у викладачі правильно оцінюють і аналізують результати навчання, роблять відповідне корегування методики, контролюють навчання студентів і позитивно впливають на навчання всіх студентів

- a *Кандидати — це особи, що готують себе до викладацької діяльності, викладачі, котрі підвищують свою кваліфікацію, і особи, які готуються до іншої професійної діяльності в школі в ролі керівників, шкільних фізіологів і шкільних бібліотекарів.*
- b *До студентства входять студенти з особливими потребами і ті, що належать до різних етнічних, расових, гендерних, релігійних, соціально-економічних і регіональних/географічних груп.*
- c *Інституційні стандарти відображені в концептуальній структурі розділу і включають професіоналізм кандидата.*

Національна Рада із стандартів у галузі професійного викладання

Місією Національної Ради із стандартів у галузі професійного викладання є:

- *установлення високих і суворих стандартів для визначення того, що кваліфіковані викладачі повинні знати і вміти;*
- *розробка і використання національної системи добровільного оцінювання і сертифікації викладачів, що задовольняє вимоги цих стандартів;*
- *реформування системи освіти з метою поліпшення навчання студентів в американських школах.*

Керована Радою директорів із 63 членів, більшість яких є шкільними викладачами, Національна Рада сприяє зростанню поваги і визнання значущості викладацької праці. Система сертифікації Національної Ради, розроблена викладачами і для викладачів, є символом професійної вчительської майстерності. Вона здійснюється на добровільних підставах і доповнює, але не заміняє державне ліцензування. Тоді як системи ліцензування штатів установлюють стандарти вхідного рівня для викладачів-початківців, система сертифікації Національної Ради затверджує досить високі і жорсткі сучасні стандарти для досвідчених викладачів [21].

Національна Рада із стандартів розробила атестаційну систему, що встановлює досить високі і жорсткі вимоги. На основі цієї системи об'єктивно визначається, чи відповідає рівень кваліфікації вчителів цим стандартним вимогам. На сьогодні триває дослідження цієї системи, а її зміст і процес упровадження привертають увагу в усьому світі. Вона вже впливає на підготовку, акредитацію, ліцензування,

взаємодію між штатами, а також на зміст і престиж професії вчителя у суспільстві. Розглянемо її детальніше.

Система атестації, розроблена Національною Радою із стандартів, базується на трьох найважливіших елементах [22], це:

- *стандарти, які встановлюють єдині вимоги до того, що повинні знати й уміти кваліфіковані вчителі та, що означає кваліфікована вчительська практика. Стандарти класифікуються за конкретними напрямками спеціалізації вчителів;*
- *критерії оцінювання, що створюють надійну і об'єктивну систему оцінювання результатів навчання з конкретних предметів залежно від рівня розвитку учнів і сприяють набуттю вчителями практичного досвіду. Система оцінювання ефективності роботи вчителів на основі прийнятих стандартів якості відкриває широкі можливості для професійного вдосконалення і визначає його зміст;*
- *професійне вдосконалення, що передбачає розробку програм навчання і впровадження дидактичних методик у вчительську практику. Програми охоплюють попередню підготовку вчителів і підвищення кваліфікації в процесі роботи, допомагають учителям досягти найвищих рівнів атестації в рамках своєї професії, окреслюють відповідні вимоги.*

П'ять принципів кваліфікованого викладання

Стандарти і критерії оцінювання Національної Ради із стандартів у галузі професійного викладання спрямовані на визначення і офіційне визнання викладачів, які здійснюють ефективне навчання студентів і демонструють високий рівень знань, навичок, здібностей і нахилів. Вони засновані на найбільш загальному уявленні про відмінне викладання, яке можна описати за допомогою п'яти найважливіших принципів.

Принцип 1. Викладачі віддані студентам та їхньому навчанню.

Кваліфіковані викладачі роблять знання доступними для всіх студентів. Вони впевнені, що всі студенти можуть навчатися. Педагоги розуміють механізми розвитку і навчання своїх учнів. Учителі піклуються про учнів і їхнє навчання. Вони усвідомлюють, що учні різні, і відповідним чином пристосовують свою викладацьку стратегію до цих розходжень. Викладачі справедливо поводяться зі студентами, визнаючи індивідуальні риси, що відрізняють одного студента від іншого, і враховують ці особливості на практиці. Вони адаптують на-

вчальний процес, ґрунтуючись на досвіді і знанні інтересів студентів, їх здібностей, навичок, рівня знань і сімейних обставин.

Кваліфіковані викладачі розуміють, як студенти розвиваються і вчать. Вони використовують на практиці загальноприйняті теорії пізнання, усвідомлюють вплив середовища і культури на поведінку, розвивають пізнавальну здібність студентів і їхнє прагнення до навчання. Їхня місія не обмежується лише наданням знань учням, а поширюється на емоційні сфери мотивації і на зосередження уваги на особистості в цілому. Вони виховують у студентів почуття власної гідності, мотивацію, характер, громадську відповідальність і повагу до особистих, культурних, релігійних, і расових відмінностей.

Принцип 2. Викладачі знають свій предмет і те, як його викладати.

Кваліфіковані викладачі демонструють глибоке розуміння свого предмета, добре розуміються на структурі змісту предмета, його внутрішній логіці і зв'язках з іншими дисциплінами, знають його місце в загальній системі знань про світ. Підвищуючи загальний культурний рівень і наголошуючи на цінності предметних знань, вони розвивають також критичні й аналітичні здібності студентів. Учителі ставлять високі цілі в навчанні, володіють спеціальними знаннями, як побудувати навчання для їх досягнення.

Кваліфіковані викладачі володіють фаховими прийомами пояснення переданих студентам знань. Вони знають, з яким багажем знань студенти зазвичай приступають до вивчення кожного предмета, обізнані з методикою і навчальними матеріалами, що можуть бути корисними. Вони розуміють, де можуть виникнути труднощі, і застосовують таку стратегію навчання, яка дозволяє викласти обговорювану тему щонайкраще. Вони мають великий досвід навчання студентів методам вирішення їхніх власних проблем.

Принцип 3. Викладачі відповідають за контроль і управління навчанням студентів.

Кваліфіковані викладачі створюють, збагачують, підтримують і змінюють методику навчання для підвищення зацікавленості студентів і для найбільш ефективного використання часу. Їхнє першочергове завдання — полегшити учням навчання. Вони також мають великий досвід щодо залучення студентів і дорослих до навчального процесу, спроможні ефективно використовувати знання і досвід своїх колег.

Кваліфіковані викладачі володіють достатньою кількістю різних дидактичних методик: індивідуальна робота з окремими учнями і гру-

пами, участь учнів у процесі навчання, регулярне оцінювання успішності учнів. Вони знають особливості застосування кожної і можуть користуватися ними у разі необхідності. Вони також обізнані з неефективними або шкідливими методами.

Викладачі знають, як організувати необхідне для вивчення дисципліни освітнє середовище, як забезпечити досягнення навчальних цілей. Вони мають великий досвід щодо впровадження нормальної взаємодії між студентами і між студентами та викладачами. Вони знають, як спонукати студентів до навчання і як підтримати їхній інтерес навіть у випадку тимчасової невдачі.

Кваліфіковані викладачі здатні оцінити досягнення окремих студентів і класу в цілому. Вони використовують різні методи для визначення успішності студентів і можуть чітко охарактеризувати роботу студентів їхнім батькам.

Принцип 4. Викладачі систематично аналізують свою практичну діяльність.

Учителі постійно осмислюють свою практичну діяльність і вчаться на основі свого досвіду. Кваліфіковані викладачі, гідні наслідування як освічені особи. Прагнуть виховати в студентах зацікавленість, толерантність, чесність, справедливість, повагу до культурного розмаїття; сприяють розвитку розумових здібностей, що є передумовами інтелектуального зростання: здатності мислити і передбачати різні перспективи, до творчості і ризику, орієнтації на експериментування і вирішення проблем. Для прийняття принципових рішень кваліфіковані викладачі використовують знання законів людського розвитку, змісту предмета і методик, а також розуміння запитів своїх студентів. Їхні рішення базуються не тільки на теорії, але і на досвіді. Вони все життя вчаться і прагнуть привчити до цього студентів.

Для вдосконалення роботи викладачі використовують обґрунтовані рекомендації, а також результати дослідницької роботи. Прагнучи поліпшити методику викладання, кваліфіковані викладачі критично оцінюють свій досвід, прагнуть розширити світогляд і поглибити знання, удосконалюють формулювання, переробляють свою методику з урахуванням нових ідей і теорій.

Принцип 5. Викладачі є членами освітнього співтовариства.

Кваліфіковані викладачі роблять свій внесок у ефективність шкільної роботи, розробляючи разом з іншими професіоналами освітню полі-

тику, навчальні плани, вирішуючи питання підготовки кадрів. У міру розуміння освітніх цілей вони можуть оцінити досягнення школи і розподіл шкільних ресурсів. Викладачі інформовані про спеціальні шкільні і суспільні ресурси, що можуть бути використані для потреб студентів, і вміють користуватися цими ресурсами в разі необхідності.

Кваліфіковані викладачі знаходять способи спільної творчої роботи з батьками, залучаючи їх до продуктивної роботи в школі та використовуючи всі ресурси, що є у розпорядженні шкільних колективів.

Ці принципи є підґрунтям для аналізу стандартних вимог до викладачів і системи ліцензування з боку суспільства.

ОСОБЛИВОСТІ ЛІЦЕНЗУВАННЯ ВИКЛАДАЦЬКОЇ ДІЯЛЬНОСТІ У США

Визначення якості викладача надзвичайно важливе для розуміння ролі ліцензійного тестування. При цьому слід відрізнити якість **викладача** від якості **навчання**. Успішне навчання залежить від багатьох факторів, включаючи рівень доступних навчальних ресурсів, укомплектованість кадрами, тривалість навчання, підтримку адміністрації і батьків. Нас, однак, цікавить саме якість підготовки викладача — його знання, навички, здібності і нахили. Саме ці характеристики викладача як професіонала і як особистості мають бути оцінені й виміряні.

Ліцензійні вимоги до певної професії складаються з трьох компонентів: освітнього, практичного і діагностичного (тестування). Вони відрізняються за обсягом освітньої та практичної підготовки та послідовністю проходження. Призначенням ліцензійних вимог є забезпечення певного рівня знань і навичок, необхідного для здійснення практичної діяльності на початковому рівні. Останній не повинен бути занадто високим, щоб не встановлювати необґрунтовані обмеження.

Для аналізу досвіду ліцензування в інших професіях і визначення того, як він може бути застосований до ліцензування вчителів, розглянемо особливості ліцензування в інших професійних сферах та вимоги, що висуваються іншими професіями до освітнього, практичного і діагностичного компонентів [23]. Порівняємо загальні вимоги до ліцензування чотирьох професій: учителів, інженерів, фізіотерапевтів і соціальних працівників, наведені у табл. 6.

Таблиця 6. Загальні вимоги до ліцензування

Професії	Вимоги		Іспити	Форма іспиту	Досвід, що враховується замість освіти	Праця без виконання всіх ліцензійних вимог	Підвищ. кваліфікації
	освіта	досвід					
1	2	3	4	5	6	7	8
Учитель	Усі штати вимагають ступінь бакалавра педагогіки або споріднену; більшість дозволяє післядипломні шляхи до педагогічної діяльності	Більшість штатів вимагає досвіду педагогічної діяльності від 9 до 18 тижнів [24]	Існує близько 600 іспитів; більшість штатів вимагає один або більше тестів з базових навичок, загальних знань, знань з предмета або педагогічних знань	Головним чином — багатовибіркові та зі сконструйованою відповіддю	Ступінь бакалавра є мінімальною вимогою для всіх штатів. Вимоги до освіти і досвіду для програм альтернативного ліцензування значно відрізняються у різних штатах	Більшість штатів дозволяє нелицензованим індивідуалам тимчасово не виконувати одну або більше з ліцензійних вимог [25]	Більшість штатів вимагає [24]
Інженер	Більшість штатів вимагає ступінь бакалавра інженерії	Усі штати вимагають наявність досвіду, як правило — чотири роки	Іспит із двох частин; одна може бути складена до отримання ступеня, а інша — після; приблизно половина штатів вимагає складання іспитів із права, етики або спеціальності	Багатовибіркові; 16 предметів; 8 годин на кожну частину; для другої частини допускається користування підручником	Приблизно половина штатів дозволяють заміну досвіду інженерним ступенем; інші вимагають 10 років роботи	Нелицензовані особи можуть виконувати інженерну діяльність певних видів без дозволу засвідчувати робочі креслення	Менш ніж половина штатів вимагає

1	2	3	4	5	6	7	8
Фізіотерапевт	Усі штати вимагають виконання відповідних акредитованих програм; після 2002 р. — тільки після дипломний рівень освіти	Деякі штати вимагають клінічну практику	Один іспит; деякі штати вимагають усного або практичного іспиту до ліцензування	Багатовибіркові тестові завдання; комп'ютерне тестування; чотири години	Досвід не може бути замінений освітою	Приблизно половина штатів дозволяє індивідуумам, що задовольняють освітнім вимогам, тимчасову практичну діяльність до проходження тестування	Приблизно половина штатів вимагає
Соціальний працівник	Чотири категорії практики; вимоги до практики змінюються залежно від штату. Більшість штатів вимагають ступеня бакалавра	Вимоги до досвіду залежать від рівня практики	Чотири рівні іспитів: базовий, перехідний, підвищений та клінічний; кандидати складають тільки один для кожного рівня ліцензування	Багатовибіркові тестові завдання; комп'ютерне тестування; чотири години за рівень	Більшість штатів не зараховує досвід замість освіти	Існує чотири рівні практики, що визначають послідовність вимог до освіти/досвіду	Більшість штатів вимагає (15 — 20 годин кожні два роки)

Перш за все треба зауважити, що для всіх цих професій ліцензуючою структурою є штат. Саме штат для кожної професії (у тому числі і для викладачів) установлює певні вимоги до освіти і досвіду практичної діяльності, необхідні для одержання ліцензії. Є, однак, різниця, між вимогами штату до ліцензійного тестування вчителів та вимогами до ліцензійного тестування інженерів, фізіотерапевтів і соціальних працівників. Для кожної з трьох останніх професій у всіх штатах використовуються ті ж самі тести або серії тестів (хоча деякі штати доповнюють національні іспити додатковими компонентами). Для двох із цих професій (інженерів і фізіотерапевтів) прохідні бали для тестування однакові по всій країні [26]. Для викладачів складається суттєво інша картина. При тестуванні останніх вимоги розрізняються у штатах, за предметними областями та рівнями знань. Прохідні бали також різні в різних штатах.

Що стосується вимог до здійснення викладацької діяльності без ліцензії, то вона допускається, як це робиться і в інших професійних сферах. Однак є певні обмеження. Так, викладач без ліцензії може працювати у деяких штатах тільки у приватних школах [27]. Багато штатів дозволяють неліцензованим учителям тимчасово викладати також і в середніх школах, якщо там для заповнення вакантних посад не вистачає ліцензованих викладачів.

Як і в інших галузях, акредитація програм підготовки викладачів є механізмом перевірки й атестації їхньої якості. Однак, на відміну від інших професій, для програм підготовки викладачів не вимагається акредитація в національних професійних організаціях. Більшість штатів визнає тільки програми підготовки викладачів, які акредитовані в місцевих організаціях. На національному рівні акредитовано менш ніж 40% програм підготовки викладачів [28].

У понад 40 штатах програми підготовки викладачів у коледжах і університетах можуть бути акредитовані як штатом, так і на національному рівні. Відповідальність за акредитацію покладена на Національну Раду з акредитації. Остання являє собою 33 спеціалізовані професійні асоціації викладачів, методистів і законодавців локального рівня і рівня штату. Акредитація на національному рівні може бути замінена акредитацією на рівні штату. Крім того, штати можуть установлювати свої особливі вимоги до зарахування на програми підготовки викладачів, а також до навчальних дисциплін і навчальних/педагогічних практик. Замість вимог до навчальних дисциплін деякі штати визначають стандарти на знання та вміння, які мають продемонструвати випускники. Ці стандарти, як правило, відповідають вимогам INTASC.

Окрім цього, більшість штатів вимагають від кандидатів на отримання ліцензії успішного проходження одного чи більше тестувань із ба-

зових навичок, загальних знань, знань із предмета або з педагогічної компетентності. Вимоги до отримання ліцензії з боку штатів можуть включати американське громадянство, вікові обмеження, наявність рекомендації тощо.

Більшість штатів має дво- або триступеневу систему ліцензування, що передбачає додаткові вимоги до отримання кожного типу ліцензії. У 31 штаті вимагається наявність початкової ліцензії (що діє від двох до п'яти років) і одержання при виконанні додаткових вимог стандартної професійної ліцензії. Ще у 13 штатах вимагається додатковий атестат підвищеного рівня. Тільки три штати надають довічну ліцензію підвищеного рівня. Більшість додаткових вимог до отримання ліцензії стосується шляхів одержання вченого ступеня (вище бакалавра) або подальшого професійного зростання. Деякі штати для отримання ліцензії наступного рівня вимагають наявності необхідної викладацької практики.

Застосування додаткових вимог до отримання ліцензії зменшує загальну кількість дипломованих викладачів, яка необхідна для заповнення вакантних посад. Особливо це стосується викладання математики, природничих наук, двомовної та спеціальної освіти у деяких міських і сільських районах. У зв'язку з цим майже всі штати дозволяють видавати обмежені ліцензії для тимчасового залучення викладачів, що не склали деякі тести. Крім того, для забезпечення ліцензійної мобільності 40 штатів уклали угоду щодо визнання документів про завершення програми підготовки вчителів і ліцензій, виданих одним штатом, на території іншого.

Ліцензійні педагогічні тести, що використовуються у США

Зупинимось детальніше на вимогах до тестування — на діагностичному компоненті. Практика тестування викладачів, проходження якого необхідно для одержання ними ліцензії, має свої особливості в кожному штаті. Для перевірки різних аспектів викладацької діяльності використовуються різні тести. Концепція добору змісту іспитів заснована на урахуванні тих знань і навичок, якими повинні оволодіти викладачі. Подібно до ліцензійних тестів у інших професійних сферах, акцент робиться на перевірці комплексу знань і навичок, відібраних експертними групами педагогів-методистів. Вони визнані мінімально необхідними для того, щоб розпочати викладацьку діяльність, і не призначені для визначення рівня (середній, високий) кваліфікації викладачів.

На сьогодні 42 штати при прийнятті різних адміністративних рішень щодо кандидатів у викладачі використовують результати ліцензійного тесту-

РОЗДІЛ. 1

вання. Останнє є обов'язковим при зарахуванні на програму перепідготовки викладачів, для підтвердження успішного завершення навчання і при одержанні ліцензії на професійну (викладацьку) діяльність. Кожен штат при цьому вибирає свої тести і встановлює власні прохідні бали [23]. Зараз використовується кілька сотень ліцензійних тестів для викладачів-початківців. Цими тестами визначається практично усе: починаючи від основних навичок з читання, писемності та математики — до знання конкретного предмета, а також загальних понять або знань із дидактики навчання. Формати тестових завдань (ТЗ) можуть бути різними — від багатовибіркових (multiple-choice) до ТЗ зі сконструйованою відповіддю (constructed-response). Таке тестування позитивно впливає на кандидатів у викладачі та потенційно на студентів американських коледжів і шкіл.

Для отримання ліцензії більшість штатів вимагають складання від одного тестового іспиту з базових навичок (Алабама) до чотирьох різних іспитів, що включають тестування з базових навичок, педагогічної компетентності і предметних знань. Кожен із цих іспитів може складатися більше ніж з одного тестового екзамену.

Існують два основні розробники тестів для ліцензування вчителів: ETS³⁴ і NES³⁵. ETS є розробником тестів серії «Praxis» («Праксис»). У табл. 7 представлено розподіл кількості штатів, що використовують різні типи тестових іспитів для отримання початкової ліцензії, за їх розробниками (1998 — 1999 рр.) [23]. Порівняння робиться для п'яти різних типів ліцензійного тестового іспиту: базові навички, загальна компетентність, знання з предмета, педагогічна компетентність, педагогічна компетентність з предмета.

Таблиця 7. Кількість штатів, що використовують різні типи тестових іспитів залежно від розробників

Розробник	Тип тестового іспиту					Усього штатів
	базові навички	загальна компетентність	знання з предмета	педагогічна компетентність	педагогічна компетентність з предмета	
ETS	28	10	20	20	7	32
NES	9	4	9	6	—	8
Штат	1	—	2	2	—	2
Усього	38	14	31	28	7	42

³⁴ Служба освітнього тестування — *Educational Testing Service*.

³⁵ Системи національного оцінювання — *National Evaluation Systems*.

Із табл. 7 видно, що 32 штати використовують один або більше компонентів ETS серії «Praxis», 8 штатів для ліцензування своїх майбутніх учителів використовують тести NES. Деякі штати розробляють тести самостійно.

У табл. 8 представлено розподіл різних типів тестових іспитів (тестів) для отримання початкової ліцензії, що використовуються у різних штатах, за їх розробниками (1998 — 1999 рр.) [23]. Порівняння робиться для тих же п'яти різних типів ліцензійного тестового іспиту.

Найбільша кількість штатів для перевірки базових навичок з писемності, комунікативності та математики використовує тестування. У 38 штатах для отримання первинної ліцензії кандидати мають відповідати мінімальним вимогам базових навичок. У 28 штатах використовують першу частину серії «Praxis-I», що оцінює базові навички з математики, читання і письма. Існують як паперові, так і комп'ютерні версії. У 14 штатах використовують тестування для перевірки **загальної компетентності**, що передбачає оцінку знань із гуманітарних наук на додипломному рівні. Ці тести можуть також перевіряти загальнонаукову освіченість з математики, історії, суспільних наук. Серія «Praxis-II» цих тестів складається з трьох частин, що перевіряють: загальну компетентність, професійну компетентність і комунікаційні навички. У 31 штаті від кандидатів у викладачі вимагається складання тестів із загальних знань з предмета.

Таблиця 8. Розподіл тестових іспитів (тестів) за типами та за їх розробниками

Розробник	Тип тестового іспиту					Усього штатів
	базові навички	загальна компетентність	знання з предмета	педагогічна компетентність	педагогічна компетентність з предмета	
ETS	7	1	126	4	6	144
NES	27	4	362	12	—	405
Штат	6	—	69	1	—	76
Усього	40	5	557	17	6	625

З табл. 8 видно, що загальна кількість тестів, що використовуються, складає понад 600. ETS адмініструє 144 різних тести, а NES — понад 400 різних тестових іспитів для вчителів. Понад 70 різних тестових екзаменів було розроблено штатами самостійно.

Для перевірки базових навичок зараз використовується 40 різних тестів, а для перевірки загальної компетентності — тільки 5. Проте іс-

нує понад 500 різних тестів для перевірки загальних знань з предмета. Як і у випадку тестування для перевірки базових навичок і загальної компетентності, тут можуть використовуватися тести від різних розробників, зокрема — 126 ETS тестів з конкретних предметів, що охоплюють 50 різних предметних галузей (серія «Praxis-II»), і близько 360 NES тестів. Причому ETS нещодавно оголосила про оновлення тестів із конкретних предметів серії «Praxis-II» з метою приведення їх у відповідність до стандартів і включення більшої кількості ТЗ з педагогіки.

Тестування з **педагогічної компетентності** перевіряє рівень педагогічної освіченості майбутніх учителів, знання і навички з психології освіти, методики планування процесу викладання, методики здійснення викладання, методики оцінювання результатів навчання, управління тими, хто навчається, і освітнім середовищем. Воно охоплює такі питання, як подання змісту навчання у зручному для викладання вигляді, створення сприятливого освітнього середовища, володіння методиками викладання, учительський професіоналізм. 28 штатів використовують тестування з педагогічної компетентності. 20 — із них застосовують для цього компоненти серії «Praxis-II» «Принципи навчання і викладання», що розроблені для трьох рівнів: К-6, 5-6 та 7-12 [29]. Тестування комунікативних навичок майбутніх учителів перевіряє їх здатність добре слухати, читати і писати. Використовується 17 різних тестів для перевірки педагогічної компетентності. Переважна більшість їх розробляється NES.

Нещодавно INTASC були розроблені нові тести з педагогічної компетентності, що отримали назву ТТК³⁶. Дослідницька версія ТТК включає ТЗ, побудовані за INTASC принципами. Ці тести включають ТЗ, що пов'язані з дослідженням випадків (case study exercise), оглядом документів, а також ТЗ зі сконструйованою відповіддю, які стосуються навчання, розвитку, мотивації, сценаріїв навчання і прикладів учнівських робіт. Останні дають змогу оцінити, чи розуміють майбутні викладачі основи розвитку дитини, мотивацію їхньої поведінки, теорію навчання, загальні труднощі у навчанні, принципи управління класом і стратегії оцінювання. ТЗ із дослідження випадків стосуються ретельного розгляду ситуацій, що трапляються у класі, а ТЗ з огляду документів представляють собою підбір документів, пов'язаних із комплексом навчальних ситуацій в окремій класній кімнаті. Зараз проходить пілотне дослідження цих тестів.

Тільки 7 штатів вимагають від майбутніх учителів складання тестів з **педагогічної компетентності з предмета**. Тестування з педа-

³⁶ Тести викладацької компетентності — *Test of Teaching Knowledge*.

логічної компетентності з предмета перевіряє аспекти педагогічної компетентності, пов'язані і зі змістом навчання предмета, і з педагогікою. Воно стосується не стільки фактичних знань з предмета і формулювання тем, що робить їх зрозумілими для слухачів, скільки Єшляхів подачі цих знань і формулювання завдань для пояснення матеріалу учням. Це також стосується розуміння того, що полегшує або утруднює вивчення учнями того чи іншого розділу. Наприклад, спотворене уявлення про Сонячну систему може перешкодити вивченню астрономії. На жаль, існує тільки шість різних тестів для перевірки педагогічної компетентності з предмета. Усі вони розроблені ETS (серія «Praxis II») і включають тести з біології, іноземної мови, математики, фізики, суспільних наук та іспанської мови.

ВИСНОВКИ

Таким чином, ми бачимо, що в різних країнах існують різні критерії визначення вимог до викладача як з боку суспільства, так і з боку суб'єктів навчання та їхніх батьків. Проте акцентується, що вчитель:

- повинен знати свій предмет;
- здатний передати іншим свої знання, навички, досвід;
- спроможний здійснювати індивідуальне і групове навчання учнів, приділяючи основну увагу загальному розвитку дитини;
- може проводити оцінювання досягнутих результатів;
- спроможний відповідати за управління процесом навчання і здійснювати контроль над ним;
- може визначати освітні цілі шкільних навчальних програм, аналізувати їх і переглядати у міру необхідності;
- здатний осмислювати свою практичну діяльність і вчитися на основі свого досвіду;
- у співдружності з іншими професіоналами сприяє ефективній роботі школи.

У США діють три взаємопов'язані системи стандартів, які описують послідовні етапи викладацької кар'єри. Стандарти INTASC описують знання, навички, здібності і схильності викладачів-початківців. Стандарти NCATE спрямовані на вимоги до програм підготовки викладачів і опи-

сують їхні знання і навички. Стандарти NBPTS визначають вимоги до викладачів, які вже займаються професійною діяльністю.

Базові стандарти INTASC складаються з 10 норм, яким має відповідати викладання незалежно від предмета або рівня, кожна з яких містить вимоги до компетентності, здібностей і майстерності викладача.

Стандарти NCATE оцінюють якість навчання і підготовки викладачів шляхом професійної акредитації шкіл, коледжів і департаментів освіти. Вони визначають вимоги до знань, умінь і можливостей кандидатів, що планують працювати у школах як викладачі або інші шкільні професіонали. Кожен елемент стандарту оцінюється за ступенем прояву: неприйнятний, прийнятний, запланований.

Система сертифікації NBPTS здійснюється на добровільній основі і доповнює, але не заміняє державне ліцензування. Тоді як системи ліцензування штатів визначають стандарти вхідного рівня для викладачів-початківців, система сертифікації NBPTS установлює стандарти для досвідчених викладачів.

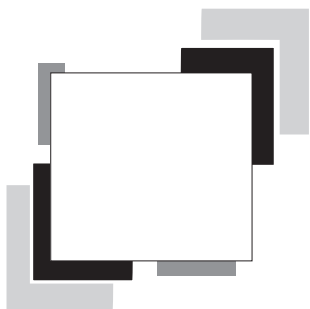
Ці системи стандартів є основою для реформи системи підготовки і підвищення професійної кваліфікації викладачів у США.

Основною метою ліцензування викладачів-початківців є гарантія їхньої компетентності. У США ліцензування викладацької діяльності, добір тестів і вибір прохідного бала належить до компетенції окремих штатів. Тести, що використовуються на сьогодні, вимірюють базові навички, загальну компетентність, знання з предмета, педагогічну компетентність і педагогічну компетентність з предмета, але рішення про видачу ліцензії ґрунтується не тільки на результатах тестування.

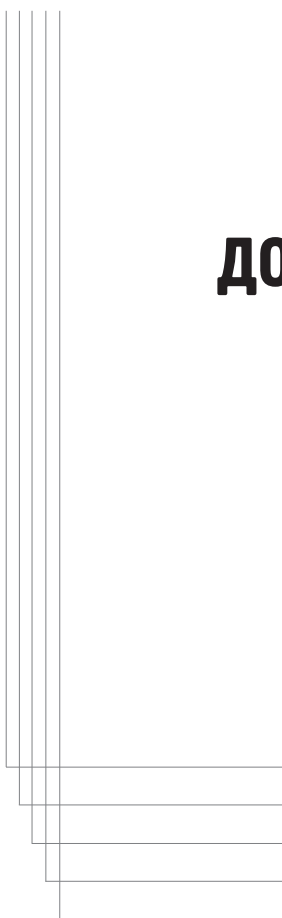
Більшість штатів визнає тільки програми підготовки викладачів, які акредитовані в місцевих організаціях. Кожен штат установлює різні вимоги до освіти і досвіду практичної діяльності, які необхідні для одержання ліцензії на здійснення викладацької діяльності.

Ліцензійні тести для викладачів-початківців призначені для визначення кандидатів, які мають знання і навички, мінімально необхідні для початку викладацької практики, і відокремлення їх від тих, хто таких знань і навичок не має. Такі тести не дають змоги відрізнити середньокваліфікованих кандидатів у викладачі від висококваліфікованих. Вони не призначені також для оцінювання всіх компетентностей, необхідних для початку практичної діяльності.

Формати тестових завдань можуть бути різними: багатовибіркові ТЗ (multiple-choice), ТЗ зі сконструйованою відповіддю (constructed-response), ТЗ на дослідження випадків (case study exercise), портфоліо.



ДОДАТКИ



- A.** Міжнародні документи з регулювання педагогічного оцінювання.
- B.** Інші регуляторні документи.
- C.** Англомовні варіанти стандартів.

А. МІЖНАРОДНІ ДОКУМЕНТИ З РЕГУЛЮВАННЯ ПЕДАГОГІЧНОГО ОЦІНЮВАННЯ

Стандарти освітнього та психологічного тестування

Стандарти освітнього та психологічного тестування [1] (AERA, APA, NCME) розроблено спільно з:

Американською асоціацією освітніх досліджень (AERA).

Американською психологічною асоціацією (APA).

Національною Радою з вимірювань в освіті (NCME).

Стандарти освітнього та психологічного тестування є сучасним регуляторним документом, який увібрав у себе принципи та завдання щодо рекомендацій із розробки та використання тестів, представлені у попередніх документах.

Першим документом були **Технічні рекомендації для психологічних тестів і діагностичних методик** (*Technical Recommendations for Psychological Tests and Diagnostic Techniques*), підготовлені Комісією Американської психологічної асоціації (APA) та опубліковані нею ще у 1954 р.

Другим документом були **Технічні рекомендації для тестів з оцінювання успішності** (*Technical Recommendations for Achievement Tests*), підготовлені Комісією при Американській асоціації освітніх досліджень (AERA) та Національною радою вимірювань в освіті та опубліковані Національною Освітньою Асоціацією у 1955 р.

Третім документом були **Стандарти для освітніх і психологічних тестів та підручників** (*Standards for Educational and Psychological Tests and Manuals*). Цей документ замінив два попередні. Його було підготовлено Комісією, створеною APA, AERA, Національною радою з вимірювань в освіті (NCME) і опубліковано APA в 1966 р. Четвертий документ — **Стандарти для освітніх та психологічних тестів** (*Standards for Educational and Psychological Tests*) — є переглянutoю версією третього документа. Він був під-

готовлений трьома спонсорськими організаціями та опублікований у 1974 р.

До 1977 р. стало очевидно, що видання 1974 р. застаріває через виникнення нових проблем та питань у зв'язку з розвитком тестування. Технічний прогрес у тестуванні та споріднених галузях, нові сфери застосування тестів і зростаюче громадське занепокоєння щодо ролі тестування в досягненні соціальних цілей обумовили необхідність перегляду Стандартів для освітніх та психологічних тестів від 1974 р. Відповідно Комісія з психологічного тестування та оцінювання американської психологічної асоціації (APA Committee on Psychological Tests and Assessment) звернулася до AERA та NCME з метою утворення спільної Комісії для перегляду Стандартів 1974 р. Редакційна комісія, у свою чергу, визначила вимоги до Стандартів, які повинні:

1. *Стосуватися питань щодо використання тестів для різноманітного призначення.*
2. *Бути переліком технічних стандартів для належної професійної практики, а не приписом соціальної діяльності.*
3. *Уможливлювати визначення технічної адекватності тесту, доцільність та належність конкретного застосування і раціональність (reasonableness) висновків, зроблених на основі результатів тестування.*
4. *Зобов'язувати розробників тестів, видавців та користувачів збирати та надавати достатню інформацію, яка дасть можливість кваліфікованому оглядачеві (reviewer) визначити дотримання стандартів, що вимагаються.*
5. *Бути втіленням сильного етичного імперативу, хоча, за погодженням, самі Стандарти не повинні містити механізмів примусу.*
6. *Визнавати, що не всі стандарти будуть однаково впроваджуватися для всього широкого спектру інструментів та застосувань.*
7. *Бути представленими на такому рівні, який дасть можливість широкому колу людей, що працюють з тестами чи результатами тестування, застосовувати ці Стандарти.*
8. *Не пригнічувати експериментаторство у розробці, використанні тестів та інтерпретації результатів.*
9. *Відображати існуючий рівень консенсусу визнаних експертів.*

10. Бути документом-наступником Стандартів для освітніх та психологічних Тестів від 1974 р. (1974 Standards for Educational and Psychological Tests).

Стандарти освітнього та психологічного тестування від 1999 р. зазнали значних змін порівняно із Стандартами освітнього та психологічного тестування від 1985 р. Кожний розділ більш обґрунтований, включає більшу кількість стандартів, значно розширений список термінів/глосарій та показчик. Нові **Стандарти** відображають зміни у федеральному законодавстві та тенденції в оцінюванні, що впливають на: валідність тесту; тестування осіб з обмеженими можливостями чи з іншого мовного середовища; нові типи тестів, а також нове використання чинних тестів. **Стандарти** написані для професіоналів та освічених непрофесіоналів і стосуються професійних та технічних питань розробки і використання тестів в освіті, психології та у працевлаштуванні. Ці стандарти є надзвичайно важливим документом для професійних розробників тестів, спонсорів, видавців, користувачів, розробників політики та процедур (policymakers), роботодавців, студентів з педагогіки та психології.

Мета Стандартів

Нещодавні суперечки довкола тестування роблять розробку Стандартів життєво необхідною для всіх учасників процесу тестування. Багато питань повинні вирішуватися таким способом, який не заважає дотриманню політики. Загалом Стандарти відстоюють таку позицію: необхідна технічна інформація надається в розумних межах сторонам для їх повної поінформованості у разі дискусії з приводу політичних питань. Стандарти не ставлять за мету дати психометричні відповіді на політичні запитання. Проте повне розмежування наукових та соціальних проблем неможливе. Вимоги Стандартів щодо документування та наукового аналізу часом можуть ускладнити певні питання щодо політики.

Застереження щодо застосування Стандартів

Стандарти — це технічні рекомендації, які можуть використовуватися як основа для оцінювання практик тестування. Оцінка прийнятності тесту чи його застосування не полягає у буквальній відповідності кожному базовому стандарту з цього документа, прийнятність не може бути визначена за переліком вимог (check list).

Використання цих Стандартів у судових справах уникнути неможливо; проте слід наголосити, що в юридичних процесах, як і будь-де, професійне вирішення на основі загальноприйнятого масиву знань завжди відіграє основну роль у визначенні доцільності конкретних стандартів в конкретних ситуаціях. Стандарти мають характер рекомендацій для прийняття таких рішень. Розробникам чи користувачам тестів некоректно буде стверджувати, що тест, підручник чи процедура задовольняє вимоги цих стандартів чи відповідає їм. Такі рішення повинні прийматися іншими особами в професійному колі.

Ці Стандарти стосуються галузі, яка розвивається, тому існує постійна потреба моніторингу та перегляду цього документа в міру розвитку знань. Новий розвиток найбільш імовірний у таких сферах: норми за гендерною ознакою чи змішані (gender-specific or combined-gender norms), культурні відхилення, комп'ютерна інтерпретація результатів, генералізація валідності, передбачення відмінності (differential prediction) та зниження тестового бала (flagging test scores) для осіб з обмеженими можливостями.

Категорії Стандартів

Важливість деяких Стандартів для побудови тестового іспиту (test construction) та оцінювання змінюватиметься залежно від застосування. Ці Стандарти визначаються як умовні. Такі Стандарти повинні вважатися первинними для одних ситуацій і вторинними для інших. Тобто, щоб вирішити як сприймати окремий умовний стандарт, як первинний чи вторинний, необхідно ретельно зважити можливість дотримання цього стандарту у зв'язку з потенційними наслідками для всіх учасників процесу тестування. Для деяких тестових програм дотримання певних умовних стандартів може бути нереальним з технічної чи фінансової точки зору, особливо для тих програм, які проводять тестування в незначних обсягах. Проте якщо використання тесту може мати серйозні наслідки для екзаменованих, особливо коли їхня кількість досить велика, умовні стандарти стають особливо важливими.

Тести та користувачі тестів, які стосуються Стандартів

Хоча ці Стандарти стосуються у першу чергу сконструйованих завдань на успішність (constructed performance tasks), анкет/опитувальних форм (questionnaires) і **структурованих** зразків поведінки (structured behavior samples), вони можуть успішно застосовуватися

до всього спектру оцінювальних технік. Проте загалом буде неможливо безоглядно застосувати ці Стандарти до широкого спектру **неструктурованих** зразків поведінки (unstructured behavior samples), що використовуються в деяких формах клінічного та шкільного психологічного оцінювання, або до тестів, розроблених викладачем (instructor-made tests), що використовуються для оцінювання успішності студентів з освітнього курсу чи тренінгу.

Структура документа

Документ складається з трьох частин.

Частина I — «Технічні Стандарти для побудови тестів та оцінювання» містить стандарти валідності, надійності, розробки тестів, шкалювання, нормування, порівнюваності, урівнювання та видання/публікації. **Частина II** — «Професійні стандарти для використання тестів». **Частина III** — «Стандарти для окремих випадків застосування» містить стандарти для тестування мовних меншин та людей з обмеженими можливостями. Технічні та професійні стандарти, що розглядаються в Частинах I і II, стосуються окремих випадків застосування Частини III.

Хоча у всіх частинах розглядаються технічні питання, у кожній з них ідеться також про те, як технічні аспекти, професійна практика та професійна етика взаємодіють між собою.

Зміст документа

Частина I: Побудова тесту, оцінювання та документування.

Валідність.

Надійність та похибка вимірювання.

Розробка та перегляд тесту.

Шкали, норми та порівнюваність балів.

Адміністрування тесту, розрахунок бала, звітність.

Допоміжна документація для тестів.

Частина II: Справедливість у тестуванні.

Справедливість у тестуванні та використанні тестів.

Права та обов'язки екзаменованих.

Екзаменовані з різних мовних середовищ.

Екзаменовані з обмеженими можливостями.

Частина III: Застосування тестування.

Відповідальність користувачів тестів.
Психологічне тестування та оцінювання.
Освітнє тестування та оцінювання.
Тестування у працевлаштуванні та підтвердженні кваліфікації.
Тестування в оцінюванні програм та політика щодо громадськості (Public Policy).

ETS Стандарти якості та справедливості

ETS Стандарти якості та справедливості [2] були покладені в основу корпоративної політики Радою повірених ETS у 1981 р. і продовжують служити еталоном (benchmark) якості для всіх продуктів та послуг ETS. Ці Стандарти були переглянуті у 2000 р., слідом за переглядом Стандартів освітнього та психологічного тестування, опублікованих Американською асоціацією освітніх досліджень (AERA), Американською психологічною асоціацією (APA) та Національною радою з вимірювань в освіті (NCME).

Відмінності у цих наборах стандартів відображають специфіку продуктів та послуг ETS. Стандарти ETS використовуються персоналом ETS щоденно, а успішність і всебічність їхнього дотримання перевіряється жорсткою програмою аудиту.

Рада повірених ETS переглядає результати програми аудиту. Для того, щоб ETS залишався світовим лідером у сфері забезпечення справедливого та валідного оцінювання, досліджень та відповідних продуктів і послуг, ми повинні продовжувати нести персональну відповідальність за досягнення найвищого рівня якості та справедливості в наших продуктах.

Ці продукти створюються з метою допомоги окремим особам, батькам, викладачам, навчальним закладам, бізнес-структурам, урядам, країнам, штатам, шкільним округам, спеціалістам в галузі вимірювань та дослідникам.

ETS Стандарти якості та справедливості у редакції 2000 р. багато чим завдячують попереднім версіям Стандартів ETS, перша з яких була прийнята Радою повірених у 1981 р. та оновлена у 1987 та 2000 рр. Нові Стандарти ETS мають два специфічні рівні: перший належить до основної мети кожного розділу, а другий — до власне стандартів.

ETS намагається дотримуватися Стандартів освітнього та психологічного тестування, спільно опублікованих Американською асоціацією освітніх досліджень (AERA), Американською психологічною асоціацією (APA) та Національною радою з вимірювань в освіті (NCME). Через це Стандарти ETS розроблялися відповідно до розділів Спільних стандартів, проте були адаптовані до особливих потреб та умов ETS. Стандарти ETS більше концентрують увагу на бажаних результатах і менше — на конкретних способах досягнення цих результатів порівняно із Спільними стандартами. Стандарти ETS включають, наприклад, матеріал, про службу підтримки клієнтів (customer service), якого немає в Спільних стандартах. Більше того, Стандарти ETS вилучають матеріал, який хоч і міститься в Спільних стандартах, проте не стосується продуктів чи послуг ETS.

Після вступу, цього огляду та опису процесу аудиту йдуть 13 розділів. Словник термінів (глосарій) пояснює технічні та спеціальні терміни. Таблиця підсумовує особливості застосування стандартів у кожному розділі до оцінювальних та неоцінювальних продуктів і послуг.

Зміст документа

Розділ 1. Процедури розробки.

Розділ 2. Придатність до використання.

Розділ 3. Служба підтримки клієнтів.

Розділ 4. Справедливість.

Розділ 5. Використання та захист інформації.

Розділ 6. Валідність.

Розділ 7. Розробка оцінювання.

Розділ 8. Надійність.

Розділ 9. Урізані бали (cut scores), шкалювання та урівнювання.

Розділ 10. Адміністрування оцінювання.

Розділ 11. Звітність про результати оцінювання.

Розділ 12. Використання оцінювання.

Розділ 13. Права та обов'язки екзаменованих.

Глосарій

Окремі тематичні розділи можуть викликати певне непорозуміння. Справедливість, наприклад, є невід'ємним аспектом тестування загалом, і стандарти, що стосуються справедливості, доцільно було б включати в кожний розділ. Виокремлення стандартів справедливості

в окремий розділ мало на меті підкреслити їх важливість, а не ізолюваність від інших питань.

Рекомендації з освітнього та психологічного тестування

Під час підготовки Рекомендацій з освітнього та психологічного тестування [3] Канадська психологічна асоціація (CPA) узяла до уваги стандарти, рекомендації, звіти, принципи та політичні підходи, розроблені різними організаціями, зокрема Американською психологічною асоціацією, а також численні коментарі до цих документів. Дані рекомендації були сформульовані відповідно до Стандартів для освітніх та психологічних тестів, розроблених Американською психологічною асоціацією у 1974 р. Хоча останній документ послужив вагомим внеском у розвиток педагогічного оцінювання і є корисним для користувачів тестів в усьому світі, CPA вирішила підготувати власні технічні рекомендації для належної професійної практики, які більше прив'язані до канадського контексту. Ці рекомендації відображають наявний рівень консенсусу серед канадських професіоналів та як такі використовуються в Канаді замість Стандартів APA. Але потрібно пам'ятати, що Рекомендації не залишаються незмінними. Вони повинні застосовуватися на практиці та переглядатися; коментарі щодо внесення змін та поліпшення дадуть змогу Рекомендаціям продовжувати розвиватися та залишатися максимально корисними для цієї професії.

Мета згаданих Рекомендацій — надати критерії оцінки тестів, тестових практик та впливу використання тестів.

Частина I цих Рекомендацій містить загальні настанови щодо валідності, надійності, розробки тестів, порівнюваності, нормування, шкалювання та видання/публікації. **Частина II** представляє рекомендації щодо використання тестів. **Частина III** представляє рекомендації щодо адміністративних процедур. Усі три частини дають технічні рекомендації, і водночас в кожній із них приділяється увага взаємопов'язаності технічних питань, професійної практики та професійної етики. Питанням професійної практики приділяється більше уваги в Частині II, а питанням професійної етики — у Частині III.

Кожний розділ починається вступним текстом, який обґрунтовує наступні рекомендації. Цей текст має на меті допомогти *інтерпретації*

рекомендацій і не нав'язує додаткових порад. Багато рекомендацій мають коментарі, які не містять додаткових вимог, а тільки ілюструють відповідні рекомендації.

Зміст документа

Частина I. Тестовий інструментарій.

Валідність.

Надійність та похибка вимірювання.

Розробка та перегляд тесту.

Шкалювання, нормування, порівнюваність балів та урівнювання.

Публікація тесту, технічні посібники та рекомендації користувачам.

Частина II. Використання тесту.

Загальні принципи використання тесту.

Тестування осіб з обмеженими можливостями.

Тестування та дві офіційні мови Канади.

Освітнє тестування та психологічне тестування у школах.

Використання тесту для клінічного оцінювання та консультативної допомоги (counseling).

Тестування з метою працевлаштування.

Професійне та виробниче ліцензування та сертифікація.

Оцінювання програм.

Частина III. Адміністративні процедури.

Адміністрування тесту, розрахунок бала та звітність.

Глосарій.

Кодекс справедливого тестування в освіті

Кодекс справедливого тестування в освіті [4] підготовлений Об'єднаним комітетом з практики тестування (JCTP).

Цей Комітет був утворений у 1985 році Американською асоціацією освітніх досліджень (AERA), Американською психологічною

асоціацією (APA) та Національною радою з вимірювань в освіті (NCME). Американська асоціація з консультування³⁷ (ACA) та Американська асоціація мовлення — мова — слух³⁸ (ASHA) приєдналися до JCTP у 1987 р. Національна асоціація шкільних психологів³⁹ (NASP) і Національна асоціація тестових директорів⁴⁰ (NATD) пізніше приєдналися до JCTP, яка тепер складається з ACA, AERA, APA, ASHA, NASP, NATD та NCME. Кожна асоціація може призначати двох представників до JCTP, один з яких бажано має відношення до розробки тестів та їх публікацій. За персонал JCTP відповідає Директорат з питань науки APA. Щорічно відбувається два чергових засідання JCTP, одне — під час засідання APA, інше — під час засідання однієї з членських асоціацій.

Згідно зі статутом JCTP надає «можливість, завдяки якій професійні організації та видавці тестів можуть співпрацювати для поліпшення використання тестів в оцінюванні».

Одна робоча група розробила та переглянула на відповідність Стандартам освітнього та психологічного тестування від 1999 р. **Кодекс справедливого тестування в освіті**. Переглянутий Кодекс схвалено спонсорськими асоціаціями, а також багатьма іншими професійними організаціями та видавцями тестів.

Друга робоча група, яка займалася проблемами користувачів тестів, розробила два документи: **Кваліфікації користувача тестів: просування належного використання тестів на основі даних** (A Data-Based Approach to Promoting Good Test Use) та **Відповідальне використання тестів: ситуативні завдання для оцінювання поведінки людини** (Responsible Test Use: Case Studies for Assessing Human Behavior). Перший описує кваліфікації користувача тестів на основі наукової методології та містить практичні поради щодо визначення кваліфікованих користувачів тестів, а також для класифікації тестів. Другий — 244-сторінковий том у м'якій обкладинці, окрім іншого можливого використання, дає додатковий матеріал викладачу з курсу тестування.

³⁷ Американська асоціація з консультування (*American Counseling Association*) — <http://www.counseling.org>.

³⁸ Американська асоціація мовлення — мова — слух (*American Speech-Language-Hearing Association*) — <http://www.asha.org>.

³⁹ Національна асоціація шкільних психологів (*National Association of School Psychologists*) — <http://www.nasponline.org>.

⁴⁰ Національна асоціація тестових директорів (*The National Association of Test Directors*) — <http://www.natd.org>.

Третя робоча група, що називається «Розуміння тестування» виготовила відеозапис під назвою «Ази шкільного тестування» та супровідний посібник «Рекомендації керівнику». Ці матеріали формують живе та інформативне уявлення про запитання, які найчастіше виникають у батьків щодо шкільного тестування. Відео та посібник можна придбати в NCME за \$20. Придбані примірники можуть копіюватися для некомерційного використання (за умови, що відеозапис буде копіюватися повністю і визнаватиметься авторство JCTP).

Четверта робоча група розробила **Права та обов'язки екзаменованих: рекомендації та вимоги**. У цьому документі пояснюються права та обов'язки екзаменованих під час процесу тестування, а також загальні вимоги до них з боку розробників, адміністраторів та користувачів тестів.

Кодекс справедливого тестування в освіті має бути узгоджений з відповідними частинами Стандартів освітнього та психологічного тестування (Американська асоціація освітніх досліджень [AERA], Американська психологічна Асоціація [APA] та Національна рада з вимірювань в Освіті [NCME], 1999).

Кодекс не має на меті додавати нові принципи до тих, що містяться в Стандартах APA, чи змінювати їх суть. Кодекс радше має представляти основний зміст вибраних частин Стандартів для розробників та користувачів тестів, а також для екзаменованих та/або їх батьків чи опікунів. Штати, округи, школи, організації чи окремі фахівці закликаються до дотримання справедливості у тестуванні та забезпеченні прав екзаменованих. Кодекс призначений надавати допомогу у виконанні таких зобов'язань.

Кодекс розроблено Об'єднаним комітетом з практик тестування, мета діяльності якого — сприяти забезпеченню інтересів громадськості для поліпшення якості практик тестування. Кодекс справедливого тестування в освіті — це рекомендації для фахівців у виконанні їхніх обов'язків, які полягають у наданні та використанні тестів, що розповсюджуються на всіх екзаменованих незалежно від їх віку, статі, фізичного стану (disability), раси, етнічного походження, громадянства, релігії, сексуальної орієнтації, мовного середовища чи інших особистісних характеристик. Справедливість є першочерговим міркуванням у всіх аспектах тестування. Ретельна стандартизація тестів та умов їх проведення допомагає забезпечити надання всім екзаменованим рівних можливостей у демонстрації своїх знань та вмінь у домені, що тестується. Справедливість означає, що кожен екзаменований має можливість підготуватися до тесту та отримати інформацію про

загальний характер та зміст тесту відповідно до мети тестування. Справедливість також передбачає точне звітування про індивідуальні та групові результати тестування. Справедливість не є ізольованою концепцією, а повинна враховуватись у всіх аспектах тестування.

Кодекс застосовується до тестування в освіті у широкому значенні (вступ, освітнє оцінювання, освітня діагностика та розподіл студентів) незалежно від способу представлення. Тому він стосується традиційних письмових тестів (conventional paper-and-pencil tests), комп'ютерних тестів та тестування практичних навичок (performance tests). Він не розрахований на тестування з метою працевлаштування, ліцензування, сертифікації чи інші типи тестування поза освітою. Кодекс насамперед спрямований на професійно розроблені тести, що використовуються у формально адміністрованих тестових програмах. Хоча Кодекс не охоплює тести, розроблені викладачами для використання їх на заняттях, викладачам рекомендовано використовувати його для поліпшення їхньої практики тестування.

Кодекс розглядає окремо ролі розробників тестів та користувачів тестів. Розробники тестів — це особи або організації, що конструюють тести або визначають політику тестових програм. Користувачі тестів — це особи або установи, які відбирають тести, адмініструють/проводять тестування, замовляють послуги з розробки тестів чи приймають рішення на основі тестових балів. Ролі розробника та користувача тесту можуть співпадати, наприклад, коли освітня установа рівня штату чи місцевого рівня замовляє послуги розробника тесту, визначає політику щодо процесу розробки тесту та приймає рішення на основі тестових балів.

Багато положень Кодексу стосуються відбору та використання існуючих тестів. Коли розробляється новий тест, модифікується існуючий тест чи спосіб адміністрування тесту, Кодекс надає рекомендації щодо цього процесу. Кодекс не є обов'язковим, вичерпним чи остаточним і може не поширюватися на всі ситуації. Натомість Кодекс визначає загальну мету і не має переваги над рішеннями тих, хто компетентний у розв'язуванні відповідних питань.

Кодекс окремо дає рекомендації розробникам та користувачам тестів у чотирьох критичних сферах:

- A. Розробка та відбір відповідних тестів.
- B. Адміністрування тестів та розрахунок тестових балів.
- C. Звітність та інтерпретація результатів.
- D. Інформування екзаменованих.

А. Розробка та відбір відповідних тестів

РОЗРОБНИКИ ТЕСТІВ	КОРИСТУВАЧІ ТЕСТІВ
1	2
Розробники тестів повинні надавати інформацію та допоміжні докази, які потрібні користувачам тестів для відбору відповідних тестів	Користувачі тестів повинні відбирати тести, які відповідають обраній меті та доцільні для передбачених екзаменованих
А-1. Надати докази того, що вимірює тест, рекомендоване використання, передбачені екзаменовані, переваги та обмеження тесту, включаючи рівень точності (precision) тестових балів	А-1. Визначити мету тестування, зміст та навички, що перевірятимуться, а також осіб, які братимуть участь у тестуванні. Вибирати та використовувати найдоцільніший тест, виходячи з ретельного аналізу наявної інформації
А-2. Описати, як відбиралися зміст та навички, що будуть тестуватися, і як розроблялися тести	А-2. Переглянути та відібрати тести відповідно до змісту, навичок, що перевірятимуться, та перевірити відповідність змісту обраній меті тестування
А-3. Повідомити інформацію про характеристики тесту на рівні деталізації, доцільному для передбачуваних користувачів тесту	А-3. Переглянути матеріали, надані розробниками тестів, та відібрати тести, щодо яких надано чітку, точну та повну інформацію
А-4. Надати рекомендації щодо рівнів навичок, знань та тренінгу, необхідних для відповідного перегляду, відбору та адміністрування тестів	А-4. Відібрати тести, залучаючи осіб з відповідними знаннями, навичками та тренінгом
А-5. Надати докази того, що технічна якість тесту, включаючи надійність і валідність, відповідає обраній меті	А-5. Оцінити докази технічної якості тесту, надані розробником тесту та іншими незалежними оглядачами
А-6. Надати кваліфікованим користувачам тесту типові зразки тестових запитань чи претестів, указівок, відповідей (answer sheets), посібників та звітів	А-6. Оцінити типові зразки тестових запитань чи претестів, указівок, відповідей (answer sheets), посібників та звітів перед тим, як відібрати тест

1	2
А-7. Під час розробки тестових запитань та супровідних матеріалів уникати потенційно образливих змісту та висловлювань	А-7. Оцінити процедури та матеріали, що використовуються розробниками тесту, та сам тест і переконатися у відсутності потенційно ображаючих висловлювань як за змістом, так і за формою
А-8. Розробити форми тестів чи процедур їх адміністрування, модифіковані належним чином для екзаменованих з обмеженими можливостями, які потребують спеціальних умов	А-8. Відібрати форми тестів чи процедури їх адміністрування, модифіковані відповідно до вимог екзаменованих з обмеженими можливостями, які потребують спеціальних умов
А-9. Отримати та надати докази успішності екзаменованих з різних підгруп, докладаючи активних зусиль до отримання достатніх зразків, що дозволяють проводити аналіз таких підгруп. Оцінити докази, щоб переконатися, що коливання в успішності пов'язані з оцінюваними навичками	А-9. Оцінити наявні докази успішності екзаменованих із різних підгруп. Визначити у межах можливого, які коливання в успішності могли виникнути через вплив факторів, не пов'язаних із оцінюваними навичками

В. Адміністрування тестів та розрахунок тестових балів

РОЗРОБНИКИ ТЕСТІВ	КОРИСТУВАЧІ ТЕСТІВ
1	2
Розробники тестів повинні пояснити, як правильно та справедливо адмініструвати тести та розраховувати тестові бали	Користувачі тестів повинні правильно та справедливо адмініструвати тести та розраховувати тестові бали
В-1. Дати чіткий опис детальних процедур адміністрування тестів стандартизованим способом	В-1. Дотримуватися встановлених процедур адміністрування тестів стандартизованим способом
В-2. Дати рекомендації щодо раціональних (reasonable) процедур для оцінювання осіб з обмеженими можливостями, які потребують особливих умов, чи для осіб з інших мовних середовищ	В-2. Дати та задокументувати відповідні процедури для оцінювання осіб з обмеженими можливостями, які потребують особливих умов, чи для осіб з інших мовних середовищ. Деякі з умов можуть вимагатися законом чи нормативними актами

ДОДАТКИ

1	2
В-3. Дати інформацію екзаменованим чи користувачам тестів про формати тестових запитань та процедури надання відповідей, включаючи інформацію про використання будь-яких необхідних матеріалів та обладнання	В-3. Дати екзаменованим можливість ознайомитися з форматами тестових запитань та будь-якими необхідними матеріалами та обладнанням, яке може використовуватися під час тестування
В-4. Установити та впровадити процедури для забезпечення схоронності тестових матеріалів на всіх етапах розробки тесту, його адміністрування, розрахунку тестових балів та звітності	В-4. Забезпечувати надійне зберігання тестових матеріалів, враховуючи дотримання авторських прав та унеможливлення отримання екзаменованими тестового результату шахрайським способом
В-5. Надати процедури, матеріали та рекомендації для розрахунку тестових балів і для моніторингу точності процесу розрахунку. Якщо за розрахунок тестових балів несе відповідальність розробник тесту, забезпечити відповідний тренінг для осіб, що будуть розраховувати тестові бали	В-5. Якщо за підрахунок тестових балів відповідає користувач тесту, провести відповідний тренінг для осіб, що будуть підраховувати тестові бали, забезпечити та перевірити точність процесу підрахунку
В-6. Виправити помилки, які впливають на інтерпретацію тестових балів, та негайно повідомити виправлені результати	В-6. Виправити помилки, які впливають на результат підрахунку тестових балів, та негайно повідомити про виправлення
В-7. Розробити та впровадити процедури для забезпечення конфіденційності тестових балів	В-7. Розробити та впровадити процедури для забезпечення конфіденційності тестових балів

С. Звітність та інтерпретація результатів

РОЗРОБНИКИ ТЕСТІВ	КОРИСТУВАЧІ ТЕСТІВ
1	2
Розробники тестів мають повідомляти точні результати тестування та надавати користувачам тестів інформацію, яка допомагатиме їм правильно інтерпретувати результати тестування	Користувачі тестів повинні повідомляти результати тестування точно і чітко

1	2
С-1. Надавати інформацію для обґрунтування рекомендованих інтерпретацій результатів, включаючи характер змісту, норми або групи порівняння/контрольні групи, та інші технічні докази. Повідомляти користувачів тестів про переваги та обмеження щодо результатів тестування та їх інтерпретації. Попереджати, що не слід розраховувати на більшу точність від гарантованої	С-1. Інтерпретувати значення результатів тестування, беручи до уваги характер змісту, норми або групи порівняння/контрольні групи, інші технічні докази, а також переваги та обмеження щодо результатів тестування
С-2. Давати рекомендації щодо інтерпретації результатів для тестів, що проводилися з модифікаціями. Поінформувати користувачів тестів про потенційні проблеми в інтерпретації результатів тестування, якщо тести або процедури адміністрування тестів були модифіковані	С-2. Інтерпретувати результати тестування модифікованих тестів чи тестів із модифікованими процедурами адміністрування, беручи до уваги можливий вплив цих модифікацій на результати тестування
С-3. Уточняти доцільне використання результатів тестування та застерегти користувачів тесту від можливих некоректних використання	С-3. Уникати використання тестів з іншою метою, ніж це рекомендовано розробником тесту, окрім випадків, коли існують докази на підтвердження доцільності використання для обраної мети чи інтерпретації
С-4. Коли розробники тестів вводять стандарти, необхідно окреслити логічне обґрунтування, процедури та докази для встановлення стандартів успішності чи прохідних балів. Уникати використання кліше	С-4. Переглянути процедури для встановлення стандартів успішності чи прохідних балів. Уникати використання кліше
С-5. Заохочувати користувачів тестів до прийняття рішень щодо екзаменованих на основі різних джерел відповідної інформації, а не тільки на основі тестового бала	С-5. Уникати використання тестового бала як єдиної детермінанти для прийняття рішень щодо екзаменованих. Інтерпретувати тестові бали у поєднанні з іншою інформацією щодо цих осіб

ДОДАТКИ

1	2
С-6. Надати інформацію, яка дозволить користувачам тестів точно інтерпретувати та звітувати про результати тестування щодо груп екзаменованих, включаючи інформацію про те, хто був включений до різних порівнюваних груп, а хто ні, та інформацію про фактори, які могли вплинути на інтерпретацію результатів	С-6. Визначити передбачену інтерпретацію та використання результатів тестування щодо груп екзаменованих. Уникати групування результатів тестування для цілей, не рекомендованих розробником тесту, окрім тих випадків, коли отримані докази, що підтверджують можливість їхнього використання для певної мети. Доповісти про процедури, які використовувалися для визначення осіб, включених до різних порівнюваних груп, а також тих, хто не був включений, та інформацію про фактори, які могли вплинути на інтерпретацію результатів
С-7. Вчасно надавати результати тестування у формі, зрозумілій екзаменованому	С-7. Вчасно повідомити результати тестування у формі, зрозумілій екзаменованому
С-8. Надавати рекомендації користувачам тестів про те, як відслідкувати ступінь, в якому тест задовольняє обраній меті	С-8. Розробити та впровадити процедури для контролю за використанням тесту, ураховуючи відповідність обраній меті тестування

D. Інформування екзаменованих

За певних обставин розробники тестів мають прямий контакт з екзаменованими та/або впливають на тести, процес тестування та результати тестування. За інших обставин ці зобов'язання покладаються на користувачів тестів.

Розробники тестів чи користувачі тестів повинні поінформувати екзаменованих про характер тесту, права та обов'язки екзаменованих, належне використання тестових балів та процедури апеляції щодо тестових балів

D-1. Поінформувати екзаменованих заздалегідь, до проведення тесту, про те, що охоплює тест, типи форматів запитань, указівки та відповідні стратегії екзаменування. Надати цю інформацію всім екзаменованим

D-2. Якщо тест не обов'язковий, надати екзаменованим чи їх батькам/опікунам інформацію, яка допоможе вирішити, чи слід скласти цей тест, включаючи попередження про будь-які наслідки нескладання тесту (наприклад, недопущення до конкурсу на певний освітній курс (scholarship)) і про наявність існуючих альтернатив цьому тесту
D-3. Надати екзаменованим чи їх батькам/опікунам інформацію про право екзаменованих отримати екземпляри тестів та заповнених бланків відповідей, перескладання тесту, перерахунок тестових балів чи ануляцію тестових балів
D-4. Надати екзаменованим чи їх батькам/опікунам інформацію про обов'язки екзаменованих, таких як: усвідомлення передбачуваної мети та правил використання тесту, докладання максимальних зусиль, слідування вказівкам та щодо нерозголошення тестових завдань чи заважання іншим екзаменованим
D-5. Інформувати екзаменованих чи їх батьків/опікунів по те, як довго тестові бали будуть зберігатися і кому та за яких обставин і в якій формі тестові бали та інша пов'язана інформація буде чи не буде розголошена. Охороняти тестові бали від неавторизованого розголошення та доступу
D-6. Описати процедури для розслідування та розв'язання ситуацій, які можуть призвести до ануляції чи відміни результатів, таких як недотримання вказаних процедур тестування
D-7. Описати процедури, за допомогою яких екзаменовані, батьки/опікуни та інші зацікавлені сторони можуть отримати більше інформації про тест, подати скарги та вирішити проблеми

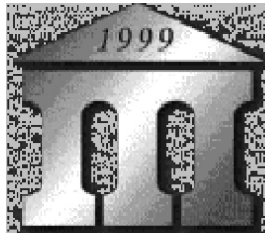
Рекомендації Міжнародної комісії з тестування

Далі наведено **офіційний** переклад Міжнародних Рекомендацій ІТС з використання тестів, що здійснено Центром Тестування за дозволом Президента ІТС prof. Jacques Grégoire. Оригінальний текст цих документів та копії дозволів на переклад представлено у додатку С.

Міжнародні Рекомендації ІТС з використання тестів



INTERNATIONAL TEST COMMISSION



ЦЕНТР ТЕСТУВАННЯ ПРИ МОЗ УКРАЇНИ

**Рекомендації Міжнародної тестової комісії (ІТС)
з використання тестів**

Українська мова

Версія 2000

© Офіційні права на переклад належать
Центру Тестування при МОЗ України

Рекомендації з використання тестів (Версія 2000) було офіційно прийнято на засіданні Ради ІТС у червні 1999 р. в м. Грац (Австрія).

Робоча Група EFPPA з тестів та тестування затвердила Рекомендації на засіданні 7 липня 1999 р. в Римі.

Офіційно рекомендації було представлено на Генеральних Зборах ІТС 24 липня 2000 р. в Стокгольмі. Усі права на Рекомендації належать ІТС, що є неакціонерною компанією, зареєстрованою у США.

Для отримання подальшої інформації про МТК чи дозволу цитувати або відтворювати зміст цього документа звертатися за адресою:

Prof Jacques Grégoire, ITC Secretary,
Université catholique de Louvain,
Faculté de Psychologie,
Place du Cardinal Mercier, 10,
1348 Louvain-la-Neuve,
Belgium.
Email: gregoire@deva.ucl.ac.be

Для отримання подальшої інформації про Рекомендації звертатися за адресою:

Prof Dave Bartram, ITC President,
SHL Group plc, The Pavilion, 1 Atwell Place,
Thames Ditton, KT7 0NE, Surrey, England.
Email: dave.bartram@shlgroup.com

Копію Рекомендацій можна отримати на web-сайті
ІТС — <http://www.intestcom.org>.

Зміст документа

Вступ та передісторія

Потреба у Міжнародних Рекомендаціях

Цілі

Розробка Рекомендацій

Рекомендації

Ключове призначення

Сфера застосування

Для кого призначені Рекомендації

Контекстуальні фактори

Знання, розуміння та навички

1. Відповідальність за етичне використання тестів

1.1. Діяти професійно та етично

1.2. Забезпечувати компетентність у використанні тестів

1.3. Брати відповідальність за використання тестів

1.4. Забезпечувати надійність зберігання тестових матеріалів

1.5. Забезпечувати конфіденційність обробки результатів тестування

2. Керуватися належною практикою (good practice) у використанні тестів

2.1. Оцінювати потенційну користь від тестування в ситуаціях, де здійснюється оцінювання (assessment situation)

2.2. Вибирати технічно якісний тест, відповідає ситуації

2.3. Приділяти достатньо уваги проблемам справедливості в тестуванні

2.4. Робити необхідні приготування до проведення тестування

2.5. Проводити тестування належним чином

2.6. Точно визначати бали та аналізувати результати тестування

2.7. Відповідно інтерпретувати результати тестування

2.8. Чітко і точно повідомляти результати тестування відповідним сторонам

2.9. Переглядати доцільність тесту та його використання

Посилання

Додаток А: Рекомендації щодо основ політики тестування

Додаток В: Рекомендації щодо розробки контрактів між сторонами, задіяними у процесі тестування

Додаток С: Питання, які слід брати до уваги під час підготовки до тестування осіб з обмеженими можливостями

Додаток D: Умови, якими слід керуватися при перекладі Рекомендацій з використання Тестів МТК

Подяка

Для Ради ІТС Рекомендації підготував проф. Дейв Бартрам (Dave Bartram). Занадану допомогу в реалізації проекту автор висловлює подяку Йєну Койну (Iain Coyne), а також учасникам Дублінського семінару в 1997 р., які зробили цінний внесок у розробку цих Рекомендацій.

Ms Dusica Boben, Produktivnost, СЛОВЕНІЯ;
Mr Eugene Burke, British Psychological Society, АНГЛІЯ;
Dr Wayne Camara, The College Board, США;
Mr Jean-Louis Chabot, ANOP, ФРАНЦІЯ;
Mr Iain Coyne, University of Hull, АНГЛІЯ;
Dr Riet Dekker, Swets and Zeitlinger, НІДЕРЛАНДИ;
Dr Lorraine Eyde, US Office of Personnel Management, США;
Prof Rocio Fernandez-Ballesteros, ЕРА, ІСПАНІЯ;
Mr Ian Florance, NFER-NELSON, АНГЛІЯ;
Prof Cheryl Foxcroft, Test Commission of South Africa, ЮАР;
Dr John Fremer, The College Board, США;
Ms Kathia Glabeke, Commissie Psychodiagnostiek, БЕЛЬГІЯ;
Prof Ron Hambleton, University of Massachusetts at Amherst, США;
Dr Karin Havenga, Test Commission of South Africa, ЮАР;
Dr Jurgen Hogrefe, Hogrefe & Huber Verlagsgruppe, НІМЕЧЧИНА;
Mr Ralf Horn, Swets and Zeitlinger, НІМЕЧЧИНА;
Mr Leif Ter Laak, Saville and Holdsworth Ltd, АНГЛІЯ;
Dr Pat Lindley, British Psychological Society, АНГЛІЯ;
Mr Reginald Lombard, Test Commission of South Africa, ЮАР;
Prof Jose Muniz, Spanish Psychological Association, ІСПАНІЯ;
Ms Gill Nyfield, Saville & Holdsworth Ltd, АНГЛІЯ;
Dr Torleiv Odland, Norsk Psykologforening, НОРВЕГІЯ;
Ms Berit Sander, Danish Psychologists' Association, ДАНІЯ;
Prof Francois Stoll, Federation Suisse des Psychologues, ШВЕЙЦАРІЯ.

Автор також висловлює подяку багатьом іншим особам та організаціям за надану допомогу на різних етапах консультацій та під час презентацій на конференціях.

Вступ та передісторія

Потреба у Міжнародних Рекомендаціях

Мета проекту Міжнародної комісії з тестування (ІТС) щодо розробки Рекомендацій — забезпечити належне використання тестів та впровадження якісних практик в оцінюванні. Робота, проведена ІТС на цей момент для поширення належної практики адаптації тестів [29, 30], є важливим кроком уперед у забезпеченні стандартизації якості тестів, адаптованих для використання в різних культурних та мовних середовищах. На своєму засіданні в Афінах у 1995 р. Рада ІТС прийняла пропозицію посилити вимоги до якості тестів шляхом урахування Рекомендацій зі справедливого та етичного їх використання, на основі яких можна розробити стандарти для підготовки та визначення компетентності користувачів тестів.

Потреба у створенні міжнародних Рекомендацій викликана низкою причин.

Ступінь офіційного контролю над застосуванням тестів та їх наслідків для екзаменованих у різних країнах неоднаковий (або відсутній зовсім). Не всі національні професійні товариства офіційно реєструють психологів; не всі мають механізми контролю над стандартами використання тестів неpsychологами. Застосування Рекомендацій, що здобули міжнародне визнання, допоможе

- національним асоціаціям психологів та іншим відповідним професійним установам і організаціям у розробці стандартів у країнах, де такі стандарти існують не в повному обсязі або взагалі не існують.

Права на придбання чи використання тестових матеріалів значно відрізняються

- у різних країнах. У деяких країнах доступ може надаватися виключно психологам, в інших — користувачам, зареєстрованим в офіційних національних тестових дистриб'юторів, у третіх екзаменовані можуть вільно, без обмежень отримати матеріали від місцевих або безпосередньо від закордонних постачальників.
- низка відомих інструментів з'явився в Інтернеті з порушенням авторських прав, без посилання на авторів чи видавців тестів та без належного забезпечення захисту тестової інформації (test security).
- зростаюча міжнародна мобільність робочої сили вимагає застосування тестів у сфері професійного тестування (occupational testing) для потенційних кандидатів з різних країн. Часто тестування проводиться в певній країні від імені потенційного роботодавця з іншої країни.

- у США та Великобританії проводиться розробка методик щодо використання Інтернету для дистанційного (віддаленого) оцінювання як у професійній, так і в освітній сферах. Це викликає масу запитань стосовно стандартів проведення та контролю процесу тестування, зокрема щодо забезпечення захисту тестової інформації.

Цілі

Довгостроковою метою цього проекту є розробка Рекомендацій з визначення компетентності (знання, навички, уміння та інші особисті характеристики) користувачів тестів. Ці компетентності визначені у вигляді вимірюваних критеріїв якості функціонування (*assessable performance criteria*). Ці критерії складають основу для розробки специфікації ознак компетентності особи, що претендує на статус користувача тестів. Компетентність має охоплювати такі питання:

- професійні та етичні стандарти у тестуванні;
- права екзаменованого та інших учасників процесу тестування;
- вибір та оцінювання альтернативних тестів;
- адміністрування тестів, розрахунок відповідного бала (*scoring*) та інтерпретація результатів;
- написання звітів та зворотний зв'язок.

У тій мірі, в якій Рекомендації прямо стосуються використання тестів, вони також впливають на:

- стандарти побудови тестів;
- стандарти розробки документації для користувачів (наприклад, технічні та прикладні посібники);
- стандарти правил надання та розповсюдження тестів, а також інформацію про них.

Ці Рекомендації представляють роботу спеціалістів у галузі психологічного та освітнього тестування (психологів, психометристів, розробників та видавців тестів) з багатьох країн. Цей документ не є «винаходом» нових рекомендацій, а спробою об'єднати загальні тенденції, що містяться у вже існуючих рекомендаціях, кодексах практик, стандартах та інших відповідних документах, і створити когерентну структуру для їх розуміння та використання.

Розробка Рекомендацій

Рекомендації слід розглядати як певні еталони (*benchmarks*), відносно яких можна порівнювати діючі місцеві стандарти щодо повноти охоплен-

ня та відповідності міжнародним нормам. Використовуючи Рекомендації як еталон чи основу для розробки документів місцевого застосування (наприклад, стандартів, кодексів практик, декларації прав екзаменованих), можна досягти високого рівня їх узгодженості в різних країнах.

Робота над Рекомендаціями розпочалася зі збору матеріалів з питань стандартів тестування, кодексів практик та застосування тестів у ряді країн⁴¹ тощо. Крім використання цих джерел, значний вплив на ці Рекомендації справили такі документи:

- *Додаток до Рекомендацій із використання психологічних тестів* [31]. Австралійське Психологічне Товариство (APS), 1997;
- *Стандарти рівня А та рівня Б для виробничого тестування* [32, 33]. Британська психологічна спілка (BPS);
- *Стандарти освітнього та психологічного тестування* [1]. Американська асоціація освітніх досліджень (AERA), Американська психологічна асоціація (APA) та Національна рада з вимірювань в освіті (NCME), 1985;
- *Відповідальність користувача стандартизованих тестів* [34]. Американська Асоціація психологічної підтримки та розвитку (AACD), 1992;
- *Рекомендації з освітнього та психологічного тестування* [3]. Канадська психологічна асоціація (CPA), 1987.

Публікація APS була особливо цінною, оскільки вона значною мірою узагальнює зміст британських та американських праць, публікацій Південноафриканського національного інституту психологічних досліджень (NIPR) та інструкцій з використання тестів різних видавців. Вона також містить основні положення базового документа про заохочення належного використання тестів на основі баз даних (data-based approach to promoting good test use) Робочої групи з визначення кваліфікацій користувачів тестів (TUQWG) Об'єднаного комітету з практики тестування (JCTP) [35, 36, 37] та публікації JCTP щодо *Кодексу справедливого тестування в освіті* [38]. Додаток Б базується на роботі JCTP 2000 р. про права та обов'язки екзаменованих.

На основі аналізу змісту першоджерел основні положення (Statements) було структуровано за 14 головними розділами. У разі необхідності спільний зміст положень з різних джерел об'єднувався в одному, спеціально сформульованому положенні. Формат положень також був

⁴¹ Перелік усіх матеріалів, на яких ґрунтувався цей процес, можна отримати від авторів.

уніфікований у вигляді завершення спільного стверджувального речення (наприклад: «Компетентні користувачі тестів намагатимуться...» чи «Компетентні користувачі тестів можуть...»).

Цю початкову структуру з 14 основних розділів та її зміст було втілено у першому проекті Рамкового документа, що слугував матеріалом для міжнародного семінару в Дубліні в липні 1997 р. Мета семінару — розгляд і критична оцінка всіх аспектів Рамкового документа для подальшої розробки проекту Рекомендацій, які мали б міжнародне застосування та визнання. На семінарі відбувся детальний розгляд Рамкового документа та було прийнято поправки щодо його форми, структури та змісту. Після семінару документ суттєво доопрацьовано (Версія 2.0) і розповсюджено серед учасників з метою отримання їх коментарів. Проект консультативного документа (Версія 3.1) поєднав усі коментарі та пропозиції до Версії 2.0.

Примірники Версії 3.1 та структурована форма для відповіді були надіслані відповідним особам та організаціям для коментарів. Загалом було розповсюджено 200 примірників і отримано 28 детальних відповідей, включаючи «корпоративні» від APA, BPS та деяких інших європейських професійних асоціацій. Улітку 1998 р. Рекомендації було переглянуто зважаючи на ці коментарі, унаслідок чого 200 примірників (Версія 4.1) розіслано для подальших консультацій. На другому колі консультацій отримано 18 офіційних відповідей. Крім цього, багатьма одержувачами надано неформальні коментарі щодо розглядуваних документів електронною поштою чи під час зустрічей.

Під час розробки цієї версії Рекомендацій (Версія 2000), усі зусилля було зосереджено на урахуванні цих відгуків, досить корисних та конструктивних⁴².

Запропонована структура вирізняє три основних аспекти компетентності:

1. Професійні та етичні стандарти належної практики, які впливають на проведення процесу тестування та спосіб взаємодії користувачів тестів з іншими учасниками процесу.
2. Знання, розуміння та вміння, що стосуються процесу тестування (що саме користувачі тестів повинні вміти робити).
3. Знання та розуміння, необхідні для інформування та підтримки процесу тестування.

⁴² Детальний звіт про результати перших консультацій було подано до Ради ІТС на засіданні в серпні 1998 р. Звіт про другі консультації разом з Версією 5.0 Рекомендацій було подано до Ради ІТС на засіданні в червні 1999 р. Версія 2000 містить деякі незначні редакторські правки до Версії 5.0.

Ці три компоненти відрізняються, проте на практиці вони взаємопов'язані.

Рекомендації починаються з *ключового призначення*. Його також можна назвати «характеристикою місії» (mission statement) користувача тестів. Це — відправний пункт розробки Рекомендацій. Кожна рекомендація окреслює певний аспект компетентності користувача тестів, з яких складається ключове призначення.

Сфера застосування разом із ключовим призначенням визначає аудиторію Рекомендацій, форми та контексти оцінювання, яких стосуються Рекомендації. Цей документ містить:

- ключове призначення та сферу застосування;
- опис компетентності користувачів тестів стосовно етичного використання тестів;
- опис компетентності користувачів тестів стосовно належної практики їх використання тестів.

Рекомендації

Ключове призначення

Компетентний користувач тестів мусить використовувати тести доцільно, професійно та етично, зважаючи на потреби та права учасників процесу тестування, причини тестування та всю ситуацію, в якій відбувається тестування.

Цей результат буде досягнуто за умови, якщо користувач тестів володітиме необхідною компетентністю для проведення тестування, знаннями та розумінням стосовно використання тестів та тестування, необхідними для інформування та підтримки процесу тестування.

Сфера застосування

Будь-яка спроба дати точне визначення «тесту» чи «тестування», як процесу, марна, оскільки певні необхідні процедури можуть бути опущені, а зайві — включені. Тому терміни «тест» та «тестування», що використовуються у Рекомендаціях, слід розуміти у широкому значенні. Неважливо, чи має процедура оцінювання назву «тест», чи ні. Ці Рекомендації підходять до багатьох процедур оцінювання, які не називаються «тест». Замість того, щоб давати єдине визначення, наступні

твердження намагатимуться окреслити домен (сферу), на який поширюються Рекомендації.

- Тестування включає широкий спектр процедур для використання в психологічному, професійному та освітньому оцінюванні.
- Тестування може включати процедури для вимірювання як нормальної, так і аномальної чи дисфункціональної поведінки.
- Процедури зазвичай розробляються для проведення тестування в ретельно контрольованих чи стандартизованих умовах. Вони представлені як відповідні систематичні кількісні протоколи (systematic scoring protocols).
- Ці процедури відображають процес вимірювання та дають змогу робити висновки на підставі одержаних результатів.
- Тестування також включає процедури якісної класифікації чи ранжування людей (наприклад, за типами).

Будь-яка процедура, що використовується для «тестування» у вищезазначеному сенсі, має вважатися «тестом», незважаючи на спосіб її проведення; незалежно від того, чи була вона розроблена професійним розробником тесту, незалежно від використання в ній набору запитань, завдань для виконання або операцій (наприклад, зразків робіт (work samples), або психомоторних тестів (psycho-motor tracking tests)).

Тести повинні мати підтвердження надійності та валідності для обраної мети. Надання доказів необхідне для обґрунтування висновків, зроблених на основі тестового бала. Ці докази мають бути доступними і вільно надаватися користувачеві тестів для незалежного розгляду та оцінки. Якщо важливі докази містяться у важкодоступних технічних звітах, дистриб'ютор тестів має надавати повністю підтверджені анотації (fully referenced synopses).

Представлені тут Рекомендації з використання тестів мають застосовуватися до всіх процедур такого рівня, незалежно від їх назви («психологічні тести» чи «освітні тести») та наявності доступних технічних доказів.

Багато з цих Рекомендацій стосуватимуться також процедур оцінювання, які лежать поза доменом «тестів». Вони можуть також застосовуватися до будь-яких процедур оцінювання в ситуаціях, де оцінювання має суттєве значення і у випадку неадекватного застосування може призвести до персональних збитків чи спричинити психологічні страждання (наприклад, співбесіда при відборі кандидатів на роботу, оцінка професійних досягнень, діагностичне оцінювання потреб у педагогічній підтримці).

Рекомендації не стосуються використання матеріалів, зовні подібних до тестів, проте всіма учасниками використовуються тільки з розважальною метою (наприклад, визначення стилю життя в журналах чи газетах).

Для кого призначені Рекомендації

Рекомендації стосуються використання тестів у професійній практиці. І тому вони спрямовані передусім на допомогу:

- покупцям та власникам тестових матеріалів;
- відповідальним за відбір тестів та визначення мети їх використання;
- тим, хто проводить тестування, визначає тестові бали чи інтерпретує результати тестування;
- консультантам, які використовують результати тестування (наприклад, рекрутери, спеціалісти з освітньої та професійної орієнтації, тренерам, фахівцям з планування спадкоємства у керівництві);
- особам, які повідомляють результати тестування та надають екзаменованим інформацію зі зворотного зв'язку.

Рекомендації можуть стосуватися й інших сторін, задіяних у застосуванні тестів згідно з викладеним вище, а тому також розраховані на:

- розробників тестів;
- постачальників тестів;
- тих, хто займається тренінгом користувачів тестів;
- осіб, які проходять тестування, та їх оточення (батьки, подружжя, партнери);
- професійні установи та інші асоціації, зацікавлені у використанні психологічного та освітнього тестування;
- політиків і законотворців.

Хоча передусім Рекомендації призначені для використання в професійній практиці, більшість аспектів належного використання (good practice), що втілені у них, можуть стосуватися й тих, хто використовує тестування в дослідницьких цілях.

Метою Рекомендацій не є охоплення кожного типу оцінювання (наприклад, неструктуровані чи напівструктуровані співбесіди, оцінювана групова діяльність) чи кожної ситуації, де має місце оцінювання (наприклад, оцінювання в центрах зайнятості). Проте вони можуть бути корисними в ситуаціях, де здійснюється оцінювання, у більш широкому сенсі, ніж це передбачено першочерговим призначенням психологічного та освітнього тестування (наприклад, використання оцінювання в центрах зайнятості для працевлаштування чи відбору кандидата, напівструктуровані та структуровані співбесіди, або оцінювання з метою відбору, кар'єрного спрямування та консультативної допомоги).

Контекстуальні фактори

Рекомендації набули міжнародного застосування. Вони можуть використовуватися для розробки особливих місцевих стандартів через процес контекстуалізації. Визнано, що існує багато факторів, які впливають на застосування та реалізацію стандартів на практиці. Ці контекстуальні фактори слід брати до уваги на місцевому рівні при інтерпретації Рекомендацій та визначенні їх реального практичного значення в певному середовищі.

При перетворенні Рекомендацій на певні конкретні стандарти необхідно враховувати такі фактори:

- соціальні, політичні, інституційні, мовні та культурні відмінності між середовищами, в яких проводиться оцінювання;
- закони країни, в якій відбувається тестування;
- існуючі національні рекомендації та стандарти, установлені професійними психологічними товариствами та асоціаціями;
- відмінності між індивідуальним та груповим оцінюванням;
- відмінності, пов'язані з умовами, в яких відбувається тестування (освітнє, клінічне, професійне та інше оцінювання);
- основні одержувачі результатів тестування (наприклад, екзаменовані, їх батьки чи опікуни, розробник тесту, роботодавець чи інша третя сторона);
- відмінності у використанні результатів тестування, наприклад, для прийняття рішень (як передвиборче опитування), для надання інформації, необхідної для орієнтаційної чи консультативної допомоги (guidance or counseling);
- різниці ступеня, з яким умови оцінювання дають змогу уточнювати та коригувати інтерпретацію результатів з урахуванням нової інформації.

Знання, розуміння та навички

Знання, розуміння та навички становлять базу всіх компетентностей користувача тестів. Характер змісту та рівень деталізації може бути різним у різних країнах та різних сферах застосування залежно від рівня компетентності, необхідного для використання тесту.

Рекомендації не містять детального опису знань, розуміння та навичок. Проте під час застосування Рекомендацій в конкретних ситуаціях необхідно уточнювати відповідні знання, розуміння та навички. Таке уточнення є частиною процесу контекстуалізації, за допомогою якого загальні Рекомендації розвиваються у конкретні стандарти. Основні об-

ласті, які повинні охоплюватися описом знань, умінь та навичок, включають:

Важливі декларативні знання (relevant declarative knowledge):

- знання базових психометричних принципів та процедур, а також технічних вимог до тестів (наприклад, надійність, валідність, стандартизація);
- знання тестів та принципів вимірювання, достатні для адекватного розуміння результатів тестування;
- знання та розуміння відповідних теорій та моделей здібностей (ability), особистості (personality) та психологічних конструктів (psychological constructs) або психопатології, згідно з необхідністю обґрунтувати вибір тестів та інтерпретувати результати тестування;
- знання тестів та їх постачальників у відповідній області діяльності.

Інструментальні знання та навички (instrumental knowledge and skills):

- знання та вміння стосовно конкретних процедур чи інструментів оцінювання, включаючи використання процедур комп'ютерного оцінювання;
- спеціалізовані знання та практичні навички в області використання тестів, які входять до набору інструментів оцінювання;
- знання та розуміння конструкту чи конструктів, що лежать в основі тестових балів (test scores), необхідних для отримання валідних висновків за результатами тестування.

Рекомендації охоплюють:

Загальні особисті практичні навички (general personal task-related skills):

- виконання відповідних дій, таких як проведення тесту, звітність, надання зворотного зв'язку екзаменованим та іншим клієнтам;
- навички усної та письмової комунікації, достатні для належної підготовки екзаменованих, проведення тестування, повідомлення результатів та взаємодії з іншими учасниками процесу;
- навички міжособистісного спілкування, достатні для належної підготовки екзаменованих, проведення тестування та надання зворотного зв'язку у вигляді результатів тестування.

Контекстуальні знання та навички (contextual knowledge and skills):

- знання щодо того, коли слід використовувати тести;

- знання щодо того, як інтегрувати тестування та інші, менш формальні компоненти оцінювання (наприклад, біографічні дані, неструктурована співбесіда та рекомендаційні посилання, тощо);
- знання з актуальних професійних, юридичних та етичних питань, що стосуються використання тестів та їх практичного застосування.

Навички практичного управління (task management skills):

- знання кодексу поведінки та належної практики, що стосуються використання тестів, даних тестування, надання зворотного зв'язку, виготовлення та зберігання звітності, зберігання та відповідальність за збереження тестових матеріалів та даних;
- знання соціального, культурного та політичного контексту, в якому використовується тестування, та способу, в який ці фактори можуть вплинути на результати, їх інтерпретацію та використання.

Навички кризового управління (contingency management skills):

- знання, як долати проблеми, труднощі та невдачі у повсякденній роботі;
- знання, як реагувати на запитання екзаменованого під час проведення тестування тощо;
- знання, як діяти в ситуації, де існує імовірність некоректного використання тестування (test misuse) чи неправильної інтерпретації результатів.

1. Відповідальність за етичне використання тестів

Компетентні користувачі тестів повинні:

1.1. Діяти професійно та етично

- 1.1.1. Пропагувати та підтримувати професійні та етичні стандарти.
- 1.1.2. Достатньою мірою розуміти (working understanding) актуальні професійні та етичні питання і дискусії щодо використання тестів у відповідній сфері застосування.
- 1.1.3. Упроваджувати прозору і детальну політику з тестування та використання тестів⁴³.

⁴³ Приклад базової політики наведено в Додатку А цих рекомендацій.

- 1.1.4. Забезпечувати дотримання відповідних професійних та етичних стандартів поведінки колегами та підлеглими.
- 1.1.5. Ураховувати під час спілкування можливу вразливість екзаменованого та інших сторін.
- 1.1.6. Представляти тести та тестування у повідомленнях та через ЗМІ у позитивний та збалансований спосіб.
- 1.1.7. Уникати ситуацій, в яких може скластися враження приватного зацікавлення у результатах оцінювання або відбуватись псування стосунків з клієнтом через оцінювання.

1.2. Забезпечувати компетентність у використанні тестів

- 1.2.1. Працювати в рамках наукового принципу та обґрунтованого досвіду.
- 1.2.2. Установлювати та підтримувати високі персональні стандарти компетентності.
- 1.2.3. Знати межі власної компетентності і діяти в цих межах.
- 1.2.4. Бути в курсі зміни нормативних досягнень у галузі тестування, включаючи зміни у законодавстві та політиці, що можуть впливати на тести та їх використання.

1.3. Брати відповідальність за використання тестів

- 1.3.1. Пропонувати лише ті послуги з тестування та використовувати тільки ті тести, які відповідають їх кваліфікації.
- 1.3.2. Брати на себе відповідальність за вибір тестів та надані рекомендації.
- 1.3.3. Забезпечувати чітке та адекватне інформування учасників тестового процесу про етичні принципи та законодавчі норми, якими керується психологічне тестування.
- 1.3.4. Забезпечувати чіткість та зрозумілість суті контракту між екзаменованим та екзаменатором⁴⁴.
- 1.3.5. Бути уважним до будь-яких непередбачуваних наслідків використання тесту.
- 1.3.6. Намагатися уникати ситуацій, що можуть погіршити психологічний стан учасників процесу тестування.

1.4. Забезпечувати надійність зберігання тестових матеріалів

- 1.4.1. Забезпечувати надійне зберігання та контрольований доступ до тестових матеріалів.

⁴⁴ Приклад базового «контракту» між користувачем тестів та екзаменованим наведено в Додатку В цих рекомендацій

- 1.4.2. Поважати авторські права та угоди стосовно тесту, включаючи будь-які заборони копіювання або передачі матеріалів в електронному чи іншому вигляді будь-яким особам, незалежно від їхньої кваліфікації.
- 1.4.3. Зберігати цілісність тесту (integrity of the test), не допускати тренування окремих осіб на справжніх тестових чи тренувальних матеріалах, якщо це може надати їм незаконну перевагу у виконанні тесту.
- 1.4.4. Переконаватися у тому, що доступний широкому загалу опис техніки тестування не впливає на їх дієвість.

1.5. Забезпечувати конфіденційність обробки результатів тестування

- 1.5.1. Визначати права на доступ до результатів та рівні конфіденційності.
- 1.5.2. Повідомляти про всі рівні конфіденційності до проведення тестування.
- 1.5.3. Обмежувати доступ до результатів тільки тим, хто має на це право.
- 1.5.4. Отримувати відповідні дозволи перед наданням результатів іншим особам.
- 1.5.5. Захищати дані, які зберігаються в системі, від несанкціонованого доступу до них.
- 1.5.6. Установлювати чіткі правила щодо терміну зберігання даних тестування в системі.
- 1.5.6. Вилучати імена та інші персональні відомості з масивів результатів, які архівуються, використовуються з дослідницькою метою, а також для розробки норм або іншої статистичної обробки.

2. Керуватися належною практикою (good practice) у використанні тестів

2.1. Оцінювати потенційну користь від тестування в ситуаціях, де здійснюється оцінювання (assessment situation)

Компетентні користувачі тестів повинні:

- 2.1.1. Надавати обґрунтування необхідності використання тестування.
- 2.1.2. Забезпечувати проведення ретельного аналізу потреб клієнта, причин для звернення, діагностичної категорії, умов або роботи для яких використовується оцінювання.

- 2.1.3. Установлювати, що виміряні тестом знання, навички, здібності, здатності та інші характеристики корелюють з відповідною поведінкою в цьому контексті, про що і будуть робитися висновки.
- 2.1.4. Шукати інші відповідні допоміжні джерела інформації.
- 2.1.5. Оцінювати переваги та недоліки використання тестів порівняно з іншими джерелами інформації.
- 2.1.6. Гарантувати, що всі відповідні наявні допоміжні джерела інформації повністю використовуються.

2.2. Вибирати технічно якісний тест, що відповідає ситуації

Компетентні користувачі тестів повинні:

- 2.2.1. Вивчати актуальну інформацію про спектр відповідних тестів (наприклад, обрати за зразками, незалежними оглядами, порадами експертів) до того, як вибрати тест для використання.
- 2.2.2. Визначати, що відповідна документація про технічні характеристики та документація для користувачів містить достатньо інформації, щоб оцінити:
 - а) діапазон або область охоплення та репрезентативність змісту тесту, правомірність груп для визначення норм, рівень складності змісту тощо;
 - б) точність вимірювання та надійність, продемонстровані на релевантних вибірках;
 - в) валідність (продемонстрована на релевантних вибірках) та доцільність для обраного використання;
 - г) вільність від систематичної похибки (systematic bias) стосовно обраної групи екзаменованих;
 - д) прийнятність для задіяних у використанні тестів, включаючи сприйняття їх як справедливих та доцільних;
 - е) практичність, включаючи необхідний час, вартість та відповідні ресурси.
- 2.2.3. Уникати використання тестів, які мають недостатню або нечітку технічну документацію.
- 2.2.4. Використовувати тести виключно за наявності достатніх та відповідних доказів валідності.
- 2.2.5. Уникати суджень про тест лише на основі зовнішніх характеристик, свідчень користувачів тесту чи порад з боку тих, хто має комерційну зацікавленість.
- 2.2.6. Відповідати на запити від зацікавлених сторін (наприклад екзаменованих, батьків, менеджерів), надаючи їм достатньо інформації для розуміння вибору саме цього тесту.

2.3. Приділяти достатньо уваги проблемам справедливості в тестуванні

Якщо тести мають використовуватися для осіб з різних груп (наприклад, груп, що відрізняються за статтю, культурним спадком, освітою, етнічним походженням чи віком), компетентні користувачі тестів мають докладати всіх можливих зусиль для забезпечення того, щоб:

- 2.3.1. Тести були неупередженими (unbiased) та доцільними для різних груп екзаменованих.
- 2.3.2. Конструкти (constructs), що оцінюються, мали сенс для кожної з груп екзаменованих.
- 2.3.3. Існували докази можливих розбіжностей в успішності проходження тесту для різних груп.
- 2.3.4. Існували, де можливо, докази здатності диференціювати тестові завдання (differential item functioning (DIF)).
- 2.3.5. Існували докази валідності для обґрунтування використання певного тесту для різних груп.
- 2.3.6. Зменшити вплив відмінностей між групами, що не має значення для тестування (наприклад, різна мотивація до відповіді, чи вміння читати).
- 2.3.7. В усіх випадках Рекомендації стосовно справедливого використання тестів інтерпретували в контексті місцевої політики та законодавства⁴⁵.

Якщо тестування відбувається більш, ніж однією мовою (в одній країні, чи в різних⁴⁶), компетентні користувачі тестів мають докладати всіх можливих зусиль для забезпечення того, щоб:

- 2.3.8. Кожна версія певною мовою чи діалектом була розроблена з використанням ретельної методології, що відповідає вимогам найліпшої практики (best practice).
- 2.3.9. Розробники тесту враховували питання змісту, культури та мови.
- 2.3.10. Особи, що проводять тестування (test administrators), могли вільно спілкуватися мовою, якою проводиться тест.

⁴⁵ Рекомендації в цьому розділі зосереджуються на «найліпших практиках». Проте в багатьох країнах, визначаючи справедливе використання тестів, слід брати до уваги національне законодавство (наприклад, the Americans with Disabilities Act, 1990, у США, або the Race Relations Act, 1976, у Великобританії).

⁴⁶ Ці Рекомендації стосуються не тільки різних національних мов та діалектів, а також особливих форм комунікації, таких як мова жестів, що використовуються для подолання наслідків обмежених можливостей.

- 2.3.11. Рівень володіння екзаменованою мовою, якою проводиться тестування, систематично визначався, підбиралася версія відповідною мовою, або в разі необхідності проводилося двомовне оцінювання.

Якщо тестування проводитиметься для осіб з обмеженими можливостями, компетентні користувачі тестів мають докладати всіх можливих зусиль для того, щоб забезпечити:

- 2.3.12. Отримання консультації від відповідних спеціалістів стосовно впливу обмежених можливостей на успішність складання тестового іспиту.
- 2.3.13. Проведення консультацій з потенційними екзаменованими, урахування їхніх потреб та побажань.
- 2.3.14. Уживання адекватних заходів, якщо серед екзаменованих є особи з порушеннями слуху, зору, моторики тощо (наприклад, труднощі у навчанні (learning impairments), дислексія).
- 2.3.15. Розгляд доцільності використання альтернативних процедур оцінювання, на противагу модифікаціям того самого тесту (наприклад, інші більш прийнятні завдання чи інакше структуровані форми оцінювання).
- 2.3.16. Одержання порад відповідних професіоналів, якщо ступінь модифікацій тесту для осіб з обмеженими можливостями перевищує досвід користувача тестів.
- 2.3.17. Відповідність характеру модифікацій (у разі необхідності) характеру інвалідності та меті мінімізації її впливу на валідність тестового бала.
- 2.3.18. Надання інформації стосовно будь-яких модифікацій до тесту чи процедури тестування тим, хто інтерпретує тестовий бал чи приймає рішення на його основі, якщо ненадання такої інформації призведе до упередженої інтерпретації чи несправедливого рішення.

2.4. Робити необхідні приготування до проведення тестування

Компетентні користувачі тестів мають докладати всіх можливих зусиль для того, щоб забезпечити:

- 2.4.1. Отримання відповідними сторонами вчасно чіткої інформації про мету тестування, найкращий спосіб підготовки до нього та процедуру тестування.
- 2.4.2. Інформування екзаменованих з мовних чи діалектних груп, для яких тест вважається доцільним.

- 2.4.3. Отримання екзаменованими погоджених зразків тренувальних або підготовчих матеріалів, якщо такі наявні і це не суперечить рекомендованій практиці для використання цих тестів.
- 2.4.4. Роз'яснення екзаменованим їх прав та обов'язків⁴⁷.
- 2.4.5. Отримання явної згоди екзаменованих, їхніх юридичних представників чи опікунів до проведення тестування.
- 2.4.6. Пояснення відповідним сторонам у разі необов'язкового тестування наслідків складання тестового іспиту для їх свідомого вибору.
- 2.4.7. Вживання необхідних заходів шляхом забезпечення:
 - а) проведення підготовки згідно з інструкцією видавця тестів;
 - б) підготовлення заздалегідь спеціального приміщення та обладнання для тестування, до яких був би вільний доступ, які були б безпечними, тихими, не відволікали та відповідали призначенню;
 - в) достатньої кількості перевірених матеріалів, які не містять позначок, залишених попередніми користувачами в тестових буклетах чи формах відповідей;
 - г) наявності компетентного персоналу для проведення тестування;
 - д) необхідних заходів для тестування осіб з обмеженими можливостями⁴⁸.
- 2.4.8. Передбачення всіх можливих проблем та способів їх вирішення шляхом підготовки матеріалів та інструкцій.

2.5. Проводити тестування належним чином

Компетентний користувач тестів повинен:

- 2.5.1. Установлювати контакт з екзаменованими під час привітання та інструктажу через позитивне ставлення до них.
- 2.5.2. Зменшувати хвилювання екзаменованого та уникати створення чи підсилення зайвого напруження.
- 2.5.3. Вимикати потенційні джерела відволікання (наприклад, звуки будильників у годинниках, мобільні телефони, пейджери).
- 2.5.4. Забезпечувати екзаменованих матеріалами до початку тестування.
- 2.5.5. Проводити тестування під відповідним наглядом.
- 2.5.6. По можливості проводити інструктаж рідною мовою екзаменованих, навіть якщо тест має на меті демонстрацію знань чи навичок нерідною мовою.

⁴⁷ Див. додаток В.

⁴⁸ Див. Додаток С (цих Рекомендацій).

- 2.5.7. Суворо дотримуватися правил, зазначених в інструкції до тесту, допомагаючи особам з обмеженими можливостями.
- 2.5.8. Читати інструкції чітко та спокійно.
- 2.5.9. Надавати достатньо часу для заповнення зразків.
- 2.5.10. Спостерігати та фіксувати відхилення від процедури тестування.
- 2.5.11. У разі необхідності вести облік та чітко фіксувати час відповіді.
- 2.5.12. Забезпечувати облік усіх матеріалів по завершенні тестування.
- 2.5.13. Проводити тестування на адекватному та відповідному рівні контролю з установами особистості екзаменованого.
- 2.5.14. Забезпечувати належну підготовку всіх осіб, які проводять тестування.
- 2.5.15. Переконаватися, що екзаменовані не залишаються без нагляду і не відволікаються під час контрольного тестування.
- 2.5.16. Надавати допомогу екзаменованим, які виявляють ознаки зайвого хвилювання чи тривоги.

2.6. Точно визначати бали та аналізувати результати тестування

Компетентні користувачі тестів повинні:

- 2.6.1. Чітко дотримуватися стандартизованої процедури розрахунку бала (scoring).
- 2.6.2. Виконувати доцільну конвертацію первинного «сирого» бала (raw score) в інші відповідні шкали.
- 2.6.3. Вибирати типи шкал відповідно до мети використання тестових балів.
- 2.6.4. Перевіряти конвертацію балів між шкалами та їх точність.
- 2.6.5. Переконаватися у валідності всіх висновків у результаті порівняння балів з нормами, що не стосуються групи екзаменованих або вже є застарілими.
- 2.6.6. Розраховувати в разі необхідності комбіновані бали (composite scores) за стандартними формулами та рівняннями.
- 2.6.7. Застосовувати процедури скринінгу результатів тестування для виявлення неможливих чи нереальних балів.
- 2.6.8. Чітко та точно вказувати назви шкал у звітних формах, чітко ідентифікувати використані норми, типи шкал та рівнянь.

2.7. Відповідно інтерпретувати результати тестування

Компетентні користувачі тестів повинні:

- 2.7.1. Професійно розуміти теоретичну та концептуальну базу тестування, технічну документацію та інструкції з використання та інтерпретації тестових балів за різними шкалами (scale scores).

- 2.7.2. Добре орієнтуватися у використанні шкал, характеристиках груп для визначення норм та їх порівняння (norm or comparison groups), та враховувати обмеженість балів (limitations of the score).
- 2.7.3. Вживати заходів для мінімізації впливу на інтерпретацію тестових результатів можливої упередженості осіб, які здійснюють інтерпретацію стосовно членів групи екзаменованих певного культурного походження.
- 2.7.4. Використовувати у разі можливості відповідні групи, норми та порівняння.
- 2.7.5. Інтерпретувати результати, ураховуючи наявну інформацію про екзаменованого (включаючи вік, стать, освіту, культурне походження та інші фактори), а також обмеження тесту, контексту оцінювання та потреб тих, хто має законні права на отримання результатів тестування.
- 2.7.6. Уникати надмірного узагальнення та перенесення результатів одного тестування на людські риси чи характеристики, що не вимірюються цим тестом.
- 2.7.7. Ураховувати надійність кожної шкали, похибку вимірювання та інші характеристики, які можуть штучно занижити чи підвищити результати під час їх інтерпретації.
- 2.7.8. Уважно розглядати наявні підтвердження валідності, зважаючи на конструкт (construct), що вимірюється у членів відповідної демографічної групи екзаменованих (наприклад, культурні, вікові, соціальні та гендерні групи).
- 2.7.9. Використовувати прохідний бал (cut-scores) в інтерпретації тесту тільки за наявності доказів валідності прохідного бала та доцільності його використання.
- 2.7.10. Усвідомлювати негативні соціальні стереотипи, щодо членів груп екзаменованих (наприклад, культурні, вікові, соціальні та гендерні групи), та уникати інтерпретації тестів за підтримки цих стереотипів.
- 2.7.11. Зважати на будь-яке індивідуальне чи групове відхилення від стандартної процедури в проведенні тестування.
- 2.7.12. Ураховувати будь-яке підтвердження попереднього досвіду проходження таких тестів, якщо існують дані про його вплив на успішність тестування.

2.8. Чітко і точно повідомляти результати тестування відповідним сторонам

ДОДАТКИ

Компетентні користувачі тестів повинні:

- 2.8.1. Визначати відповідні сторони, які офіційно мають право на отримання результатів тестування.
- 2.8.2. За попередньою згодою екзаменованих чи їхніх офіційних представників надавати письмові або усні звіти відповідним зацікавленим сторонам.
- 2.8.3. Переконаватися, що технічний та мовний рівень будь-яких звітів відповідає рівню розуміння отримувачів.
- 2.8.4. Давати чітко зрозуміти, що тестові дані є лише одним джерелом інформації і завжди повинні розглядатися разом з іншою інформацією.
- 2.8.5. Пояснювати, яким чином важливість результатів тестування враховується в контексті іншої інформації про екзаменованих.
- 2.8.6. Використовувати форму та структуру звітності, узгоджену з контекстом оцінювання.
- 2.8.7. У разі необхідності надавати особам, які приймають рішення (decision-makers), інформацію про використання результатів тестування для підтримки їх рішень.
- 2.8.8. Пояснювати та обґрунтовувати використання результатів тестування для класифікації людей за категоріями (наприклад, з діагностичною метою або для відбору на роботу).
- 2.8.9. Включати до письмового звіту стислий підсумок або по можливості конкретні рекомендації.
- 2.8.10. Надавати екзаменованим усний зворотний зв'язок у конструктивній та підтримуючій формах.

2.9. Переглядати доцільність тесту та його використання

Компетентні користувачі тестів повинні:

- 2.9.1. Відслідковувати зміни, що відбуваються з часом у середовищі екзаменованих, та періодично переглядати використовувані критерії вимірювання.
- 2.9.2. Запобігати небажаного впливу від тестування.
- 2.9.3. Усвідомлювати необхідність переоцінки використання тесту, якщо змінюється його форма, зміст чи спосіб проведення.
- 2.9.4. Усвідомлювати необхідність переоцінки доказів валідності, якщо змінюється мета використання тестування.
- 2.9.5. Намагатися підтверджувати валідність тестів для мети їх проведення або брати участь у формальних дослідженнях валідності.
- 2.9.6. Допомогати оновлювати інформацію стосовно норм, надійності та валідності тесту шляхом надання відповідних даних розробникам тестів, дослідникам та видавцям.

Посилання

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1985). *Standards for Educational and Psychological Testing*. Washington DC: American Psychological Association.
- Bartram, D. (1995). The Development of Standards for the Use of Psychological Tests in Occupational Settings: The Competence Approach. *The Psychologist*, May, 219-223.
- Bartram, D. (1996). Test Qualifications and Test Use in the UK: The Competence Approach. *European Journal of Psychological Assessment*, **12**, 62-71.
- Canadian Psychological Association. (1987). *Guidelines for Educational and Psychological Testing*. Ottawa: Canadian Psychological Association.
- Eyde, L. D., Moreland, K. L. & Robertson, G. J. (1988). *Test User Qualifications: A Data-based Approach to Promoting Good Test Use*. Report for the Test User Qualifications Working Group. Washington DC: American Psychological Association.
- Eyde, L. D., Robertson, G. J., Krug, S. E. et al (1993). *Responsible Test Use: Case Studies For Assessing Human Behaviour*. Washington DC: American Psychological Association.
- Fremer, J., Diamond, E. E. & Camara, W. J. (1989). Developing a Code of Fair Testing Practices in Education. *American Psychologist*, **44**, 1062-1067.
- Hambleton, R. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment*, **10**, 229-244.
- Joint Committee on Testing Practices. (1988). *Code of Fair Testing Practices in Education*. Washington DC: Joint Committee on Testing Practices.
- Joint Committee on Testing Practices. (2000). Rights and Responsibilities of Test Takers: Рекомендації and Expectations. Washington DC: Joint Committee on Testing Practices.
- Kendall, I., Jenkinson, J., De Lemos, M. & Clancy, D. (1997). *Supplement to Guidelines for the use of Psychological Tests*. Australian Psychological Society.
- Moreland, K. L., Eyde, L. D., Robertson, G. J., Primoff, E. S. & Most, R. B. (1995). Assessment of Test User Qualifications: A Research-Based Measurement Procedure. *American Psychologist*, **50**, 14-23.
- Schafer, W. D. (1992). *Responsibilities of Users of Standardized Tests: RUST Statement Revised*. Alexandria, VA: American Association for Counseling and Development.
- Van de Vijver, F. & Hambleton, R. (1996). Translating tests: some practical Guidelines. *European Psychologist*, **1**, 89-99.

Додаток А. Рекомендації щодо основ політики тестування

Ці Рекомендації стосуються потреб організацій розглядати їх політику тестування систематично та забезпечувати чітке розуміння її суті всіма особами, задіяними в процесі. Необхідність існування прозорої політики тестування стосується не тільки великих організацій. Малі та середні підприємства, що використовують тестування, так само, як і великі, повинні приділяти увагу політиці тестування поряд з питаннями охорони здоров'я чи безпеки виробництва, рівних прав, непрацездатності та інших питань, що стосуються належної практики управління та розвитку персоналу.

Хоча наступні міркування або вимоги повинні бути адаптовані для використання індивідуальними користувачами тестів як єдиними професійними практиками, необхідно мати чітке розуміння власної політики, а також здатність повідомляти її іншим.

Політика тестування розробляється для:

- досягнення цілей окремих осіб та організацій;
- уникнення потенційного неналежного використання;
- демонстрації відданості належній практиці;
- забезпечення використання тестування з належною метою;
- уникнення несправедливої дискримінації;
- забезпечення оцінювання на основі повної та доцільної інформації;
- забезпечення використання тестів тільки кваліфікованим персоналом.

Політика тестування повинна стосуватись більшості з наступних питань:

- *належне використання тестів;*
- безпека матеріалів та результатів (security of materials and scores);
- які особи можуть проводити тестування, розраховувати тестовий бал та інтерпретувати результати;
- кваліфікаційні вимоги до користувачів тестів;
- тренування користувачів тестів;
- підготовка екзаменованих;
- доступ до матеріалів та безпека;

- забезпечення конфіденційності доступу до результатів тестування та тестових балів;
- повідомлення результатів екзаменованим;
- визначення обов'язків екзаменованих до, під час та після тестування;
- визначення обов'язків та підзвітності кожного окремого користувача.

Будь-яка політика повинна регулярно переглядатися та оновлюватися відповідно до розвитку тестування та змін у практиці.

Зацікавлені сторони повинні мати доступ до інформації щодо політики тестування.

Відповідальність за політику тестування будь-якої організації повинна покладатися на кваліфікованого користувача тестів, який має повноваження забезпечити її впровадження та виконання.

Додаток В. Рекомендації щодо розробки контрактів між сторонами, задіяними у процесі тестування

Контракти між користувачем тесту та екзаменованим повинні узгоджуватися з належною практикою, законодавством та політикою тестування користувача тесту. Нижче наводиться приклад переліку питань, які може охоплювати такий контракт. Деталі можуть змінюватися залежно від контексту оцінювання (наприклад, професійне, освітнє, клінічне, судове) та місцевого чи національного законодавства.

Контракти між користувачем тесту, екзаменованим та іншими сторонами часто є неявними та розуміються за умовчанням (принаймні частково). Чітке визначення, ролей та обов'язків усіх сторін допоможе уникнути непорозумінь, шкоди та судових позовів.

Зі свого боку, користувач тесту намагатиметься:

ДОДАТКИ

- b.1 інформувати екзаменованих про їх права щодо використання тестових балів та прав доступу до них⁴⁹;
- b.2 надавати адекватне попередження про будь-які фінансові видатки, викликані процесом тестування, указувати особу, відповідальну за їх сплату та термін сплати;
- b.3 поводитися з екзаменованим ввічливо, шанобливо та неупереджено, незалежно від раси, статі, віку, дієздатності тощо;
- b.4 використовувати тести доведеної якості, доцільні для екзаменованих та мети оцінювання;
- b.5 інформувати екзаменованих перед тестуванням про мету оцінювання, характер тесту, особу, якій будуть повідомлятися результати тестування, та про плановане використання результатів;
- b.6 заздалегідь повідомляти про час проведення тестування, коли будуть готові результати, про можливість екзаменованих або інших осіб отримати примірники тесту, заповнені форми відповідей або тестові бали⁵⁰;
- b.7 проводити тестування за допомогою спеціально навченої особи і інтерпретувати результати із залученням кваліфікованої особи;
- b.8 переконатися, що екзаменований усвідомлений щодо обов'язковості (необов'язковості) тестування, а також наслідків участі та неучасті у складанні тестового іспиту, якщо воно є обов'язковим;
- b.9 переконатися, що екзаменовані розуміють умови (за їх наявності) для перескладання тестового іспиту, перерахунку тестового бала чи його ануляції;
- b.10 переконатися, що екзаменований знає, що результати тестування будуть йому роз'яснені в найкоротший термін після складання тестового іспиту за допомогою зрозумілої термінології;
- b.11 переконатися, що екзаменовані розуміють конфіденційність результатів тестування в межах дозволеного законом та найліпшими практиками;
- b.12 інформувати екзаменованих щодо осіб, що матимуть доступ до їх результатів, та умов надання тестового балу;

⁴⁹ У різних країнах існує різне законодавство з цього приводу. Наприклад, чинний у Великобританії Data Protection Act надає інше право доступу до даних, збережених у комп'ютері, ніж до даних на паперових носіях.

⁵⁰ Зазвичай усюди тестові матеріали та бланки відповідей не надаються стороннім. Водночас, у практиці різних країн існують відмінності стосовно того, що надається екзаменованим чи іншим сторонам. Проте ще більші відмінності існують в інформації, яку вони отримують. Тому важливо в контракті чітко вказати, яку саме інформацію отримують екзаменовані, а яку — ні.

- b.13 переконатися, що екзаменовані ознайомлені з процедурою подання скарг, або вирішення проблем;
Користувач тестів буде інформувати екзаменованих про те, що від них очікується, а саме:
- b.14 під час процесу тестування ставитися до інших ввічливо та з повагою;
- b.15 у разі виникнення сумнівів ставити запитання до початку тестування; запитання можуть стосуватися того, навіщо проводиться тестування, як воно проводитиметься, що вони повинні робити, як будуть використовуватись результати.
- b.16 сповіщати відповідну особу про умови, за яких, на їхню думку, результати тестування втрачають чинність, чи до яких вони хотіли б привернути увагу;
- b.17 виконувати вказівки особи, що проводить тестування;
- b.18 усвідомлювати та приймати наслідки нескладання тестового іспиту (неучасті в іспиті);
- b.19 забезпечити оплату послуг тестування (за наявності) у зазначений термін.

Додаток С. Питання, які слід брати до уваги під час підготовки до тестування осіб з обмеженими можливостями

Значні обережність та досвід необхідні для зміни способу проведення тестування відповідно до потреб осіб з обмеженими можливостями. Як завжди, необхідно враховувати місцеве та національне законодавство та практику⁵¹, а також поважати право особи на захист особистого життя. Збираючи інформацію про типи та рівні обмежених можливостей, необхідно враховувати тільки ті факти, що стосуються здатності особи виконувати необхідну під час тестування діяльність. Особлива обережність потрібна в разі тестування з метою працевлаштування⁵².

⁵¹ У США, наприклад, слід брати до уваги положення the Americans with Disabilities Act (1990). У Великобританії, the Disability Discrimination Act (1995), Employment Code of Practice вказує, що «роботодавці повинні переглядати тести — чи спосіб оцінювання результатів таких тестів — для врахування кандидатів з обмеженими можливостями».

⁵² Для детального роз'яснення цього питання в США див., see Eyde, Nestor, Heaton and Nelson (1994).

Не існує простого правила справедливого проведення тестування для осіб з усіма типами обмежених можливостей. Це справа професійного судження, щодо використання альтернативної форми оцінювання або модифікації тесту або способу його проведення. На практиці рідко вдається нормувати модифіковані тести на достатній вибірці осіб з обмеженими можливостями для забезпечення порівняння зі стандартизованою версією тесту. Проте, якщо існують дані щодо, наприклад, впливу зміни часових обмежень, використання тестів абеткою Брайля чи в аудіозапису, слід керуватися ними для вжиття необхідних заходів. Хоча повна стандартизація модифікованої версії і не буде можлива, по можливості потрібно проводити пілотне тестування на невеликій вибірці.

Ураховуючи брак інформації щодо успішності осіб з обмеженими можливостями на тестуванні (модифікованому чи ні), часто доцільніше використовувати результати тестування у більш квалітативний спосіб. Вони можуть використовуватися для визначення оцінюваних характеристик (здатність, мотивація, особистість, тощо) з доповненням інформацією, отриманою іншими методами.

Для індивідуального оцінювання екзаменатор може зазвичай «підігнати» процедуру оцінювання під можливості оцінюваного. Проте в груповому тестуванні виникають певні проблеми (наприклад, при виборі для працевлаштування). Тут можуть з'являтися труднощі зі зміною способу проведення тестування для окремих осіб із групи. Більше того, усі учасники можуть розцінювати відмінність у ставленні як несправедливість. Наприклад, якщо на виконання тесту дається більше часу, особи з обмеженими можливостями можуть сприймати це, як «особливе» ставлення, а інші вважатимуть, що додатковий час дає несправедливу перевагу.

Поради стосовно особливих потреб можна отримати у відповідних організаціях інвалідів, а також від окремих екзаменованих. Загалом найкраще (якщо це дозволено законодавством) довідатися в особи, чи існують певні моменти, які слід прийняти до уваги⁵³. У багатьох випадках такі консультації дозволяють запровадити прийнятні модифікації до умови проведення тестування без унесення змін до самого тесту.

Наступний базовий протокол дає загальні рекомендації для прийняття рішення про доцільність та спосіб модифікації тестування. В основному, обмежені можливості можуть не спричинити жодних варіатив-

⁵³ У Великобританії, the Disability Discrimination Act (1995) також покладає певну відповідальність на особу за привернення уваги до її потреб.

ностей тестового бала або спричинити варіативність щодо конструкту (construct relevant variance) чи не пов'язану з ним (construct irrelevant variance). У першому випадку, модифікації не потрібні. У другому — (construct relevant variance) вони вплинуть на релевантність тестового бала. В останньому випадку модифікації повинні бути спрямовані на усунення іррелевантного джерела варіативності (irrelevant source of variance) (відповідною модифікацією умов тестування чи заміни більш зручним тестом).

- c.1 Чи вплинуть обмежені можливості на успішність складання тестового іспиту? Багато людей мають обмежені можливості, що не впливають на успішність складання тесту. У таких випадках недоцільно створювати їм окремі умови.
- c.2 Якщо обмежені можливості можуть вплинути на успішність складання тестового іспиту, то чи є вплив на успішність при-
таманним конструкту, що вимірюється? Наприклад, особа з артритом руки під час іспиту буде мати проблеми зі швидкістю писання. Якщо здатність швидко виконувати ручні дії є частиною вимірюваного конструкту, тест змінювати не треба. Проте, якщо метою є оцінка швидкості візуального реагування, тоді доцільно застосувати іншу форму відповіді.
- c.3 Якщо певні обмежені можливості не стосуються вимірюваного конструкту, але можуть вплинути на індивідуальну успішність складання тестового іспиту, слід розглянути можливість модифікації процедури.
- c.4 Користувачі повинні керуватися тестовою інструкцією чи радитися з видавцем тесту стосовно модифікацій чи інформації про альтернативні формати та процедури.
- c.5 Користувачі повинні завжди звертатися до відповідних організацій інвалідів для отримання порад щодо можливих наслідків певних обмежених можливостей, відповідної літератури, документації та інформації про пристосування чи спеціальні умови, які можуть знадобитися.
- c.6 Будь-які модифікації до тесту чи процедури тестування повинні бути ретельно задокументовані та обґрунтовані.

Додаток D. Умови, якими слід керуватися при перекладі Рекомендацій з використання Тестів МТК

Наступні умови стосуються офіційних версій Рекомендацій Міжнародної тестової комісії (МТК) з використання тестів. Такі умови деле-

гують право перевірки якості та точності перекладу місцевим Національним Психологічним Асоціаціям.

1. У кожній країні повинен бути тільки один офіційний переклад.
2. Авторські права на оригінальну версію залишаються у МТК. МТК надає дозвіл на створення офіційної версії під егідою місцевої Національної Психологічної Асоціації, копія якої надається в МТК. Відповідальність за точність перекладу несе Психологічна Асоціація.
3. Лист від відповідного службовця місцевої Національної Психологічної Асоціації, що офіційно підтверджує точність перекладу, повинен надаватися МТК.
4. МТК зберігає за собою право розповсюджувати копії перекладу без матеріального відшкодування Психологічній Асоціації та розміщувати Рекомендації на веб-сайті МТК.
5. Офіційна версія повинна носити назву:
«Рекомендації Міжнародної тестової комісії (ІТС) з використання тестів [Мова] Версія. Офіційні права на переклад належать [повна назва Психологічної Асоціації]».
6. Офіційна версія мусить мати чіткі зображення логотипів Психологічної Асоціації та МТК.
7. Власне Рекомендації розповсюджуються безкоштовно або в безприбутковий спосіб. Доповнюючі документи, форми заяв, кваліфікаційні процедури та інші документи, розроблені на місцях на основі Рекомендацій, можуть розповсюджуватися на комерційній основі.
8. Діють звичайні правила застереження авторських прав. Для публікації частин необхідно звертатися за дозволом. МТК передає право на надання дозволів стосовно версії у місцевому перекладі місцевій Психологічній Асоціації, відповідальній за переклад. Примітка про це повинна міститися на документі Рекомендацій.

Переклад Н. В. Александрович.

Редактор перекладу М. Р. Мруга.

Науковий редактор перекладу І. Є. Булах.

© Центр тестування при МОЗ України, 2007.

Рекомендації ІТС з адаптації тестів

Рекомендації ІТС з адаптації тестів [6] (редакція 1996 року) були розроблені комісією з 13 представників низки міжнародних організацій. Метою була розробка детальних рекомендацій щодо адаптування психологічних та освітніх тестів для використання в різних мовних та культурних середовищах [39]. Це надзвичайно важливе питання, оскільки тести, розроблені в одній країні, перекладаються та адаптуються до використання у все більшій і більшій кількості інших країн. Адаптація повинна враховувати всі культурні аспекти середовища, в якому тест буде використовуватися. Hambleton [40] детально описує проект та визначає 22 рекомендації, що отримані в ході проекту. Ці рекомендації підпадають під чотири основні категорії, які стосуються:

- культурного контексту;
- технічних деталей розробки інструменту та адаптації;
- адміністрування тесту;
- документування та інтерпретації.

Усі категорії, за винятком другої, мають прямий вплив на використання тесту та його користувачів.

У 2006 році ІТС було проведено у Брюсселі конференцію з адаптації тестів для різних мов та культур⁵⁴. На цій конференції були розглянуті основні напрями оновлення Рекомендацій ІТС з адаптації тестів.

Міжнародні Рекомендації ІТС з комп'ютерного та інтернет-тестування

В останні роки відбувся значний та швидкий розвиток у сфері надання незалежного (stand-alone) та комп'ютерного інтернет-тестування. Це ставить низку запитань стосовно стандартів адміністрування, охорони (security) тестів та тестових результатів, а також контролю процесу тестування. Оскільки ринок такого тестування збільшується, а технологічна складність відповідних продуктів підвищується, зростає необхідність у дотриманні належних практик тими, хто розробляє, розповсюджує, використовує та проходить такі тести і інструменти оцінювання. У зв'язку з цим, Рада ІТС вирішила збільшити інвестиції у науково-дослідну програму, метою якої є розробка міжнародно узгоджених рекомендацій, спеціально спрямованих на комп'ютерне/Інтернет тестування.

⁵⁴ The ITC 5th International Conference on Psychological and Educational Test Adaptation across Language and Cultures — http://www.intestcom.org/itc_conference_2006.htm

Кінцевою метою проекту було:

- розробити набір міжнародно визнаних рекомендацій, що підкреслюють важливість належної практики у комп'ютерному (Computer-Based) тестуванні (CBT) та тестуванні через Інтернет;
- підвищити серед усіх зацікавлених сторін процесу тестування усвідомлення поняття складових відповідної практики.

Завдяки детальним та ретельним дослідженням деяких осіб та організацій, що займаються комп'ютерним та інтернет-тестуванням, у 2005 році була випущена узгоджена версія Міжнародних Рекомендацій ІТС з комп'ютерного та інтернет-тестування [7]. Для доступу на офіційний веб-сайт Рекомендацій ІТС слід використовувати зв'язок — <http://www.intestcom.org/guidelines>.

EFPA. Модель аналізу та оцінювання психологічних тестів

Модель аналізу та оцінювання психологічних тестів [8] та критерії за формою і змістом були значною мірою розроблені на основі Критеріїв аналізу тестів Британської психологічної спілки (BPS) та Критеріїв, розроблених Комітетом із справ тестування⁵⁵ (COTAN) Голландської асоціації психологів⁵⁶ (NIP).

Дейв Бартрам (Dave Bartram) і Патрісія Ліндлі (Patricia Lindley) спочатку розробили критерії та процедури аналізу для Служби працевлаштування Великобританії (UK Employment Service), а пізніше розширили їх для використання BPS; Арне Еверс (Arne Evers) редагував голландську рейтингову систему якості тестів. EFPA висловлює вдячність BPS та NIP за дозвіл використання їхніх критеріїв для розробки європейської моделі. EFPA також висловлює вдячність Дейву Бартраму, Арне Еверс та Патрісії Ліндлі за їхній внесок у розробку цієї моделі. Усі права інтелектуальної власності оригінальних критеріїв BPS та NIP визнаються та залишаються за цими організаціями.

Постійна Комісія з тестів та тестування EFPA схвалила ці коментарі

⁵⁵ Комітет із справ тестування — *Committee of Test Affairs (COTAN)* — <http://www.psynip.nl/default.asp?URL2=subpage.asp&topmenuID=2&submenuID=11>.

⁵⁶ Голландська асоціація психологів — *Dutch Association of Psychologists (Nederlands Instituut van Psychologen, NIP)* — <http://www.psynip.nl/>.

для оглядачів та відповідний формат аналізу. Психологічні Асоціації-члени можуть використовувати їх як основу для власних процедур аналізу тестових інструментів. Поширення цього документа має на меті сприяти гармонізації процедур та критеріїв аналізу в Європі. Вітається надходження коментарів до цих документів, оскільки досвід оглядачів тестових інструментів дасть змогу поліпшити та прояснити процеси.

Оригінальна версія скомпільована та відредагована Дейвом Бартрамом. Її було оновлено та переглянуто Патрісією Ліндлі, Дейвом Бартрамом та Наталі Кеннеді у квітні 2004-го та у 2005 р.⁵⁷ Сучасна версія 3.41 була прийнята у серпні 2005 р. і представляє собою анкету для оцінювання якості психологічних тестів. Деякі з таких тестів можуть використовуватися і при педагогічному оцінюванні.

Щодо деталей процедури аналізу та внесення будь-яких змін до змісту та критеріїв необхідно звертатися до місцевої психологічної асоціації. Названий документ описує процедури, що залучають двох оглядачів для аналізу кожного тесту, а третя особа — консультуючий редактор (the «Consulting Editor») — контролює процес аналізу. Старший редактор (Senior Editor) відповідає за забезпечення однакового застосування критеріїв у різних випадках аналізу. На місцях деякі з цих функцій можуть поєднуватися.

EFPA відзначає, що оцінювання в цих оглядах орієнтоване на кваліфікованих користувачів тестів — практиків, проте може становити інтерес і для вчених, фахівців з психометрії та психологічного тестування.

⁵⁷ Ця версія є інтеграцією двох окремих документів: Форми огляду (Review Form) та Коментарів для Оглядачів (Notes for Reviewers)). Додатково зміст був відредагований та змінений для використання BPS для їх онлайн-оглядачів тестів. Версія 3.4 була прийнята на заміну всіх попередніх версій на Генеральній Асамблеї EFPA в липні 2005 р.

В. ІНШІ РЕГУЛЯТОРНІ ДОКУМЕНТИ

BUROS CENTER FOR TESTING

**BUROS
INSTITUTE***Buros Institute of Mental Measurements*

На сайті цієї авторитетної американської організації наведені регуляторні документи, що стосуються унормування багатьох аспектів в галузі педагогічних вимірювань та оцінювання, і які, на думку авторів цієї книги, є дуже важливими для освітян України. Це такі документи, як:

1. *Кодекс професійної відповідальності при педагогічних вимірюваннях (Code of Professional Responsibilities in Educational Measurement).*
2. *Стандарти компетентності вчителів з оцінювання учнів (Standards for Teacher Competence in Educational Assessment of Students).*
3. *Стандарти компетентності освітніх адміністраторів з оцінювання учнів (Competency Standards in Student Assessment for Educational Administrators).*

Нижче наводяться повні тексти перекладів цих документів українською мовою, які здійснено Центром тестування при МОЗ України.

Оригінали документів наведено у додатку С.

Кодекс професійної відповідальності при педагогічних вимірюваннях

Підготовлено Комісією з розробки Кодексу етики Національної Ради з вимірювань в освіті (1995) [9].

Автори документа: Cynthia B. Schmeiser, ACT — Chair, Kurt F. Geisinger, State University of New York, Sharon Johnson-Lewis, Detroit Public Schools, Edward D. Roeber, Council of Chief State School Officers, William D. Schafer, University of Maryland.

Будь-яка частина цього Кодексу може бути відтворена та розповсюджена для освітніх цілей.

Як організація, що займається поліпшенням практики вимірювання та оцінювання в освіті, Національна Рада з вимірювань в освіті (NCME) прийняла цей Кодекс для впровадження професійно відповідальної практики у педагогічних вимірюваннях.

Професійно відповідальна практика — це поведінка, що керується професійними стандартами в певній сфері, загальними етичними принципами або тим і тим разом.

Прийняття **Кодексу професійної відповідальності при педагогічних вимірюваннях**, який у подальшому іменуватиметься Кодекс, є результатом діяльності членів Національної Ради з вимірювань в освіті (NCME), котрі займаються різноманітними видами оцінювальної діяльності в галузі освіти. NCME також безоплатно надає цей Кодекс у користування всім, хто займається оцінювальною діяльністю, сподіваючись, що така діяльність проводитиметься у професійно відповідальний спосіб. Коло осіб, які займаються такою діяльністю, включає місцевих працівників освіти (шкільних учителів, директорів шкіл, інспекторів); фахівців (шкільних психологів та радників); службовців державного та національного рівнів, які займаються технічними, юридичними та політичними питаннями в освіті; працівників дослідних, оцінювальних та тестових організацій; провайдерів послуг із підготовки до тестування; викладачів та адміністраторів коледжів та університетів; фахівців із бізнесу і промисловості, що розробляють та впроваджують освітні та тренінгові програми.

Кодекс стосується різних форм та типів оцінювання, що є частиною освітнього процесу, включаючи формальні та неформальні, традиційні та альтернативні технології збору інформації для використання при прийнятті освітніх рішень усіх рівнів. Ці технології включають (проте не обмежуються переліченим): широкомасштабне оцінювання на рівні школи, району, штату, національному та інтернаціональному рівнях; стандартизовані тести; методики спостереження; оцінювання вчителем (teacher-conducted assessments); допоміжні матеріали з оцінювання (assessment support materials); вимірювання досягнень, здатності, зацікавленості та інших характеристик особистості, що використовуються в освіті та з освітньою метою.

Хоча NCME розповсюджує цей Кодекс серед своїх членів, вона на постійно закликає й інші організації та осіб, що займаються освітнім

ДОДАТКИ

оцінюванням, дотримуватися його в професійній діяльності. Оскільки Кодекс стосується тільки використання оцінювання в освіті, припускається, що використання оцінювання поза освітнім контекстом, наприклад для працевлаштування, сертифікації чи ліцензування, може включати додаткові аспекти професійної відповідальності, що не входять в названий Кодекс.

Кодекс повинен виконувати освітню функцію: інформувати та нагадувати тим, хто займається освітнім оцінюванням, про їхній обов'язок підтримувати сумління ставлення (*uphold the integrity of the manner*) до розробки, використання, оцінювання для визначення якості (*evaluate*) та маркетингу інструментів оцінювання. Окрім того, очікується, що **Кодекс буде стимулювати ґрунтовне обговорення складових професійно відповідальної практики оцінювання на всіх освітніх рівнях.**

Кодекс визначає типи професійної відповідальності у восьми основних сферах оцінювальної діяльності. А саме — описує професійну відповідальність тих, хто займається переліченою нижче діяльністю щодо інструментів педагогічного оцінювання успішності навчання (*assessments*):

1. Розробка.
2. Маркетинг та продаж.
3. Відбір.
4. Адміністрування/проведення оцінювання.
5. Розрахунок тестового бала (*scoring*).
6. Інтерпретація, використання та повідомлення результатів оцінювання.
7. Навчання інших в галузі оцінювання.
8. Оцінка освітніх програм та проведення досліджень в галузі оцінювання.

Хоча структура Кодексу базується на диференціації цих видів діяльності, вони надзвичайно взаємопов'язані, і ті, хто користується цим Кодексом, повинні розглядати їх цілісно. Предметний покажчик в кінці Кодексу вказує список тем із різних видів оцінювальної діяльності, які становлять особливий інтерес у педагогічному вимірюванні.

Типи загальної відповідальності

Професійна відповідальність, проголошена цим Кодексом у восьми основних сферах оцінювальної діяльності, передбачає, що фахівці, які займаються освітнім оцінюванням, повинні:

1. *Захищати здоров'я та безпеку всіх екзаменованих.*
2. *Діяти згідно з місцевими та федеральними законами щодо здійснення професійної діяльності.*
3. *Підтримувати та підвищувати рівень професійної компетентності в галузі освітнього оцінювання.*
4. *Надавати послуги з оцінювання лише у сфері своєї компетенції та досвіду, що сприятиме повному застосуванню своєї професійної кваліфікації.*
5. *Пропагувати розуміння надійних практик оцінювання в освіті.*
6. *Дотримуватися найвищих стандартів діяльності та пропагувати професійно відповідальну діяльність в освітніх закладах та агентствах, що надають освітні послуги.*
7. *Виконувати професійні обов'язки чесно, надійно та сумлінно.*

Відповідальна професійна практика включає поінформованість про діяльність згідно з Кодексом справедливих практик тестування в освіті (Code of Fair Testing Practices in Education, Joint Committee on Testing Practices, 1988), Стандартами освітнього та психологічного тестування (The Standards for Educational and Psychological Testing, American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 1985), або їхніми наступними переглянутими версіями, а також згідно з відповідним законодавством штату чи федеральним законодавством, яке регулює розробку, адміністрування та використання оцінювання. Стандарти освітнього та психологічного тестування і Кодекс справедливих практик тестування в освіті мають на меті встановлення критеріїв щодо технічної адекватності тестів, доцільного використання тестів та результатів тестування. Завдання цього Кодексу — описати професійну відповідальність (професійні обов'язки) осіб, які займаються оцінювальною діяльністю. Абсолютно логічно, що між професійно відповідальною практикою та надійним освітнім оцінюванням існує тісний зв'язок і цей Кодекс має узгоджуватися з відповідними частинами цих двох документів.

NCME не ставить своїм завданням упроваджувати професійну відповідальність, описану в цьому Кодексі, чи розслідувати випадки порушення Кодексу.

Оскільки Кодекс стосується принципів оцінювання для визначення якості (frame of reference for the evaluation) відповідності/доречності поведінки, NCME визнає, що названий Кодекс може використовуватися в судових та інших подібних справах.

ІНСТРУМЕНТИ ТА ПОСЛУГИ

Розділ 1: Відповідальність тих, хто розробляє оцінювання

Розробники інструментів та послуг з оцінювання (наприклад, шкільні вчителі та інші фахівці з оцінювання) повинні розробляти інструменти оцінювання найвищої якості. Професійна відповідальність розробників інструментів оцінювання полягає у такому:

- 1.1. *Забезпечувати розробку інструментів та послуг з оцінювання за відповідними професійними, технічними та правовими стандартами.*
- 1.2. *Розробляти інструменти та послуги з оцінювання без упередженості щодо характеристик, які не мають прямого відношення до вимірюваного конструкту (до предмета оцінювання), — таких, як стать, національність, раса, соціоекономічний статус, обмежені можливості, віросповідання, вік та країна походження.*
- 1.3. *При розробці враховувати можливі модифікації/поправки для груп екзаменованих з обмеженими можливостями чи іншими особливими потребами.*
- 1.4. *Повідомляти зацікавленим сторонам про наявний чи потенційний конфлікт інтересів, який може вплинути на рішення чи діяльність розробників.*
- 1.5. *Використовувати в інструментах та послугах з оцінювання матеріали із застереженням авторським правом лише згідно із законодавством штату чи федерального законодавства.*
- 1.6. *Надавати відповідним особам інформацію про вжиті заходи з розробки інструментів оцінювання та визначення тестових балів (score), включаючи необхідну інформацію про докази надійності, валідності, процеси розрахунку тестового бала (scoring) та звітності, а також інші доцільні характеристики оцінювання.*
- 1.7. *Забезпечувати захист права екзаменованих на конфіденційність інформації (privacy), визнаючи це як частину процесу розробки інструментів оцінювання.*
- 1.8. *Чітко і однозначно попереджувати користувачів інструментів оцінювання про найбільш імовірні помилки в інтерпретації та використанні даних, що впливають з процесу розробки інструментів оцінювання.*
- 1.9. *Уникати неправильних чи непідтверджених висловлювань щодо оцінювання, його використання та інтерпретацію ре-*

зультатів у матеріалах для підготовки до тестування, супровідних матеріалах до програми та послуг.

- 1.10. Виправляти будь-які суттєві неточності в інструментах оцінювання чи супровідних матеріалах до нього в найкоротші можливі терміни.
- 1.11. Розробляти форми звітності (score reports) та супровідні матеріали в доступній для розуміння результатів оцінювання формі.

Розділ 2: Відповідальність тих, хто займається маркетингом і продажем інструментів та послуг з оцінювання

Маркетинг інструментів та послуг з оцінювання, таких, як тести та інші інструменти, послуги з розрахунку тестового бала (scoring services), послуги з підготовки до тестування, консультації та послуги з інтерпретації результатів тестування, повинен базуватися на точній, повній та прийнятній для користувачів інформації. Професійна відповідальність осіб, які займаються маркетингом і продажем інструментів та послуг з оцінювання, полягає у такому :

- 2.1. Надавати точну інформацію потенційним покупцям про інструменти та послуги з оцінювання, їх рекомендоване використання та можливості.
- 2.2. Не приховувати навмисно відповідну інформацію про інструменти та послуги з оцінювання, що може вплинути на відповідне рішення щодо вибору.
- 2.3. Базувати всі висловлювання (claims) про властивості інструментів та послуг з оцінювання на валідній оцінці відкритої та доступної широкому загалу інформації.
- 2.4. Надавати кваліфікованим користувачам рівний доступ до придбання інструментів та послуг з оцінювання.
- 2.5. Установлювати розумні ціни на інструменти та послуги з оцінювання.
- 2.6. Повідомляти потенційним користувачам заздалегідь, до придбання чи користування, про всі можливі платежі стосовно інструментів та послуг з оцінювання.
- 2.7. Намагатися нікого не позбавляти можливостей використання інструментів та послуг з оцінювання через неспроможність за них сплатити.

ДОДАТКИ

- 2.8. Установити критерії продажу інструментів та послуг з оцінювання, які б обмежували коло покупців тими, хто має право (*who are qualified*) на їхнє використання за призначенням та від кого очікується відповідне використання та інтерпретація.
- 2.9. Інформувати потенційних користувачів про відомі випадки використання інструментів та послуг з оцінювання не за призначенням та надавати рекомендації щодо уникнення таких ситуацій.
- 2.10. Підтримувати актуальний рівень розуміння інструментів та послуг з оцінювання та їх використання в освіті.
- 2.11. Публікувати інформацію, що містить схвалення інструментів та послуг з оцінювання користувачами тільки за згодою користувачів.
- 2.12. Уникати повідомлень про те, що інструменти та послуги з оцінювання були схвалені (*endorsed*) іншими організаціями до отримання офіційного підтвердження схвалення.
- 2.13. Уникати рекламування підготовчих матеріалів та послуг, результати яких призвели б до спотвореного уявлення про реальний рівень успішності екзаменованих.

Розділ 3: Відповідальність тих, хто займається відбором інструментів та послуг з оцінювання

Ті, котрі відбирають інструменти та послуги з оцінювання для використання в галузі освіти чи допомагають іншим у цьому, несуть серйозну професійну відповідальність щодо забезпечення відповідності інструментів та послуг з оцінювання обраний меті використання. Професійна відповідальність тих, хто відбирає інструменти та послуги з оцінювання, полягає у такому:

- 3.1. Здійснювати ретельний перегляд та оцінку наявних стратегій та інструментів оцінювання, що можуть відповідати обраний меті використання.
- 3.2. Рекомендувати та/чи відбирати інструменти та послуги з оцінювання на підставі відкритих задокументованих доказів їхньої технічної якості та придатності, а не на основі непідтверджених висловлювань чи суджень.
- 3.3. Оприлюднювати будь-які зв'язків чи співпраці з авторами, видавцями тестів та іншими особами, зацікавленими в продажі інструментів та послуг з оцінювання, та утри-

муватися від участі, якщо такі зв'язки можуть вплинути на об'єктивність процесу відбору.

- 3.4. Повідомляти осіб, які приймають рішення (*decision makers*), і майбутніх користувачів про доцільність використання інструментів та послуг з оцінювання для обраної мети, можливі наслідки використання, захист прав екзаменованого, відносну вартість, матеріали та послуги, необхідні для проведення чи застосування оцінювання, включаючи можливе нецільове використання чи помилкову інтерпретацію інформації щодо інструментів та послуг з оцінювання.
- 3.5. Не рекомендувати використовувати будь-який потенційний вид оцінювання, якщо проведення, розрахунок тестового бала чи використання може виявитися невалідним для членів різних груп суспільства через їхню расу, національність, стать, вік, обмеження можливостей, рідну мову, соціоекономічний статус, віросповідання чи країну походження.
- 3.6. Дотримуватися всіх вимог безпеки (*security precautions*), що можуть стосуватися інструментів та послуг з оцінювання, які відбираються.
- 3.7. негайно повідомляти про всі спроби чинити вплив в процесі відбору інструментів та послуг з оцінювання.
- 3.8. Не рекомендувати купувати чи використовувати підготовчі матеріали та послуги, які можуть призвести до спотвореного уявлення про реальний рівень успішності екзаменованих.

Розділ 4: Відповідальність тих, хто проводить оцінювання

Ті, хто проводить підготовку до оцінювання, та ті, хто прямо чи опосередковано задіяний в адмініструванні оцінювання як частини освітнього процесу, включаючи вчителів, адміністраторів/прокторів та персонал, що проводить оцінювання, відіграють важливу роль у забезпеченні справедливого, точного проведення оцінювання. Професійна відповідальність осіб, які проводять підготовку до оцінювання, та осіб, які проводять оцінювання, полягає у такому:

- 4.1. До проведення оцінювання сповіщати екзаменованих про мету, використання і наслідки оцінювання; про те, як інформація, отримана в результаті оцінювання, буде оброблятися, як зберігатимуться результати, хто буде мати доступ до результатів, як результати будуть розповсюджуватися, які права екзаменованого до, під час та після проведення оцінювання.

ДОДАТКИ

- 4.2.Проводити тільки ті види оцінювання, на які вони мають право, отримане шляхом освіти, тренування, ліцензування чи сертифікації.
- 4.3.Вживати відповідних заходів безпеки (*security precautions*) до, під час та після проведення оцінювання.
- 4.4.Ознайомитися з процедурами, необхідними для проведення оцінювання, до його проведення.
- 4.5.Проводити стандартизоване оцінювання згідно з рекомендованими процедурами та умовами і доводити до відома відповідних осіб про виникнення будь-яких нестандартних чи обмежувальних умов.
- 4.6.Не виключати безпідставно з групи екзаменованих жодного кандидата.
- 4.7.Під час проведення оцінювання уникати будь-яких ситуацій, які можуть анулювати його результати.
- 4.8.Робити всі прийнятні і допустимі модифікації/поправки для проведення оцінювання осіб з обмеженими можливостями чи спеціальними потребами, та документувати їх.
- 4.9.Надати достатню можливість (*reasonable opportunities*) екзаменованим ставити запитання стосовно процедур оцінювання чи інструкції перед та, у вказаний час, протягом проведення оцінювання.
- 4.10.Захищати права екзаменованих на конфіденційність та належну обробку інформації (*privacy and due process*).
- 4.11. Уникати дій чи умов, які б дали можливість чи сприяли окремим особам або групам отримати результати, що спотворювали б їхній реальний рівень успішності.

Розділ 5: Відповідальність тих, хто займається розрахунком тестового бала (scoring)

Розрахунок тестового бала повинен проводитися належним чином та ефективно для забезпечення точного та вчасного повідомлення результатів. Професійна відповідальність осіб, які займаються розрахунком тестового бала та готують звіти про проведення оцінювання, полягає у такому:

- 5.1.Надавати повну та точну інформацію користувачам про те, як проводиться розрахунок тестового бала, таку, як графік звітування (*reporting schedule*), процес розрахунку тестового

бала (scoring process), обґрунтування підходу до розрахунку (rationale for the scoring approach), технічні характеристики, процедури контролю якості, формат звітності та вартість цих послуг (у разі наявності).

- 5.2. Забезпечувати точність результатів оцінювання шляхом проведення достатніх процедур контролю якості до, під час та після розрахунку тестового бала.*
- 5.3. Мінімізувати вплив факторів, не пов'язаних з метою оцінювання, на розрахунок тестового бала.*
- 5.4. Оперативно повідомляти користувачам про будь-які зміни в послугі чи графіку розрахунку тестового бала та надання звітів і знаходити рішення шляхом переговорів з користувачами.*
- 5.5. У разі виявлення помилок, які можуть вплинути на рішення, що приймається на основі тестового бала, надати екзаменованому чи замовнику виправлені розрахунки тестового бала в якомога короткий термін.*
- 5.6. Забезпечувати конфіденційність інформації, що ідентифікує особу, згідно із законодавством штату чи федеральним законодавством.*
- 5.7. Надавати підсумкові результати оцінювання тільки особам, уповноваженим законодавством штату чи федеральним законодавством мати таку інформацію, або тим, хто вказаний стороною, яка замовляла послуги з розрахунку тестового бала.*
- 5.8. Запровадити по можливості справедливий та практичний (reasonable) процес апеляції та перерахунку результатів оцінювання.*

Розділ 6: Відповідальність тих, хто займається інтерпретацією, використанням та повідомленням результатів оцінювання

Інтерпретація, використання та повідомлення результатів оцінювання повинні сприяти валідним висновкам і мінімізувати невалідні. Професійна відповідальність осіб, що займаються інтерпретацією, використанням та повідомленням результатів оцінювання, полягає у такому:

- 6.1. Проводити цю діяльність усвідомлено, об'єктивно та справедливо в контексті лімітів (обмежень) оцінювання та з усвідомленням потенційних наслідків використання.*
- 6.2. Надавати отримувачу результатів оцінювання інформацію про інструмент оцінювання, його призначення, обмеження*

ДОДАТКИ

та використання, необхідну для належної інтерпретації результатів.

- 6.3. Надавати отримувачам звітів про тестовий бал (score reports) зрозумілий письмовий опис усіх балів, які містяться в звіті, а також належну їхню інтерпретацію і можливі при цьому помилки.
- 6.4. Повідомляти споживачам результати оцінювання зрозуміло та вчасно, включаючи об'єктивну їхню інтерпретацію і можливі при цьому помилкові інтерпретації.
- 6.5. Зважувати та повідомляти адекватність та доцільність будь-яких норм та стандартів, що використовуються в інтерпретації результатів оцінювання.
- 6.6. Інформувати сторони, задіяні в процесі оцінювання, про можливий вплив результатів оцінювання на них.
- 6.7. Використовувати (в разі можливості) декілька джерел та типів відповідної інформації про осіб чи програми для прийняття рішення в освітньому контексті (educational decisions).
- 6.8. Уникати самому та активно відраджувати інших від створення неточних звітів, висловлювання необґрунтованих тверджень, неналежних інтерпретацій чи інших неправильних та оманливих свідчень про результати оцінювання.
- 6.9. Повідомляти екзаменованим та іншим особам, чи будуть, і як довго, зберігатися результати оцінювання, про процедури апеляції та перерахунку тестового бала (rescoring), які права екзаменовані та інші особи мають щодо інформації про результати оцінювання і як практично застосовуються ці права.
- 6.10. Повідомляти особам, відповідальним за процес оцінювання, про будь-які явні випадки неналежного використання інформації щодо оцінювання.
- 6.11. Захищати право осіб та організацій, задіяних в процесі оцінювання, на конфіденційність інформації (privacy).

Розділ 7: Відповідальність тих, хто займається навчанням інших у сфері оцінювання

У процесі навчання інших у сфері освітнього оцінювання як частини вищої освіти, професійного розвитку, обговорення громадської політики чи робочого тренінгу, необхідно підготувати людей до розуміння та участі в правильній практиці вимірювання та до перетворення на обізнаних користувачів тестів і тестових результатів. Професійна

відповідальність осіб, які навчають інших чи надають їм інформацію про оцінювання, полягає у такому:

- 7.1. Бути компетентними та володіти сучасними знаннями в своїх сферах викладання і застосовувати це в практичній діяльності.*
- 7.2. Подавати матеріал про оцінювання справедливо та збалансовано.*
- 7.3. Чітко розмежовувати у процесі навчання інших певним методом, засобам чи послугам з оцінювання висловлювання думки від обґрунтованого знання.*
- 7.4. Повідомляти про будь-які фінансові інтереси, які можуть бути розцінені як фактор впливу на ставлення до певного інструменту чи послуги з оцінювання, що є предметом вивчення.*
- 7.5. Уникати проведення будь-якого оцінювання, що не є частиною оцінки успішності студента в певному навчальному курсі, якщо проведення такого оцінювання може зашкодити будь-якому студенту.*
- 7.6. Уникати використання чи повідомлення результатів будь-якого оцінювання, що не є частиною оцінки успішності студента в певному навчальному курсі, якщо використання чи повідомлення результатів такого оцінювання може зашкодити будь-якому студенту.*
- 7.7. Забезпечити інформаційний захист та збереженість усіх інструментів оцінювання та матеріалів, що використовуються в навчальному процесі. (Protect all secure assessments and materials...)*
- 7.8. Подавати особистий приклад відповідальної практики оцінювання та допомагати слухачам курсу дізнатися про свої професійні обов'язки/відповідальність в освітньому оцінюванні.*
- 7.9. Справедливо та збалансовано висвітлювати проблеми оцінювання, що обговорюються політиками, батьками та іншими громадянами.*

Розділ 8: Відповідальність тих, хто займається оцінкою освітніх програм та проведенням досліджень у сфері оцінювання

Проведення досліджень у сфері оцінювання чи освітніх програм є ключовою діяльністю для поліпшення розуміння та використання оцінювання і освітніх програм. Професійна відповідальність осіб, які займаються оцінкою освітніх програм та проведенням досліджень у сфері оцінювання, полягає у такому:

ДОДАТКИ

- 8.1.Проводити оцінювання для визначення якості (evaluation) та дослідницьку діяльність поінформовано, об'єктивно та справедливо.
- 8.2.Оприлюднювати будь-які форми зв'язків чи співпраці з авторами, видавцями тестів та іншими особами, задіяними в оцінюванні, та утримуватися від участі, якщо такі зв'язки можуть негативно вплинути на об'єктивність дослідження чи оцінки.
- 8.3.Зберігати належну конфіденційність всіх оцінювальних матеріалів протягом всього терміну дослідження.
- 8.4.Вживати необхідних заходів для мінімізації потенційних джерел невалідності в дослідженні та повідомляти про відомі факти, що можуть вплинути на неупередженість результатів дослідження.
- 8.5.Представляти результати дослідження, як очікувані, так і неочікувані, справедливо, об'єктивно, у повному обсязі.
- 8.6.Повністю та належним чином визнавати авторство роботи та ідей інших.
- 8.7.Оцінювати підсумки дослідницької діяльності в рамках дослідження.
- 8.8.Використовувати в разі можливості декілька джерел інформації для оцінки та дослідницької діяльності.
- 8.9.Дотримуватися чинних стандартів для захисту прав учасників дослідження, включаючи право на конфіденційність інформації (privacy) та поінформовану згоду.

Післямова

Як було зазначено напочатку, метою Кодексу професійної відповідальності в галузі педагогічних вимірювань є скеровування діяльності фахівців, які займаються будь-яким видом оцінювання в освіті. Беручи до уваги широкий спектр освітнього оцінювання, а також різноманітність діяльності, якою можуть займатися професіонали, можна констатувати, що жоден Кодекс не охопить усі типи професійної відповідальності у певній ситуації чи виді діяльності, коли здійснюється оцінювання в освіті. Залишається сподіватися, що цей Кодекс слугуватиме базою для подальшого обговорення професійно відповідальної практики. Більш того, такі обговорення однозначно висвітлять ті сфери практики, які потребують подальшого аналізу та пояс-

нення в наступних редакціях Кодексу. У цьому разі функцію Кодексу можна вважати певною мірою виконаною.

Додаткова література

Наступний список джерел наданий для тих, хто хоче знайти додаткову інформацію про Кодекси професійної відповідальності, що були розроблені та прийняті організаціями, задіяними в освітньому оцінюванні.

American Association for Counseling and Development (now American Counseling Association). (1988). *Ethical standards of the American Counseling Association*. Alexandria, VA: Author.

American Association for Counseling and Development (now American Counseling Association) & Association for Measurement and Evaluation in Counseling and Development (now Association for Assessment in Counseling). (1989). *Responsibilities of users of standardized tests: RUST statement revised*. Alexandria, VA: Author.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1985). *Standards for educational and psychological testing*. Washington, DC: Author.

American Educational Research Association. (1992). *Ethical standards of the American Educational Research Association*. Educational Researcher, 21(7), 23 — 26.

American Federation of Teachers, National Council on Measurement in Education, & National Education Association. (1990). *Standards for teacher competence in educational assessment of students*. Washington, DC: Author.

American Psychological Association. (1992). *Ethical principles of psychologists and code of conduct*. Washington, DC: Author.

American Psychological Association President's Task Force on Psychology in Education, (in press). *Learner-centered psychological principles: Guidelines for school redesign and reform*. Washington, DC: Author.

Joint Advisory Committee. (1993). *Principles for fair assessment practices for education in Canada*. Edmonton, Alberta: Author.

Joint Committee on Testing Practices. (1988). *Code of fair testing practices in education*. Washington, DC: Author.

Joint Committee on Standards for Educational Evaluation. (1988). *The personnel evaluation standards: How to assess systems for evaluating educators*. Newberry Park, CA: Sage.

Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards: How to assess evaluations of educational programs*. Thousand Oaks, CA: Sage.

National Association of College Admission Counselors. (1988). *Statement of principles of good practice*. Alexandria, VA: Author.

Предметний покажчик до Кодексу

Цей покажчик надає список основних тем та питань відповідальності, уміщених у восьми розділах Кодексу. Список не є повним, проте може вважатися певним джерелом посилань для користувачів Кодексу.

Поінформування (Advertising): 1.9, 1.10, 2.3, 2.11, 2.12

Зсув (Bias): 1.2, 3.5, 4.5, 4.7, 5.3, 8.4

Списування (Cheating): 4.5, 4.6, 4.11

Підготовка до тестування (Coaching and Test Preparation): 2.13, 3.8, 4.11

Компетентність (Competence): 2.10, 4.2, 4.4, 4.5, 5.2, 5.5, 7.1, 7.8, 7.9, 8.1, 8.7

Конфлікт інтересів (Conflict of Interest): 1.4, 3.3, 7.4, 8.2

Наслідки використання тестування (Consequences of Test Use): 3.4, 6.1, 6.6, 7.5, 7.6

Матеріали із застережним використанням авторського права (Copyrighted Materials, Use of): 1.5, 8.6

Інваліди, що екзамнуються, їх права (Disabled Examinees, Rights of): 1.3, 4.8

Повідомлення інформації (Disclosure): 1.6, 2.1, 2.2, 2.6, 3.3, 3.7, 4.1, 5.1, 5.4, 6.2, 6.3, 6.4, 6.6, 6.9, 8.2, 8.4, 8.5

Належна обробка даних (Due Process): 4.10, 5.8, 6.9

Рівний доступ (Equity): 1.2, 2.4, 2.7, 3.5, 4.6

Плата (Fees): 2.5, 2.6, 2.7

Неналежне використання тестів (Inappropriate Test Use): 1.8, 2.8, 2.9, 3.4, 6.8, 6.10

Об'єктивність (Objectivity): 3.1, 3.2, 3.3, 6.1, 6.5, 7.2, 7.3, 7.9, 8.1, 8.2, 8.5, 8.7

Право на конфіденційність (Rights to Privacy): 1.7, 3.4, 4.10, 5.6, 5.7, 6.11, 8.9

Забезпечення конфіденційності (Security): 3.6, 4.3, 7.7, 8.3

Правдивість (Truthfulness): 1.10, 2.1, 2.2, 2.3, 2.11, 2.12, 3.2, 4.6, 7.3

Неналежний вплив (Undue Influence): 3.7

Необґрунтовані твердження (Unsubstantiated Claims): 1.9, 3.2, 6.8

Стандарти компетентності вчителів з оцінювання учнів

Розроблено Американською федерацією вчителів, Національною Радою з вимірювань в освіті, Національною освітньою асоціацією [10].

Цей матеріал не застережений авторськими правами. Відтворення та розповсюдження вітаються. 1990.

У 1987 р. професійні освітні асоціації почали працювати над розробкою стандартів компетентності з оцінювання учнів для вчителів з метою якнайповнішої реалізації потенційних освітніх переваг оцінювання учнів. Комітет⁵⁸, призначений для виконання цього проекту, аналізуючи варіанти попередніх проектів, переглянуті фахівцями у галузі вимірювання, викладання, підготовки та сертифікації вчителів, завершив свою роботу в 1990 р. Вітаються наміри паралельних комітетів причетних асоціацій у найближчому майбутньому розробити подібні описи кваліфікацій для шкільних адміністраторів, радників, директорів з тестування, інспекторів та інших фахівців у галузі освіти. Ці описи мають на меті спрямування допрактичної (preservice) та практичної (in-service) підготовки фахівців у галузі освіти, акредитації програм підготовки та майбутньої сертифікації всіх працівників освіти. Термін **стандарт** означає принцип, загальновизнаний професійними асоціаціями, відповідальними за цей документ. Термін **оцінювання** означає процес отримання інформації для прийняття освітніх рішень щодо учнів; для інформування учнів щодо їхнього прогресу, сильних чи слабких сторін; для визначення інструктивної ефективності та відповідності навчальному плану; а також для формування політики. Різноманітні **техніки оцінювання** включають формальні та неформальні спостереження успішності учнів, якісний аналіз успішності та результатів

⁵⁸ Комітет, що розробив цей документ, був призначений співпрацюючими професійними асоціаціями: Джеймс Сандерс (James R. Sanders) (Western Michigan University) очолював комітет та представляв NCME разом з Джоном Хіллом (John R. Hills) (Florida State University) та Ентоні Нітко (Anthony J. Nitko) (University of Pittsburgh). Джек Мервін (Jack C. Merwin) (University of Minnesota) представляв Американську Асоціацію Коледжів Педагогіки (American Association of Colleges for Teacher Education), Керолін Трайс (Carolyn Trice) представляла Американську Федерацію Вчителів (American Federation of Teachers) і Марсела Дайанда (Marcella Dianda) та Джефрі Шнайдер (Jeffrey Schneider) представляли Національну Освітню Асоціацію (National Education Association).

діяльності учнів, «паперові» (paper-and-pencil) тести, усне опитування та аналіз навчального прогресу учня, проте не обмежуються ними. Компетентність з оцінювання, включена сюди, представляє собою знання та вміння, необхідні для виконання освітньої ролі вчителя. Зрозуміло, що існує багато інших компетентностей, окрім компетентності з оцінювання, якою повинен володіти вчитель. Установлюючи Стандарти компетентності вчителів з оцінювання учнів, асоціації поділяють ту точку зору, коли оцінювання учнів є важливою частиною викладання, а якісне викладання не може існувати без якісного оцінювання учнів. Тренінг з розвитку компетентностей, описаних у Стандартах, повинен бути невід'ємною частиною допрактичної підготовки. Більше того, такі тренінги з оцінювання мають бути широкодоступні практикуючим учителям через програми підвищення кваліфікації на рівні окремих шкіл і районів. Стандарти призначені для використання як:

- *рекомендації для викладачів, які розробляють та затверджують програми підготовки вчителів;*
- *рекомендації щодо самооцінювання вчителів з метою визначення потреби підвищення кваліфікації в оцінювання учнів;*
- *рекомендації щодо розробки завдань для викладачів, які проводять семінари з підвищення кваліфікації практикуючих учителів;*
- *мотивація фахівцям з освітньої діагностики та з підготовки вчителів для розробки ширшої концепції оцінювання учнів та підготовки вчителів у сфері оцінювання учнів, порівняно з попереднім періодом.*

Ці стандарти необхідно впровадити в майбутні програми підготовки та сертифікації вчителів. Учителям, які отримали підготовку, що не відповідає вказаним стандартам, необхідно надати можливість та підтримку для розвитку цих компетентностей до того, як названі стандарти застосовуватимуться для оцінювання таких учителів.

Підхід, що використовувався для розробки стандартів

Члени асоціацій, які підтримували цей проект, є професійними педагогами, що займаються викладанням, освітою вчителів та оцінюванням учнів. Члени цих асоціацій стурбовані неадекватністю підготовки вчителів до оцінювання освітнього прогресу своїх учнів і тому намагалися ефективно підійти до вирішення цієї проблеми. Комітет, призначений асоціаціями, уперше зібрався у вересні 1987 р. і підтвердив свій намір визначити стандарти підготовки вчителів у сфері оцінювання учнів. Потім комітет провів огляд дослідницької літератури

з метою визначення потреби в оцінюванні учнів, поточного рівня підготовки вчителів у сфері оцінювання учнів, сфер діяльності вчителів, що вимагають компетентності у використанні оцінювання та поточного рівня компетентності вчителів у сфері оцінювання учнів.

Члени комітету, користуючись своїм колективним досвідом, сформували і переглянули перелік важливих компетентностей в оцінюванні. Проект цих компетентностей кілька разів переглядався комітетом перед тим, як його представили на громадський розгляд. Для підготовки остаточного документа використовувалися коментарі представників від кожної асоціації.

Усвідомлення професійної ролі (the scope of a teacher's professional role) та відповідальності вчителя при оцінюванні учнів

Цей документ містить сім стандартів. У зв'язку з нагальною потребою оновлення методики оцінювання безпосередньо в класі, в деяких стандартах увага зосереджується на компетентностях, обмежених роботою в класі. Оскільки зростає роль учителя у прийнятті рішень щодо освіти та її політики поза межами класу, в інших стандартах наголошується на компетентностях учителя, необхідних для прийняття рішень щодо оцінювання на рівні школи, району, штату та національному рівні.

Усвідомлення професійної ролі та відповідальності вчителя при оцінюванні учнів можна описати за допомогою таких видів діяльності. Вони вимагають від учителів компетентності з оцінювання учнів, а також потребують достатнього часу і ресурсів для їхнього професійного виконання.

Види діяльності, що мають місце до початку навчального процесу:

- (а) *розуміння культурного середовища, інтересів, навичок та здібностей учнів у тій мірі, в якій вони стосуються навчальних доменів і/або предметів;*
- (б) *розуміння мотивації учнів та їх зацікавленості в певному навчальному змісті;*
- (в) *відбір та формулювання навчальних результатів, очікуваних від учнів;*
- (г) *планування навчального процесу для окремих учнів чи груп.*

Види діяльності, що мають місце під час навчального процесу:

- (а) *моніторинг прогресу учнів у досягненні навчальних цілей;*
- (б) *визначення досягнень учнів та складнощів, з якими вони зустрічаються в процесі навчання та демонстрації досягнень;*

ДОДАТКИ

- (в) коригування навчального процесу;
- (г) представлення отриманих конкретних та правдивих схвальних відгуків і зворотного зв'язку;
- (г) мотивація учнів до навчання;
- (д) визначення ступеня досягнення учнем навчальних цілей.

Види діяльності, що мають місце після закінчення відповідного сегмента навчального процесу (наприклад, уроку, семестру, навчального року):

- (а) опис ступеня досягнення кожним учнем короткострокових і довгострокових навчальних цілей;
- (б) повідомлення учням, батькам чи опікунам про сильні і слабкі властивості учнів за результатами оцінювання;
- (в) документування результатів оцінювання та звітування з метою аналізу, оцінювання та прийняття рішень на рівні школи;
- (г) аналіз інформації, зібраної до та під час навчального процесу для розуміння реального прогресу кожного учня та для формування майбутніх навчальних планів;
- (г) оцінка ефективності навчального процесу;
- (д) оцінка ефективності навчального плану та використаних матеріалів.

Види діяльності, що пов'язані з прийняттям учителем рішень на рівні окремої школи та на рівні району:

- (а) робота в шкільному чи районному комітеті, який вивчає сильні та слабкі сторони школи чи району в розвитку учнів;
- (б) робота над розробкою чи відбором методів оцінювання для використання в окремій школі чи в районі;
- (в) оцінка навчального плану району;
- (г) інші споріднені види діяльності.

Види діяльності, пов'язані з роботою вчителя у більш великих освітніх громадах:

- (а) робота в комітеті штату над розробкою навчальних цілей та відповідних їм типів оцінювання;
- (б) участь у переглядах доцільності навчальних цілей та відповідно типів оцінювання на рівні району, штату та національному рівні;
- (в) інтерпретація результативності програм оцінювання учнів на рівні штату та національному рівні.

Кожен із наведених нижче стандартів є описом знань чи навичок з оцінювання, які повинен мати вчитель для своєї успішної роботи в п'яти щойно описаних сферах. Разом узяті стандарти вимагають від учителя демонстрації вмінь з відбору, розробки, застосування, використання, повідомлення та оцінки інформації про оцінювання учнів та щодо практик оцінювання учнів. Після кожного стандарту наводиться обґрунтування та типові поведінкові характеристики. Стандарти являють собою концептуальну основу чи «скелет», на базі якого можна формувати певні конкретні вміння. Навіть після публікації цих стандартів необхідно проводити роботу щодо введення їх у дію. Очікується, що досвід застосування цих стандартів приведе до їхнього вдосконалення та подальшого розвитку.

Текст Стандартів

1. Учителі повинні вміти вибирати методи оцінювання, доцільні для прийняття навчальних рішень.

Уміння вибирати відповідні, корисні, адміністративно зручні, технічно адекватні та справедливі методи оцінювання є основою для належного використання інформації та підтримки навчальних рішень. Учителі повинні бути добре ознайомлені з типами інформації, яку можна отримати за допомогою широкого спектра різних видів оцінювання, а також з їх перевагами та недоліками. Зокрема, вони мають бути обізнані з критеріями для оцінки та відбору методів оцінювання в світлі навчальних планів.

Учителі, які відповідають цьому стандарту, повинні мати такі теоретичні та практичні навички. Вони мають уміти використовувати поняття похибки оцінювання та валідності під час розробки чи вибору підходів до оцінювання учнів у класі. Вони повинні розуміти, як валідні результати оцінювання можуть використовуватися для підтримки таких видів викладацької діяльності, як надання належного зворотного зв'язку учням, діагностування навчальних потреб груп та окремих учнів, планування індивідуалізованих навчальних програм, мотивування учнів та оцінка навчальних процедур. Учителі мають розуміти, як невалідна інформація може вплинути на викладацькі рішення щодо учнів. Вони також повинні вміти використовувати та оцінювати доступні їм альтернативи оцінювання, беручи до уваги серед іншого культурне, соціальне, економічне та мовне середовище учнів. Учителі мають усвідомлювати, що різні оцінювальні підходи можуть бути несумісними з певними навчальними цілями і по-різному впливати на викладання.

ДОДАТКИ

Учителі повинні знати доцільність кожного оцінювального підходу для прийняття рішень щодо їхніх учнів. Більше того вчителі мають знати, де знайти інформацію про різні методи оцінювання. Оцінювання існує в різноманітних формах і включає: запитання та тести згідно з навчальним планом, стандартизовані тести з використанням критеріїв та норм, усне опитування, спонтанне та структуроване оцінювання успішності, портфоліо, презентації, покази, рейтингові шкали, письмові зразки, «паперові» тести, роботу в класі та вдома, взаємота самооцінювання, історію попередніх досягнень, спостереження, опитування, співбесіди, проекти, продукти/вироби та думку інших.

2. Учителі повинні вміти розробляти методи оцінювання, доцільні для прийняття навчальних рішень.

Хоча вчителі часто застосовують опубліковані чи інші зовнішні інструменти оцінювання, основна маса інформації з оцінювання, яку вони застосовують для прийняття рішень, надходить від розроблених та впроваджених ними методів. І справді, вимоги оцінювання в класі виходять далеко за межі готових інструментів.

Учителі, які відповідають цьому стандарту, повинні володіти такими теоретичними та практичними навичками: вони зобов'язані вміти планувати збір інформації для забезпечення прийняття рішень, знати та застосовувати відповідні принципи для розробки та використання у викладацькій роботі методів оцінювання, що дозволяють уникнути типових хиб з оцінювання учнів. Ці методи можуть включати деякі з переліку варіантів, наведених у кінці першого стандарту. Учитель зобов'язаний уміти відібрати методи, які відповідають меті його викладання.

Учителі, які відповідають цьому стандарту, повинні вміти використовувати дані учнів для аналізу якості кожного методу оцінювання, який вони використовують. Оскільки більшість учителів не мають контактів із фахівцями з оцінювання, вони повинні бути готовими проводити такий аналіз самостійно.

3. Учителі мають уміти застосовувати як зовнішні готові, так і власно розроблені методи оцінювання, розраховувати за ними тестовий бал та інтерпретувати результати.

Недостатньо, щоб учителі вміли лише вибирати та розробляти певні методи оцінювання; вони повинні також уміти відповідно їх застосовувати. Учителі зобов'язані вміти застосовувати різні методи оцінювання, розраховувати за ними тестовий бал і інтерпретувати результати. Учителі, які відповідають цьому стандарту, повинні мати такі теоретичні та практичні навички: уміти інтерпретувати отримані ними неформальні

та формальні результати оцінювання, включаючи успішність учнів у виконанні класних та домашніх завдань; уміти користуватися посібниками для розрахунку балів за творчі відповіді (essay questions) та проекти, використовувати шаблони для розрахунку балів для запитань з множинним вибором відповіді та шкали для визначення рівня успішності. Використання цих засобів повинне давати несуперечливі результати.

Учителі зобов'язані вміти застосовувати стандартизовані тести досягнень та інтерпретувати бали в їх часто вживаному представленні: процентний рейтинг (percentile ranks), процентний діапазон (percentile band scores), стандартні бали (standard scores) та їхні еквіваленти в оцінках (grade equivalents). Вони повинні мати концептуальне розуміння підсумкових показників (summary indexes), що часто вказуються разом із результатами оцінювання: вимір основної тенденції (measures of central tendency), дисперсія, взаємозв'язки, надійність та похибка вимірювання.

Учителі повинні вміти застосовувати ці поняття балів та підсумкових показників для кращого використання оцінювання, розробленого ними. Вони повинні вміти аналізувати результати оцінювання для визначення здібностей учнів та їхніх помилок. Якщо вони отримують суперечливі результати, то мають знайти пояснення такому неспівпадінню чи використати інші дані для подолання невизначеності, перш ніж приймати рішення. Вони повинні вміти використовувати методи оцінювання так, щоб заохочувати освітній розвиток учнів і уникати зайвого занепокоєння з їхнього боку.

4. Учителі повинні вміти використовувати результати оцінювання для прийняття рішень щодо окремих учнів, для планування навчального процесу, розробки навчального плану та поліпшення функціонування школи.

Результати оцінювання використовуються для прийняття освітніх рішень на декількох рівнях: на рівні класу — про учнів, на рівні громади — про школу та шкільний район (school district), загалом у суспільстві — про цілі та результати освіти. Учителі відіграють основну роль, беручи участь у прийнятті рішень на кожному з цих рівнів, і повинні вміти ефективно використовувати результати оцінювання.

Учителі, які відповідають цьому стандарту, мають володіти такими теоретичними та практичними навичками. Вони повинні вміти використовувати кумулятивну інформацію, отриману в результаті оцінювання, для організації дієвого навчального плану, що сприятиме освітньому розвитку учнів. Використовуючи результати оцінювання для планування та/або оцінки викладання і навчального плану, учи-

ДОДАТКИ

телі зобов'язані правильно інтерпретувати результати, уникаючи поширених помилок, таких, як обґрунтування рішень балами, що позбавлені тематичної валідності (curriculum validity). Вони повинні бути поінформованими про результати оцінювання локального рівня, регіонального рівня, рівня штату та національного рівня та про їх належне використання для поліпшення освіти на рівні окремого учня, класу, школи, району, штату та на національному рівні.

5. Учителі мають уміти розробляти валідні процедури визначення рівня учнів за допомогою оцінювання.

Визначення рівня учнів є важливою частиною професійної діяльності вчителів. Термін **визначення рівня** означає відображення рівня успішності учня і оцінку вчителем цієї успішності. Принципи використання оцінювання для отримання валідних показників рівня учнів відомі, і вчителі повинні їх застосовувати.

Учителі, які відповідають цьому стандарту, повинні володіти такими теоретичними та практичними навичками: уміти розробляти, впроваджувати та пояснювати процедуру визначення рівня учнів на основі оцінок, отриманих за виконання різних типів завдань — проектів, класних завдань, вікторин, тестів тощо. Вони повинні розуміти і пояснити, чому визначений ними рівень є раціональним, обґрунтованим та справедливим, визнаючи, що він відображає їхні пріоритети та судження. Учителі зобов'язані усвідомлювати та уникати неадекватних процедур визначення рівня, наприклад, використання їх з метою покарання. Вони повинні вміти оцінювати та змінювати свої методи визначення рівня з метою підвищення валідності висновків про досягнення учнів.

6. Учителі повинні вміти повідомляти результати оцінювання учням, батькам та іншим непрофесійним аудиторіям, а також іншим працівникам освіти.

Учителі зобов'язані постійно повідомляти результати оцінювання учням та їхнім батькам чи опікунам. Окрім цього, вони мають часто обговорювати результати оцінювання з іншими викладачами та різними непрофесійними аудиторіями. У разі неефективного повідомлення результати можуть використовуватися неналежним чином або не використовуватись взагалі. Для ефективного спілкування з іншими на предмет оцінювання учнів учителі повинні вміти використовувати термінологію оцінювання та пояснювати значення, обмеження і наслідки результатів оцінювання. Більше того вчителі можуть опинитися у становищі, коли їм потрібно буде

відстояти свої процедури оцінювання та їхню інтерпретацію або допомогти публіці інтерпретувати результати оцінювання належним чином.

Учителі, які відповідають цьому стандарту, повинні мати такі теоретичні та практичні навички: вони мають розуміти та вміти пояснити, як соціоекономічні, культурні, мовні та інші об'єктивні фактори можуть вплинути на оцінювання учнів. Учителі зобов'язані пояснити, що результати оцінювання не означають, що наявні фактори оточення обмежують кінцевий освітній розвиток учня. Вони мають пояснити учням та їхнім батькам чи опікунам, як сприяти освітньому прогресу учня. Учителі повинні розуміти та вміти пояснити важливість урахування похибки вимірювання під час використання оцінювання для прийняття рішень щодо окремих учнів. Учителі мають уміти пояснити обмеження різних неформальних та формальних методів оцінювання. Вони повинні надавати пояснення до друкованих звітів про оцінювання учнів на рівні класу, шкільного району, штату та на національному рівні.

7. Учителі повинні вміти визначати неетичні, незаконні та інші невідповідні методи оцінювання та способи використання інформації щодо оцінювання.

Справедливість, права всіх зацікавлених сторін та професійне етичне поводження повинні лежати в основі всіх видів діяльності, пов'язаних з оцінюванням учнів, — від початкового планування і збору інформації до інтерпретації, використання і повідомлення результатів. Учителі повинні добре усвідомлювати свою етичну та правову відповідальність в оцінюванні. Окрім цього, вони зобов'язані не допускати використання неналежних методів оцінювання іншими. Учителі повинні співпрацювати з широкими освітніми колами для встановлення рамок належної професійної поведінки в оцінюванні.

Учителі, які відповідають цьому стандарту, повинні володіти такими теоретичними та практичними навичками. Вони мають знати закони та прецеденти, що стосуються їхніх методів оцінювання на рівні школи, шкільного району, штату. Учителі повинні усвідомлювати, що різні процедури оцінювання можуть використовуватися неналежним чином, спричиняючи такі недопустимі наслідки, як згальблення учнів, порушення права учнів на конфіденційність, неналежне використання результатів стандартизованих тестів досягнень учнів для визначення ефективності викладання вчителів.

Стандарти компетентності адміністраторів у галузі освіти з оцінювання учнів

Розроблено:

*Американською асоціацією шкільних адміністраторів
Національною асоціацією директорів початкових шкіл
Національною асоціацією директорів середніх шкіл
Національною радою з вимірювань в освіті (NCME)⁵⁹
Цей матеріал [11] не застережений авторськими правами. Відтворення та розповсюдження вітаються (1997).*

Передісторія

Оцінювання учнів повсякчас використовується в школах, проте небагато з підготовчих додипломних (preservice) програм для фахівців у сфері освіти вимагають ґрунтовної роботи з оцінювання. Ці Стандарти мають на меті спрямувати додипломну (preservice) і післядипломну (in-service) підготовку освітніх адміністраторів та ініціювати створення майбутніх стандартів акредитації програм підготовки адміністраторів і стандартів сертифікації освітніх адміністраторів. Вони призначені як для працівників окремих шкіл (building-level personnel), так і центральних офісів (central office personnel), які займаються асистуванням, керівництвом та оцінюванням осіб, котрі безпосередньо навчають учнів.

Комітет⁶⁰, що сформулював Стандарти компетентності освітніх адміністраторів, уперше зібрався в 1991 р. Проект стандартів було пере-

⁵⁹ У 1990 р. Американська федерація вчителів, Національна рада з вимірювань в освіті (NCME) та Національна освітня асоціація опублікували *Стандарти компетентності вчителів з оцінювання учнів*. Об'єднаний комітет рекомендував ці стандарти як базові для додипломної (preservice) та післядипломної (in-service) підготовки вчителів. Комітет також рекомендував розробку стандартів для інших категорій фахівців у галузі освіти. Цей документ слугує доповненням до *Стандартів компетентності вчителя*.

⁶⁰ Комітет було призначено професійними асоціаціями, які співпрацюють між собою: Джеймс Імпара (James C. Impara) і Барбара Плейк (Barbara S. Plake) (Університет штату Небраска) спільно очолювали комітет та представляли Національну раду з вимірювань в освіті (NCME) разом із Джеком Мервіном (Jack C. Merwin) (Університет штату Мінесота) Річард Міллер (Richard Miller), Гейл Стівенс (Gail Stephens) і Гері Маркс (Gary Marx) представляли Американську асоціацію шкільних адміністраторів (AASA), Рон Арегладо (Ron Areglado) представляв Національну асоціацію директорів початкових шкіл (NAESP), а Джейм Кіф (James Keefe) і Пол Херсі (Paul Hersey) представляли Національну асоціацію директорів середніх шкіл (NASSP). Виконання проекту було б неможливе без фінансової підтримки Фонду Келлогга (W. K. Kellogg Foundation).

глянуто кілька разів, починаючи з простого опитування членів організацій-учасників⁶¹, через зворотний зв'язок, отриманий на форумі під час професійного засідання, та шляхом отримання відповідей на проект, що розповсюджувався для громадських коментарів.

У зазначених стандартах термін **оцінювання** означає процес отримання інформації про успішність учня для прийняття освітніх рішень щодо учнів; для інформування учнів, їхніх батьків, учителів чи інших зацікавлених осіб щодо їхнього прогресу, сильних чи слабких моментів; для визначення інструктивної ефективності та відповідності навчальному плану; а також для формування політики. **Методи чи техніка оцінювання** включають: формальні та неформальні спостереження успішності учнів; демонстрацію учнями навичок та знань; якісний аналіз успішності та результатів діяльності учнів; оцінювання на базі портфоліо, комерційно розроблених або розроблених учителем «паперових» (paper-and-pencil) тестів; усне опитування; комп'ютерні чи адаптивні тести або аналіз прогресу учня, проте не обмежуються ними.

Загальний огляд стандартів

Ці стандарти представляють компетентності оцінювання в тому адміністративному контексті, в якому оцінювання є доцільним для освітніх адміністраторів. Цей контекст включає:

- *допомогу вчителям у створенні та ефективному використанні оцінювання;*
- *забезпечення керівництва у створенні та впровадженні політики оцінювання на рівні окремих шкіл та районів;*
- *використання адміністраторами результатів оцінювання для прийняття рішень щодо учнів, учителів, навчання загалом та повідомлення результатів оцінювання різним зацікавленим сторонам і спільнотам.*

Якщо освітні адміністратори діють у контексті **допомоги вчителям** у створенні та використанні оцінювання, вони повинні мати такий самий рівень компетентності оцінювання учнів, як і вчителі. Виходячи з цієї ролі 1-й стандарт стосується огляду **Стандартів компетентності вчителів з оцінювання учнів** [10]. Освітні адміністратори повинні мати компетентності, описані у стандартах для вчителів, відповідно до рівня прямої взаємодії з учителями, керування ними та їхнього оцінювання.

⁶¹ Більше інформації про це опитування можна знайти в: Impara, J. C., Plake, B. S., & Merwin, J. C. (1994). Student assessment tasks and knowledge: Comparing superintendents and elementary and secondary principals. *Journal of School Leadership*, 4 (5), 517-528.

ДОДАТКИ

Адміністратори повинні володіти цими компетентностями як мінімум на рівні, достатньому для роботи. Такий рівень вимагається з двох причин. **По-перше**, адміністратори повинні розуміти, навіщо вчителям потрібні ці компетентності; **по-друге**, у певних випадках адміністратори зобов'язані спостерігати, визначати, реєструвати, відстежувати та оцінювати їх у учителів. Ці компетентності зазначені в Стандартах 1 та 2.

Коли освітні адміністратори **забезпечують керівництво** у створенні та розробці політики оцінювання, їм потрібен дещо інший набір знань, умінь та здібностей, ніж учителям. Оскільки сфера оцінювання зазнає швидких змін, адміністратори мають бути здатними оцінити раціональність нових технік оцінювання, що пропонуються в їхніх школах та районах. Вони повинні визначити доцільність запропонованого використання та інтерпретації як традиційних, так і нових стратегій оцінювання. Стандарти 3, 4 і 5 описують базові навички оцінювання учнів: належні принципи оцінювання, використання оцінювання та термінологія оцінювання.

Освітні адміністратори повинні володіти низкою компетентностей, якщо вони **використовують та повідомляють** результати оцінювання для прийняття рішень. Адміністратори часто повідомляють результати оцінювання іншим і використовують результати оцінювання для прийняття рішень щодо окремих учнів, груп учнів, учителів та іншого персоналу, навчального плану і освітніх програм. Решта Стандартів (з 6 по 12) стосуються різноманітних контекстів, у яких адміністратори виступають основними користувачами результатів оцінювання.

ТЕКСТ Стандартів

Компетентності, пов'язані з допомогою вчителям:

1. Мати достатній для роботи рівень компетентності, що відповідає **Стандартам компетентності вчителів з оцінювання учнів**.

Ці стандарти включають таке:

- а) вибір методів оцінювання, які можуть бути застосованими при прийнятті рішень щодо викладання;
- б) розробку методів оцінювання, які можуть бути застосованими при прийнятті рішень щодо викладання;
- в) адміністрування, розрахунок тестового бала та інтерпретацію результатів як зовнішніх, так і розроблених учителем методів оцінювання;
- г) використання результатів оцінювання під час прийняття рішень щодо окремих учнів, викладання, планування, розробки навчальних планів, поліпшення роботи школи;

- г) розробку валідних процедур визначення рівня успішності учнів (*pupil grading procedures*), що використовуються при оцінюванні учнів;
 - д) повідомлення результатів оцінювання учням, батькам та іншим непрофесійним аудиторіям, а також іншим працівникам освіти;
 - е) визначення неетичних, незаконних та інших невідповідних методів оцінювання та способів використання інформації щодо оцінювання.
2. Володіти необхідними навиками щодо створення різних засобів оцінювання.

Локально створені засоби оцінювання мають широке застосування. Учителі створюють щоденні, щотижневі та семестрові засоби оцінювання для своїх класів. Адміністратори, яким надано керівну функцію, повинні вміти визначити якість процедур оцінювання, що використовуються їхнім персоналом для прийняття рішень щодо учнів. Поширення методів оцінювання досягнень успішності (*performance assessment*), оцінювання на базі портфоліо та комп'ютерного оцінювання зобов'язує адміністраторів критично підходити до розробки та використання цих підходів.

Компетентності, пов'язані із забезпеченням керівництва у розробці та впровадженні політики оцінювання.

3. Розуміти та вміти застосовувати базові принципи вимірювання до оцінювання, що проводиться в умовах школи.
- Адміністратори повинні усвідомлювати, як ці принципи застосовуються в умовах школи, щоб, виконуючи свою керівну роль, підтримувати належне використання та інтерпретацію результатів оцінювання. Вони зобов'язані розуміти відмінності між інтерпретацією тестів на базі критеріїв та на базі норм, повинні вміти оцінювати доцільність цієї інтерпретації в певному контексті прийняття рішень. Адміністратори мають розуміти, що валідність визначається інтерпретацією, яка робиться на основі тестового балу, а не самим тестовим балом, повинні бути готові до виявлення лідерства як у підтримці належного використання, так і у припиненні спроб неналежного використання результатів оцінювання.
4. Розуміти цілі (наприклад, опис, діагностика, визначення успішності) різних видів оцінювання (наприклад, успішності, здібності, ставлення) та доцільні для отримання даних, що відповідають обраним цілям, стратегії оцінювання.

ДОДАТКИ

Цей стандарт є продовженням Стандартів 1 і 2 в **Стандартах компетентності вчителів з оцінювання учнів** [10], які зосереджуються лише на прийнятті рішень, що стосуються викладання. Більшість технік оцінювання придатні лише для однієї певної мети. Через це рішення повинні базуватися на результатах оцінювання, що відповідають меті певної техніки оцінювання. Адміністратори несуть відповідальність за належне використання результатів оцінювання в кожному окремому контексті прийняття рішень.

5. Розуміти необхідність чіткої та послідовної політики оцінювання учнів на рівні окремої школи та на рівні району.

Це є продовженням Стандарту 5 із **Стандартів компетентності з оцінювання учнів для вчителів**, що зосереджуються лише на визначенні рівня успішності учнів (grading). Політики на рівні окремих шкіл та на рівні районів повинні включати визначення рівня учнів, випускні/перевідні екзамени та інші аспекти оцінювання. Робота адміністраторів має гарантувати, що вчителі розуміють інструкції, згідно з якими повинне відбуватися присвоєння оцінок та визначення рівня учнів, і що вони керуються загальними стандартами якості роботи, що претендує на певну оцінку. Аналогічно, адміністратори повинні надати можливість учителям та іншим особам зрозуміти спосіб, у який випускні/перевідні екзамени та інші види оцінювання використовуються для прийняття рішень щодо учнів.

Компетентності, необхідні для використання оцінювання з метою прийняття рішень та для повідомлення результатів оцінювання.

6. Розуміти та пояснювати іншим технічні поняття і термінологію з оцінювання у нетехнічний, але правильний спосіб.

Адміністратори повинні розуміти технічні поняття, щоб вирішувати, які типи оцінювання використовувати і як представляти результати оцінювання іншим. Коли адміністратори інтерпретують результати оцінювання для різних аудиторій та зацікавлених сторін (наприклад, для вчителів, інших адміністраторів, батьків, шкільного керівництва, засобів масової інформації), вони повинні вміти пояснювати технічні поняття у нетехнічний, але правильний спосіб. Більше того, прийняття рішення щодо вибору стандартизованого тесту чи іншої процедури оцінювання вимагає ознайомлення з посібниками, технічними звітами чи іншою технічною літературою з оцінювання, що містить технічні поняття.

7. Розуміти та дотримуватися етичних та технічних рекомендацій щодо оцінювання.

Оскільки результати оцінювання можуть використовуватися для прийняття важливих, а інколи майже незворотних рішень стосовно окремих учнів, персоналу школи та освітніх програм, інтерпретація результатів повинна бути валідною і відповідати технічно обґрунтованим принципам оцінювання та належній освітній практиці. Адміністратори мають бути ознайомлені з чинними рекомендаціями та стандартами, які описують належну розробку та використання оцінювання і повідомлення його результатів.

Основні документи щодо освітнього оцінювання (рекомендації та стандарти), що стосуються оцінювання адміністраторів, які використовувалися у 1995 р. під час розробки цих Стандартів компетентності, — такі:

Стандарти компетентності вчителів з оцінювання учнів [10]. (1990, м. Вашингтон). Американська федерація вчителів, Національна рада з вимірювань в освіті (NCME) та Національна освітня асоціація.

Стандарти освітнього та психологічного тестування. (1985, м. Вашингтон. Зараз переглядаються⁶²). Американська психологічна асоціація, Американська асоціація освітніх досліджень та Національна рада з вимірювань в освіті.

Кодекс справедливого тестування в освіті [4]. (1988, м. Вашингтон). Об'єднаний комітет з практики тестування (JCTP). Американська психологічна асоціація.

8. Узгоджувати суперечливі результати оцінювання належним чином.
Інколи результати оцінювання здаються суперечливими і їх необхідно розв'язати для прийняття належних рішень. Наприклад, результати оцінювання за навчальним планом можуть показувати високу успішність із читання, проте бали за стандартизованим тестом із читання можуть бути низькими; оцінювання шкільного середовища може показувати позитивне сприйняття школи учнями, проте наявні численні скарги батьків щодо безпеки. Адміністратори повинні вміти визначати очевидні суперечності в результатах оцінювання в світлі контекстуальних факторів і надавати їм належну інтерпретацію.
9. Визначати важливість, доцільність та складність інтерпретації результатів оцінювання в світлі мовного і культурного середовища учнів та інших позашкільних факторів з метою врахування індивідуальних особливостей, включаючи обмежені можливості, для забезпечення валідності результатів оцінювання для всіх учнів.

⁶² Це стосується редакції 1985 р., а не редакції 1999 р. [1].

ДОДАТКИ

На результати оцінювання можуть впливати ряд соціальних, культурних та інших факторів. Ці фактори можуть прямо спричиняти або не спричиняти високу чи низьку успішність, проте їхнє знання часто допомагає інтерпретувати успішність у виконанні оцінювальних завдань. Адміністратори повинні усвідомлювати, створення яких спеціальних умов для проведення оцінювання може бути доцільне чи недоцільне для різних учнів за різних обставин.

10. Забезпечувати належне застосування оцінювальних та інформаційних технологій для проведення оцінювання учнів.

Технології включають такі інструменти комп'ютерного оцінювання: комп'ютеризоване адаптивне тестування, комп'ютерне тестування, комп'ютеризовані банки тестових завдань, комп'ютеризовані архіви даних оцінювання та комп'ютеризовані бази даних. Оскільки технології стають доступними для все більшого числа шкіл, адміністратори повинні будуть приймати рішення щодо використання таких технологій для проведення оцінювання учнів та для зберігання, систематизації, використання та безпечного збереження отриманих результатів.

11. Доцільно використовувати наявні технології для інтегрування результатів оцінювання та інших даних з метою поліпшення навчання, викладання та успішності учнів.

Адміністратори повинні вміти систематизовувати широкий спектр інформації щодо оцінювання учнів та інші учнівські дані для отримання висновків про якість успішності учнів, викладацької діяльності їхніх учителів та функціонування шкільних програм. Адміністратори мають визначати доцільність чи недоцільність використання результатів оцінювання в цьому контексті.

12. Вносити судження в межах своєї юрисдикції про якість стратегії чи програми оцінювання, що використовується для прийняття рішень.

Оцінювання вчителів в класі, внутрішньошкільні програми тестування, системне тестування та державне оцінювання потенційно можуть впливати на функціонування школи. Адміністратори повинні допомагати вчителям та іншим особам у прийнятті поінформованих рішень про стратегії оцінювання. Адміністратори зобов'язані проводити чи координувати системні дослідження, що дають змогу правильно формувати висновки щодо якості наявних чи запропонованих програм оцінювання. Зміна технології процедур оцінювання зумовлює критичне ставлення до них адміністраторів. Наприклад, адміністратори повинні запроваджувати та контролювати процедури оцінювання для того, щоб оцінюван-

ня успішності та оцінювання на базі портфоліо проводилися у технічно надійний та освітньо доцільний спосіб.

Загальна структура Стандартів Компетентності, пов'язані із допомогою вчителям:

1. Мати достатній для роботи рівень компетентності за Стандартами компетентності з оцінювання учнів для вчителів.
2. Володіти необхідними навиками у створенні різних засобів оцінювання.

Компетентності, пов'язані із забезпеченням керівництва у розробці та впровадженні політики оцінювання:

3. Розуміти та вміти застосовувати базові принципи вимірювання до оцінювання, що проводиться в умовах школи.
4. Розуміти цілі (наприклад, опис, діагностика, визначення рівня учнів) різних видів оцінювання (наприклад, успішності, здібностей, ставлення) та доцільні для отримання даних, що відповідають обраним цілям, стратегії оцінювання.
5. Розуміти необхідність чіткої та послідовної політики оцінювання учнів на рівні окремої школи та на рівні району.

Компетентності, необхідні для використання оцінювання з метою прийняття рішень та для повідомлення результатів оцінювання.

6. Розуміти та пояснювати іншим технічні поняття з оцінювання і термінологію у нетехнічний, але правильний спосіб.
7. Розуміти та дотримуватися етичних та технічних рекомендацій щодо оцінювання.
8. Узгоджувати суперечливі результати оцінювання належним чином.
9. Визначати важливість, доцільність та складність інтерпретації результатів оцінювання в світлі мовного і культурного середовища учнів та інших позашкільних факторів з метою врахування індивідуальних особливостей, включаючи обмежені можливості, для забезпечення валідності результатів оцінювання для всіх учнів.
10. Забезпечувати належне застосування оцінювальних та інформаційних технологій для проведення оцінювання учнів.
11. Доцільно використовувати наявні технології з метою інтегрування результатів оцінювання та інших даних для поліпшення навчання, викладання та успішності учнів.
12. Робити висновки в межах своєї юрисдикції про якість стратегії чи програми оцінювання, що використовується для прийняття рішень.

С. АНГЛОМОВНІ ВАРІАНТИ СТАНДАРТИВ



INTERNATIONAL TEST COMMISSION

ITC Guidelines on Adapting Tests

ITC Guidelines on Adapting Tests [6] were developed by a 13-person committee representing a number of international organizations. The objective was to produce a detailed set of guidelines for adapting psychological and educational tests for use in various different linguistic and cultural contexts (Van de Vijver & Hambleton, 1996). This is an area of major importance as tests become used in more and more countries, and as tests developed in one country get translated or adapted for use in another. Adaptation needs to consider the whole cultural context within which a test is to be used. Indeed, the adaptation guidelines apply wherever tests are moved from one cultural setting to another — regardless of whether there is a need for translation. Hambleton (1994) describes the project in detail and outlines the 22 guidelines that have emerged from it. These guidelines fall into four main categories: those concerned with the cultural context, those concerned with the technicalities of instrument development and adaptation, those concerned with test administration, and those concerned with documentation and interpretation. All but the second of these also have direct implications for test use and for test users.

In 2006, the ITC hosted a conference on Test adaptation across language and cultures in Brussels.

This conference provided information, guidance and support for updating the ITC Guidelines on Adapting Tests. For more details see the conference web site — http://www.intestcom.org/itc_conference_2006.htm.

«Дозвіл на переклад рекомендацій українською мовою»

**PRESIDENT**

Prof. Jacques Gregoire
Universite Catholique de Louvain
Faculte de Psychologie
Place du Cardinal Mercier, 10
1348 Louvain-la-Neuve
Belgium

PRESIDENT-ELECT

Dr. Marise Ph. Born
Erasmus University of Rotterdam
Institute of Psychology
Woudestein, Room T15-13
3000 DR Rotterdam
The Netherlands

PAST-PRESIDENT

Prof. Jose Muniz
Universidad de Oviedo
Facultad de Psicologia
Plaza de Feijoo, s/n
3300 Oviedo
Spain

SECRETARY

Prof. Ronald K. Hambleton
University of Massachusetts
152 Hills South
Amherst, MA 01003
U.S.A.

TREASURER

Prof. Em. Barbara M. Byrne
School of Psychology
University of Ottawa
Ottawa, Ontario
Canada K1N6N5

COUNCIL MEMBERS

Prof. Fanny M. Cheung
Chinese University of Hong Kong,
Hong Kong
Dr. Iain Coyne University of Nottingham, U.K.
Dr. David Foster Kryterion, U.S.A.
Prof. Cheryl Foxcroft
Nelson Mandela Metropolitan University,
South Africa
Prof. Frederick Leong Michigan State
University, U.S.A.
Prof. Thomas Oakland University of
Florida, U.S.A.

REPRESENTATIVES

Dr. Merry Bullock, U.S.A.
International Union of Psychological Science
Prof. Dave Bartram, U.K.
International Association of Applied
Psychology

EDITORS

International Journal of Testing
Prof. John Hattie
University of Auckland
The Hub, N452
Epsom Campus
Private Bag 92019
Auckland
New Zealand
Testing International
Dr. Jan Bogg
University of Liverpool
Department of Clinical Psychology
The Whelan Building
Liverpool L69 3GB
U.K.

INTERNATIONAL TEST COMMISSION

Brussels, October 30, 2006

Mrs. Marina Mrouga
Deputy Director, Testing Board
22 Pushkinska St., suite 307
Kiev 01601
Ukraine

Dear Colleague,

We examined your request to translate the ITC standards («ITC Guidelines for Test Use» and «ITC CBT/Internet Testing Guidelines» into Ukrainian.

On the condition that you certify that you will not changed or amended these guidelines in any way, and that you will always present them as the ITC guidelines, I am pleased to give you the official authorisation to translate the ITC guidelines into Ukrainian.

Sincerely,

Jacques Grégoire
ITC President

ДОДАТКИ

«Дозвіл на переклад рекомендацій російською мовою»

**PRESIDENT**

Prof. Jacques Gregoire
Universite Catholique de Louvain
Faculte de Psychologie
Place du Cardinal Mercier, 10
1348 Louvain-la-Neuve
Belgium

PRESIDENT-ELECT

Dr. Marise Ph. Born
Erasmus University of Rotterdam
Institute of Psychology
Woudestein, Room T15-13
3000 DR Rotterdam
The Netherlands

PAST-PRESIDENT

Prof. Jose Muniz
Universidad de Oviedo
Facultad de Psicologia
Plaza de Feijoo, s/n
3300 Oviedo
Spain

SECRETARY

Prof. Ronald K. Hambleton
University of Massachusetts
152 Hills South
Amherst, MA 01003
U.S.A.

TREASURER

Prof. Em. Barbara M. Byrne
School of Psychology
University of Ottawa
Ottawa, Ontario
Canada K1N6N5

COUNCIL MEMBERS

Prof. Fanny M. Cheung
Chinese University of Hong Kong,
Hong Kong
Dr. Iain Coyne University of Nottingham, U.K.
Dr. David Foster Kryterion, U.S.A.
Prof. Cheryl Foxcroft
Nelson Mandela Metropolitan University,
South Africa
Prof. Frederick Leong Michigan State
University, U.S.A.
Prof. Thomas Oakland University of
Florida, U.S.A.

REPRESENTATIVES

Dr. Merry Bullock, U.S.A.
International Union of Psychological Science
Prof. Dave Bartram, U.K.
International Association of Applied
Psychology

EDITORS

International Journal of Testing
Prof. John Hattie
University of Auckland
The Hub, N452
Epsom Campus
Private Bag 92019
Auckland
New Zealand
Testing International
Dr. Jan Bogg
University of Liverpool
Department of Clinical Psychology
The Whelan Building
Liverpool L69 3GB
U.K.

INTERNATIONAL TEST COMMISSION

Brussels, October 30, 2006

Mrs. Marina Mrouga

Deputy Director, Testing Board

22 Pushkinska St., suite 307

Kiev 01601

Ukraine

Dear Colleague,

We examined your request to translate the ITC standards ("ITC Guidelines for Test Use" and "ITC CBT/Internet Testing Guidelines" into Russian.

On the condition that you certify that you will not changed or amended these guidelines in any way, and that you will always present them as the ITC guidelines, I am pleased to give you the official authorisation to translate the ITC guidelines into Russian.

Sincerely,

Jacques Grégoire

ITC President

ITC International Guidelines for Test Use

Version 2000

The Council of the International Test Commission (ITC) formally adopted the Guidelines [5] at its June 1999 meeting in Graz, Austria.

The European Federation of Professional Psychologists Associations' Task Force on Tests and Testing also endorsed the Guidelines at its July 1999 meeting in Rome.

The Guidelines were officially published at the General Meeting of the ITC on 24 July, 2000 in Stockholm. The Guidelines are copyright of the ITC, 1999. The ITC is a non-stock corporation incorporated in the USA.

For further information on the ITC or for permission to quote from or reproduce the contents of this document, please contact the Secretary of the ITC:

Prof Jacques Grégoire, ITC Secretary,
Université catholique de Louvain,
Faculté de Psychologie,
Place du Cardinal Mercier, 10,
1348 Louvain-la-Neuve,
Belgium.

Email: gregoire@deva.ucl.ac.be

For further information regarding the Guidelines, contact:

Prof Dave Bartram, ITC President,
SHL Group plc, The Pavilion, 1 Atwell Place,
Thames Ditton, KT7 0NE, Surrey, England.

Email: dave.bartram@shlgroup.com

See also the ITC website (from which copies of the Guidelines can be obtained): <http://www.intestcom.org>

Acknowledgements

The Guidelines were prepared for the ITC Council by Professor Dave Bartram. The author is grateful for the assistance provided by Iain Coyne in the execution of this project and is grateful to the following individuals who took part in the 1997 Dublin workshop and who provided such valuable input to the development of the present Guidelines.

Ms Dusica Boben, Produktivnost, SLOVENIA;

Mr Eugene Burke, British Psychological Society, ENGLAND;
Dr Wayne Camara, The College Board, USA;
Mr Jean-Louis Chabot, ANOP, FRANCE;
Mr Iain Coyne, University of Hull, ENGLAND;
Dr Riet Dekker, Swets and Zeitlinger, NETHERLANDS;
Dr Lorraine Eyde, US Office of Personnel Management, USA;
Prof Rocio Fernandez-Ballesteros, EAPA, SPAIN;
Mr Ian Florance, NFER-NELSON, ENGLAND;
Prof Cheryl Foxcroft, Test Commission of South Africa, SOUTH AFRICA;
Dr John Fremer, The College Board, USA;
Ms Kathia Glabeke, Commissie Psychodiagnostiek, BELGIUM;
Prof Ron Hambleton, University of Massachusetts at Amherst, USA;
Dr Karin Havenga, Test Commission of South Africa, SOUTH AFRICA;
Dr Jürgen Hogrefe, Hogrefe & Huber Verlagsgruppe, GERMANY;
Mr Ralf Horn, Swets and Zeitlinger, GERMANY;
Mr Leif Ter Laak, Saville and Holdsworth Ltd, ENGLAND;
Dr Pat Lindley, British Psychological Society, ENGLAND;
Mr Reginald Lombard, Test Commission of South Africa, SOUTH AFRICA;
Prof Jose Muniz, Spanish Psychological Association, SPAIN;
Ms Gill Nyfield, Saville & Holdsworth Ltd, ENGLAND;
Dr Torleiv Odland, Norsk Psykologforening, NORWAY;
Ms Berit Sander, Danish Psychologists' Association, DENMARK;
Prof Francois Stoll, Federation Suisse des Psychologues, SWITZERLAND.
The author is also grateful to the many other individuals and organisations who provided feedback during the various stages of consultation and in conference presentations.

Introduction and background

The need for international Guidelines

The focus of the International Test Commission (ITC) project is on guidelines for good test use and for encouraging best practice in assessment. The work so far carried out by the ITC to promote good practice in test adaptations (Hambleton, 1994; Van de Vijver, F. & Hambleton, R., 1996) is an important step towards assuring uniformity in the quality of tests adapted for use across different cultures and languages. At its meeting

in Athens in 1995, the ITC Council accepted a proposal to broaden this concern to include guidelines on the fair and ethical use of tests, from which standards for training and specifying the competence of test users could be derived.

There are a number of reasons why guidelines on test use are needed at an international level.

Countries differ greatly in the degree, if any, of statutory control they can exercise over the use of testing and its consequences for those tested. Some national professional societies have statutory registration of psychologists, others do not; some have mechanisms for the control of standards of test use by non-psychologists, others do not. The existence of a set of internationally-accepted guidelines would provide national psychological associations and other relevant professional bodies and organisations with a degree of support in the endeavours of such organisations to develop standards in countries where such standards are currently either lacking in some respect or non-existent.

Patterns of access, in terms of the rights to purchase or use test materials, vary greatly from country to country. In some countries, access is restricted to psychologists, in others to users registered with formally approved national test distributors, in yet others, test users may be free to obtain materials without restriction from suppliers in their country or directly from suppliers abroad.

A number of well-known instruments have appeared on the Internet in violation of copyright, without acknowledgement of the test authors or publishers, and without regard to issues of test security.

Within the occupational testing arena, the greater international mobility of labour has increased the demand for tests to be used on job applicants from a number of different countries — often with the tests being administered in one country on behalf of a potential employer in another country.

Development work is being carried out in the USA and in the UK on the use of Internet for distance- or remote-assessment in both occupational and educational settings. This raises a whole host of issues relating to standards of administration and control over the testing process, including test security.

Aim and objectives

The long-term aim of this project includes the production of a set of guidelines that relate to the competencies (knowledge, skills, abilities and other personal characteristics) needed by test users. These competencies are specified in terms of assessable performance criteria. These criteria provide the basis for developing specifications of the evidence of

competence that would be expected from someone seeking qualification as a test user. Such competencies need to cover such issues as:

- *professional and ethical standards in testing,*
- *rights of the test taker and other parties involved in the testing process,*
- *choice and evaluation of alternative tests,*
- *test administration, scoring and interpretation,*
- *report writing and feedback.*

Insofar as they directly relate to test use, the Guidelines also have implications for:

- *standards for test construction,*
- *standards for user-documentation — e.g., technical and user manuals,*
- *standards for regulating the supply and availability of tests and information about tests.*

The present Guidelines represent the work of specialists in psychological and educational testing (i.e. psychologists, psychometricians, test publishers and test developers) drawn from a number of countries. The intention of this document is not to 'invent' new guidelines, but to draw together the common threads that run through existing guidelines, codes of practice, standards and other relevant documents, and to create a coherent structure within which they can be understood and used.

Development of the Guidelines

The Guidelines should be considered as benchmarks against which existing local standards can be compared for coverage and international consistency. By using the Guidelines as benchmarks or the basis from which to develop locally applicable documents (e.g. standards, codes of practice, statements on test taker rights), a high level of consistency across national boundaries will be promoted.

Work on the Guidelines began by drawing together materials concerned with test standards, codes of practice, test use, etc., from a number of countries⁶³ While drawing on all of these sources, the present Guidelines have been particularly influenced by:

The Australian Psychological Society (APS) *Supplement to guidelines on the use of Psychological Tests* (Kendall et al., 1997).

The British Psychological Society (BPS) *Level A and Level B standards for occupational test use* (Bartram, 1995, 1996).

⁶³ A list of all the materials that informed this process is available on request from the authors.

The American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME) (1985) *Standards for educational and psychological testing*.

American Association for Counseling and Development (AACD) *Responsibilities of Users of Standardized Tests* (Schafer, W.D, 1992).

The CPA (Canadian Psychological Association, 1987) *Guidelines for Educational and Psychological Testing*.

The APS document has been particularly valuable as it pulls together much of what is contained in the BPS and American publications as well as drawing on South African National Institute for Psychological Research (NIPR) publications and various publishers' guidance for test users. It also embodies much of what has come out of Joint Committee on Testing Practices (JCTP) Test User Qualifications Working Group's (TUQWG) seminal work on a data-based approach to promoting good test use (e.g., Eyde et al, 1988, 1993; Moreland et al., 1995), and the work of the JCTP on the ***Code of Fair Testing Practices in Education*** (JCTP, 1988; Fremer, Diamond, & Camara, 1989). Appendix B drew from the more recent work of the JCTP (JCTP, 2000) on test taker rights and responsibilities.

The content of the primary sources was analysed and statements categorised under 14 main headings. Where appropriate single statements were written to capture the common meaning of a number of statements from different sources. Statements were also modified in format such that they provided completion of a common stem (e.g.: «Competent test users will endeavour to....», or «Competent test users can....»).

This initial structure of 14 main sections and their content was embodied in the first draft Framework Document. This formed the material for an international workshop held in Dublin in July 1997. The purpose of the ITC Workshop was to consider and critically evaluate all aspects of a framework document, with a view to producing a draft set of guidelines that would have international currency and acceptance. During the workshop, the Framework Document was examined in detail, with refinements being proposed in terms of form, structure and content. Following the workshop, the document was extensively revised (Version 2.0) and circulated to all those who attended for comment. A draft consultation document (Version 3.1) was prepared that incorporated all the comments and suggestions submitted regarding Version 2.0.

Copies of the Version 3.1 consultation document and a structured response form were widely circulated to key individuals and organisations

for comment. A total of 200 were distributed. A total of 28 detailed responses were received including 'corporate' responses from the APA, the BPS and some other European professional associations. In the summer of 1998 the Guidelines were revised in the light of these comments, and 200 copies (Version 4.1) were sent out for further consultation. A total of 18 formal responses were received to this second round of consultation. In addition, supportive informal comments were provided by many recipients of the consultation documents by email or in meetings.

In producing the current version of the Guidelines (Version 2000), every effort has been made to take account of all these responses. Without exception, the responses were helpful and constructive.⁶⁴

These Guidelines are to be seen as supportive rather than constraining. We need to ensure that the Guidelines embody universal key principles of good test use, without attempting to impose uniformity on legitimate differences in function and practice between countries or between areas of application.

The proposed structure differentiates three main aspects of competence:

Professional and ethical standards of good practice that affect the way in which the process of testing is carried out and the way in which test users interact with others involved in the process.

The knowledge, understanding and skills relating to the process of testing: what test users need to be able to do.

The knowledge and understanding that are necessary to inform and underpin the process of testing.

These three components differ, yet are inextricably inter-twined in practice.

The Guidelines start from a **key purpose**. This can be characterised as the "mission statement" for test users. It provides the focus from which the guidelines are developed. Each guideline defines an aspect of test user competence that contributes to the key purpose.

Together with the key purpose, the **scope statement** describes to whom the Guidelines apply, the forms of assessment to which they relate, and the assessment contexts.

This document contains:

Key purpose and scope statements.

⁶⁴ A detailed report on the results of the first consultation was submitted to the ITC Council meeting in August, 1998. A report on the second consultation together with Version 5.0 of the Guidelines was submitted to the ITC Council when it met in June 1999. Version 2000 contains some minor editorial amendments to Version 5.0.

Specifications of test user competencies in relation to ethical test use.
Specifications of test user competencies in relation to good practice in the use of tests.

The Guidelines

Key purpose

A competent test user will use tests appropriately, professionally, and in an ethical manner, paying due regard to the needs and rights of those involved in the testing process, the reasons for testing, and the broader context in which the testing takes place.

This outcome will be achieved by ensuring that the test user has the necessary competencies to carry out the testing process, and the knowledge and understanding of tests and test use that inform and underpin this process.

Scope of the Guidelines

Any attempt to provide a precise definition of a 'test' or of 'testing' as a process, is likely to fail as it will tend to exclude some procedures that should be included and include others that should be excluded. For the purpose of these Guidelines, the terms 'test' and 'testing' should be interpreted broadly. Whether an assessment procedure is labelled a 'test' or not is immaterial. These Guidelines will be relevant for many assessment procedures that are not called 'tests' or that seek to avoid the designation 'test'. Rather than provide a single definition, the following statements attempt to map out the domain covered by the Guidelines.

Testing includes a wide range of procedures for use in psychological, occupational and educational assessment.

Testing may include procedures for the measurement of both normal and abnormal or dysfunctional behaviours.

Testing procedures are normally designed to be administered under carefully controlled or standardised conditions that embody systematic scoring protocols.

These procedures provide measures of performance and involve the drawing of inferences from samples of behaviour.

They also include procedures that may result in the qualitative classification or ordering of people (e.g., in terms of type).

Any procedure used for 'testing', in the above sense, should be regarded as a 'test', regardless of its mode of administration; regardless of whether it was developed by a professional test developer; and regardless of whether it involves sets of questions, or requires the performance of

tasks or operations (e.g., work samples, psycho-motor tracking tests).

Tests should be supported by evidence of reliability and validity for their intended purpose. Evidence should be provided to support the inferences that may be drawn from the scores on the test. This evidence should be accessible to the test user and available for independent scrutiny and evaluation. Where important evidence is contained in technical reports that are difficult to access, fully referenced synopses should be provided by the test distributor.

The test use Guidelines presented here should be considered as applying to all such procedures, whether or not they are labelled as 'psychological tests' or 'educational tests' and whether or not they are adequately supported by accessible technical evidence.

Many of these Guidelines will apply also to other assessment procedures that lie outside the domain of 'tests'. They may be relevant for any assessment procedure that is used in situations where the assessment of people has a serious and meaningful intent and which, if misused, may result in personal loss or psychological distress (for example, job selection interviews, job performance appraisals, diagnostic assessment of learning support needs).

The Guidelines do not apply to the use of materials that may have a superficial resemblance to tests, but which all participants recognise are intended to be used only for purposes of amusement or entertainment (e.g., life-style inventories in magazines or newspapers).

Who the Guidelines are for

The Guidelines apply to the use of tests in professional practice. As such they are directed primarily towards:

- The purchasers and holders of test materials;

- Those responsible for selecting tests and determining the use to which tests will be put;

- Those who administer, score or interpret tests;

- Those who provide advice to others on the basis of test results (e.g., recruitment consultants, educational and career counsellors, trainers, succession planners);

- Those concerned with the process of reporting test results and providing feedback to people who have been tested.

The Guidelines will be of relevance to others involved in the use of tests as defined above. These include:

- the developers of tests,*

- the suppliers of tests,*

*those involved in the training of test users,
those who take tests and their relevant others (e.g., parents, spouses,
partners),
professional bodies and other associations with an interest in the use
of psychological and educational testing, and
policy makers and legislators.*

While aimed primarily at professional practice, most aspects of the good practice embodied in the Guidelines will also be of relevance to those who use tests solely for research purposes.

The Guidelines are not intended to cover every type of assessment technique (e.g., unstructured or semi-structured interviews, assessed group activities) or every situation in which assessment occurs (e.g., employment assessment centres). Yet many of the Guidelines are likely to be applicable in assessment situations and for purposes more general than those concerned primarily with psychological and educational testing (for example, the use of assessment centres for employee placement or selection, semi-structured and structured interviews, or assessment for selection, career guidance and counselling).

Contextual factors

The Guidelines are applicable internationally. They may be used to develop specific local standards through a process of contextualisation. It is recognised that there are many factors which affect how standards may be managed and realised in practice. These contextual factors have to be considered at the local level when interpreting the Guidelines and defining what they would mean in practice within any particular setting.

The factors that need to be considered in turning Guidelines into specific standards include:

*social, political, institutional, linguistic, and cultural differences
between assessment settings;
the laws of the country in which testing is taking place;
existing national guidelines and performance standards set by
professional psychological societies and associations;
differences relating to individual versus group assessment;
differences related to the test setting (educational, clinical, work-
related and other assessment);
who the primary recipients of the test results are (e.g., the test-takers,
their parents or guardian, the test-developer, an employer or other
third party);*

differences relating to the use of test results (e.g., for decision-making, as in selection screening, or for providing information to support guidance or counselling); and

variations in the degree to which the situation provides opportunity for the accuracy of interpretations to be checked in the light of subsequent information and amended if needed.

Knowledge, Understanding, and Skill

Knowledge, understanding and skill underpin all the test user competencies. The nature of their content and level of detail are likely to vary between countries, areas of application and as a function of the level of competence required to use a test.

The Guidelines do not contain detailed descriptions of these. However, when applying the Guidelines for use in specific situations the relevant knowledge, skills, abilities and other personal characteristics will need to be specified. This specification is part of the process of contextualisation, through which generic guidelines are developed into specific standards. The main areas descriptions of knowledge, understanding and skills need to cover include the following.

Relevant declarative knowledge.

This includes:

knowledge of basic psychometric principles and procedures, and the technical requirements of tests (e.g., reliability, validity, standardisation);

knowledge of tests and measurement sufficient to enable the proper understanding of test results;

knowledge and understanding of relevant theories and models of ability, of personality or other psychological constructs, or of psychopathology, as necessary to properly inform the choice of tests and the interpretation of test results; and

knowledge of the tests and the test suppliers relevant to one's area of practice.

Instrumental knowledge and skills

These include:

knowledge and skills relating to specific assessment procedures or instruments, including the use of computer-based assessment procedures;

specialised knowledge of and practitioner skills associated with using those tests that are within one's repertoire of assessment tools; and

knowledge and understanding of the construct or constructs underlying test scores, where this is important if valid inferences are to be drawn from the test results.

The Guidelines cover:

General personal task-related skills

This includes:

the performance of relevant activities such as test administration, reporting, and the provision of feedback to test takers and other clients; oral and written communication skills sufficient for the proper preparation of test takers, test administration, the reporting of test results, and for interaction with relevant others (e.g., parents, or organisational policy makers); and interpersonal skills sufficient for the proper preparation of test takers, the administration of tests, and the provision of feedback of test results.

Contextual knowledge and skills

This includes:

knowing when and when not to use tests; knowing how to integrate testing with other less formal components of the assessment situation (e.g., biographical data, unstructured interview and references etc.); and knowledge of current professional, legal, and ethical issues relating to the use of tests, and of their practical implications for test use.

Task management skills

This includes:

knowledge of codes of conduct and good practice relating to the use of tests, test data, the provision of feedback, the production and storage of reports, the storage of and responsibility for test materials and test data; and knowledge of the social, cultural, and political context in which the test is being used, and the ways in which such factors might affect the results, their interpretation and the use to which they are put.

Contingency management skills

This includes:

knowing how to deal with problems, difficulties, and breakdowns in routine;

knowing how to deal with a test taker's questions during test administration etc.; and

knowing how to deal with situations in which there is the potential for test misuse or for misunderstanding the interpretation of test scores.

1 Take responsibility for ethical test use

Competent test users should:

1.1 Act in a professional and ethical manner

- 1.1.1 Promote and maintain professional and ethical standards.
- 1.1.2 Have a working understanding of current professional and ethical issues and debates relating to the use of tests in their field of application.
- 1.1.3 Implement an explicit policy on testing and test use.⁶⁵
- 1.1.4 Ensure that people who work for or with them adhere to appropriate professional and ethical standards of behaviour.
Conduct communications with due concern for the sensitivities of the test taker and other relevant parties.
Represent tests and testing in a positive and balanced manner in communications with and through the media.
- 1.1.7 Avoid situations in which they may have or be seen to have a vested interest in the outcome of the assessment, or where the assessment might damage the relationship with their client.

1.2 Ensure they have the competence to use tests

- 1.2.1 Work within the limits of scientific principle and substantiated experience.
- 1.2.2 Set and maintain high personal standards of competence.
- 1.2.3 Know the limits of their own competence and operate within those limits.
- 1.2.4 Keep up with relevant changes and advances relating to the tests they use, and to test development, including changes in legislation and policy, which may impact on tests and test use.

⁶⁵ An example policy outline is attached as Appendix A.

1.3 Take responsibility for their use of tests

- 1.3.1 Only offer testing services and only use tests for which they are qualified.
- 1.3.2 Accept responsibility for the choice of tests used, and for the recommendations made.
- 1.3.3 Provide clear and adequate information to participants in the testing process about the ethical principles and legal regulations governing psychological testing.
- 1.3.4 Ensure that the nature of the contract between test-taker and tester is clear and understood.⁶⁶
- 1.3.5 Be alert to any unintended consequences of test use.
- 1.3.6 Endeavour to avoid doing harm or causing distress to those involved in the testing process.

1.4 Ensure that test materials are kept securely

- 1.4.1 Ensure secure storage of and control access to test materials
- 1.4.2 Respect copyright law and agreements that exist with respect to a test including any prohibitions on the copying or transmission of materials in electronic or other forms to other people, whether qualified or otherwise.
- 1.4.3 Protect the integrity of the test by not coaching individuals on actual test materials or other practice materials that might unfairly influence their test performance.
- 1.4.4 Ensure that test techniques are not described publicly in such a way that their usefulness is impaired

1.5 Ensure that test results are treated confidentially.

- 1.5.1 Specify who will have access to results and define levels of confidentiality.
- 1.5.2 Explain levels of confidentiality to individuals before tests are administered.
- 1.5.3 Limit access to results to those with a right to know.
- 1.5.4 Obtain the relevant consents before releasing results to others.
- 1.5.5 Protect data kept on file so that only those who have a right of access can obtain them.
- 1.5.6 Establish clear guidelines as to how long test data are to be kept on file.

⁶⁶ An example 'contract' between test user and test taker is attached as Appendix B.

- 1.5.7 Remove names and other personal identifiers from databases of results that are archived, for research use, development of norms or other statistical purposes.

2 Follow good practice in the use of tests

2.1 Evaluate the potential utility of testing in an assessment situation

Competent test users will:

- 2.1.1 Produce a reasoned justification for the use of tests.
- 2.1.2 Ensure there has been a thorough analysis of the client's needs, reasons for referral, or of the diagnostic category, condition, or job for which assessment is being used.
- Establish that the knowledge, skills, abilities, aptitudes or other characteristics, which the tests are intended to measure, are correlates of relevant behaviours in the context about which inferences are to be drawn.
- Seek other relevant collateral sources of information.
- 2.1.5 Assess the advantages and disadvantages of using tests compared with other sources of information.
- 2.1.6 Ensure that full use is made of all available collateral sources of information.

2.2 Choose technically sound tests appropriate for the situation

Competent test users will:

- 2.2.1 Examine current information covering the range of potentially relevant tests (e.g., from specimen sets, independent reviews, expert advice), before selecting a test to use.
- 2.2.2 Determine that the test's technical and user documentation provides sufficient information to enable evaluation of the following:
- a) *scope or coverage and representativeness of test content, appropriateness of norm groups, difficulty level of content etc.;*
 - b) *accuracy of measurement and reliability demonstrated with respect to relevant populations;*
 - c) *validity (demonstrated with respect to relevant populations) and relevance for the required use;*
 - d) *freedom from systematic bias in relation to the intended test taker groups;*

- e) *acceptability to those who will be involved in their use, including perceived fairness and relevance;*
 - f) *practicality, including time required, costs, and resource needs.*
- 2.2.3 Avoid the use of tests that have inadequate or unclear supporting technical documentation;
- 2.2.4 Use tests only for those purposes where relevant and appropriate validity evidence is available.
- 2.2.5 Avoid judging a test solely on the basis of face value, test-user testimonials, or advice from those with a vested commercial interest.
- 2.2.6 Respond to requests from relevant interested parties (e.g. test takers, parents, managers) by providing sufficient information to allow them to understand why the test was chosen.

2.3 Give due consideration to issues of fairness in testing

When tests are to be used with individuals from different groups (e.g., groups differing in terms of gender, cultural background, education, ethnic origin, or age), competent test users will make all reasonable efforts to ensure that:

- 2.3.1 The tests are unbiased and appropriate for the various groups that will be tested.
- 2.3.2 The constructs being assessed are meaningful in each of the groups represented.
- Evidence is available on possible group differences in performance on the test.
- Evidence relating to differential item functioning (DIF) is available, where relevant.
- 2.3.5 There is validity evidence to support the intended use of the test in the various groups.
- 2.3.6 Effects of group differences not relevant to the main purpose (e.g., differences in motivation to answer, or reading ability) are minimised.

In all cases, Guidelines relating to the fair use of tests are interpreted in the context of local policy and legislation.⁶⁷

⁶⁷ The Guidelines in this section focus on what is 'best practice'. However, in many countries, issues relating to the fair use of tests must also take account of national laws (e.g., the Americans with Disabilities Act, 1990, in the USA, or the Race Relations Act, 1976, in the UK).

When testing in more than one language (within or across countries⁶⁸), competent test users will make all reasonable efforts to ensure that:

- 2.3.8 Each language or dialect version has been developed using a rigorous methodology meeting the requirements of best practice.
- 2.3.9 The developers have been sensitive to issues of content, culture and language.
- 2.3.10 The test administrators can communicate clearly in the language in which the test is to be administered.
- 2.3.11 The test taker's level of proficiency in the language in which the test will be administered is determined systematically and the appropriate language version is administered or bilingual assessment is performed, if appropriate.

When tests are to be used with people with disabilities, competent test users will make all reasonable efforts to ensure that:

- 2.3.12 Advice is sought from relevant experts on the potential effects of the various disabilities on test performance.
- 2.3.13 Potential test takers are consulted and their needs and wishes are given proper consideration.
- 2.3.14 Adequate arrangements are made when test takers include people with hearing, visual or motor impairments, or other disabilities (e.g., learning impairments, dyslexia).
- 2.3.15 Use of alternative assessment procedures, rather than modifications to tests, is considered (e.g., other more suitable tests, or alternative structured forms of assessment).
- 2.3.16 Relevant professional advice is sought if the degree of modification required for use by those with disabilities is beyond the experience of the test user.
- 2.3.17 Modifications, when necessary, are tailored to the nature of the disability and are designed to minimize impact on score validity.
- 2.3.18 Information regarding the nature of any modifications made to a test or testing procedure is provided to those who interpret or act upon the test scores whenever the withholding of such information might otherwise result in biased interpretation or an unfair decision.

⁶⁸ These Guidelines relate not only to different national languages and dialects, but also to special forms of communication, such as sign language, used to overcome the effects of forms of disability.

2.4 *Make necessary preparations for the testing session*

The competent test user will make all reasonable efforts to:

- 2.4.1 Provide relevant parties in a timely manner with clear information concerning the purpose of testing, ways in which they might best prepare for the test session, and the procedures to be followed.
- 2.4.2 Advise test takers of the linguistic or dialectic groups for which the test is considered appropriate.
- 2.4.3 Send test takers approved practice, sample, or preparation materials where these are available and where this is consistent with recommended practice for the tests concerned.
- 2.4.4 Explain clearly to test takers their rights and responsibilities⁶⁹.
- 2.4.5 Gain the explicit consent of test takers or their legal guardians or representatives before any testing is done.
- 2.4.6 Explain, when testing is optional, the consequences of taking or not taking the test to relevant parties so that they can make an informed choice.
- 2.4.7 Make the necessary practical arrangements by ensuring that:
 - a) *preparations conform to those stipulated in the publisher's manual;*
 - b) *locations and facilities for testing have been arranged well in advance, and the physical environment is accessible, safe, quiet, free from distractions and appropriate for the purpose;*
 - c) *sufficient materials are available and have been checked to ensure there are no marks left by previous users on question booklets or answer sheets;*
 - d) *staff who will be involved in the administration are competent;*
 - e) *appropriate arrangements have been made for the testing of people with disabilities*⁷⁰.
- 2.4.8 Anticipate likely problems and counteract them through thorough preparation of materials and instructions.

2.5 *Administer the tests properly*

The competent test user will:

- 2.5.1 Establish rapport by welcoming test-takers and briefing them in a positive fashion.

⁶⁹ See Appendix B.

⁷⁰ See Appendix C.

ДОДАТКИ

- 2.5.2 Act to reduce test-taker anxiety and avoid creating or reinforcing unnecessary anxiety.
- 2.5.3 Ensure potential sources of distraction (e.g., wristwatch alarms, mobile phones, pagers) are removed.
- 2.5.4 Ensure test-takers have the materials they require for taking the test before it begins.
- 2.5.5 Administer tests under appropriate supervised conditions.
- 2.5.6 Wherever possible, administer test instructions in the primary language of the test takers, even where the test content is designed to provide evidence of knowledge or skills in a non-primary language.
- 2.5.7 Adhere strictly to the directions and instructions as specified in the test manual while making reasonable accommodations for persons with disabilities.
- 2.5.8 Read instructions clearly and calmly.
- 2.5.9 Provide adequate time for examples to be completed.
- 2.5.10 Observe and record deviations from test procedures.
- 2.5.11 Monitor and record response times accurately where appropriate.
- 2.5.12 Ensure all materials are accounted for at the end of each testing session.
- 2.5.13 Administer tests by modes that permit adequate and appropriate levels of supervision and authentication of the identity of the test takers.
- 2.5.14 Ensure those assisting the administration have had proper training.
- 2.5.15 Ensure test takers are not left unattended or subjected to distracting activities during a supervised test session.
- 2.5.16 Provide appropriate assistance to test takers who show signs of undue distress or anxiety.

2.6 Score and analyse test results accurately

Competent test users will:

- 2.6.1 Follow carefully the standardised procedures for scoring.
- 2.6.2 Carry out appropriate raw score conversions to other relevant types of scale.
- 2.6.3 Choose scale types relevant to the intended use of the test scores.
- 2.6.4 Check score scale-conversions and other clerical procedures for accuracy.

- 2.6.5 Ensure that invalid conclusions are not drawn from comparisons of scores with norms that are not relevant to the people being tested or are outdated.
- 2.6.6 Compute, where appropriate, composite scores using standard formulae and equations.
- 2.6.7 Employ procedures to screen test results to recognise improbable or unreasonable scores.
- 2.6.8 Clearly and accurately label scales in reports, and provide clear identification of norms, scales types, and equations used.

2.7 Interpret results appropriately

Competent test users will:

- 2.7.1 Have a good professional understanding of the test's theoretical or conceptual basis, technical documentation and guidance on the use and interpretation of the scale scores.
- 2.7.2 Have a good understanding of the scales used, the characteristics of the norm or comparison groups, and the limitations of the scores.
- 2.7.3 Take steps to minimise the effects on test interpretation of any biases the test interpreter may have towards members of the test taker's cultural group.
- 2.7.4 Use appropriate norm or comparison groups where available.
- 2.7.5 Interpret results in the light of available information about the person being tested (including age, gender, schooling, culture and other factors) with due regard for the technical limitations of the test, the assessment context, and the needs of those with a legitimate interest in the outcome of the process.
- 2.7.6 Avoid over-generalising the results of one test to traits or human characteristics which are not measured by the test.
- 2.7.7 Consider each scale's reliability, error of measurement and other qualities which may have artificially lowered or raised results when interpreting scores.
- 2.7.8 Give due consideration to the available evidence of validity, with respect to the construct being measured for members of the test takers' relevant demographic groups (e.g., cultural, age, social class, and gender groups).
- 2.7.9 Use passing scores (cut-scores) in test interpretation only when evidence of the validity for the pass scores is available and supports its use.

- 2.7.10 Be aware of negative social stereotyping that may pertain to members of the test taker's group (e.g., cultural group, age, social class, and gender) and avoid interpreting tests in a manner that perpetuates such stereotyping.
- 2.7.11 Take into account any individual or group variations from standard procedures in test administration.
- 2.7.12 Take into account any evidence of prior experience with the test where there are data available relating to the effect of such experience on test performance.

2.8 *Communicate the results clearly and accurately to relevant others*

Competent test users will:

- 2.8.1 Identify appropriate parties who may legitimately receive test results.
With the informed consent of the test takers, or their legal representatives, produce written or oral reports for relevant interested parties.
Ensure that the technical and linguistic levels of any reports are appropriate for the level of understanding of the recipients.
- 2.8.4 Make clear that the test data represent just one source of information and should always be considered in conjunction with other information.
- 2.8.5 Explain how the importance of the test results should be weighted in relation to other information about the people being assessed.
- 2.8.6 Use a form and structure for a report that is appropriate to the context of the assessment.
- 2.8.7 When appropriate, provide decision-makers with information on how results may be used to inform their decisions.
- 2.8.8 Explain and support the use of test results used to classify people into categories (e.g., for diagnostic purposes or for job selection).
- 2.8.9 Include within written reports a clear summary, and when relevant, specific recommendations.
- 2.8.10 Present oral feedback to test takers in a constructive and supportive manner.

2.9 *Review the appropriateness of the test and its use*

Competent test users will:

- 2.9.1 Monitor and periodically review changes over time in the populations of individuals being tested and any criterion measures being used.

- 2.9.2 Monitor tests for evidence of adverse impact.
- 2.9.3 Be aware of the need to re-evaluate the use of a test if changes are made to its form, content, or mode of administration.
- 2.9.4 Be aware of the need to re-evaluate the evidence of validity if the purpose for which a test is being used is changed.
- 2.9.5 Where possible, seek to validate tests for the use to which they are being put, or participate in formal validation studies.
- 2.9.6 Where possible, assist in updating information regarding the norms, reliability and validity of the test by providing relevant test data to the test developers, publishers or researchers.

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1985). *Standards for Educational and Psychological Testing*. Washington DC: American Psychological Association.
- Bartram, D. (1995). The Development of Standards for the Use of Psychological Tests in Occupational Settings: The Competence Approach. *The Psychologist*, May, 219-223.
- Bartram, D. (1996). Test Qualifications and Test Use in the UK: The Competence Approach. *European Journal of Psychological Assessment*, **12**, 62-71.
- Canadian Psychological Association. (1987). *Guidelines for Educational and Psychological Testing*. Ottawa: Canadian Psychological Association.
- Eyde, L. D., Moreland, K. L. & Robertson, G. J. (1988). *Test User Qualifications: A Data-based Approach to Promoting Good Test Use*. Report for the Test User Qualifications Working Group. Washington DC: American Psychological Association.
- Eyde, L. D., Robertson, G. J., Krug, S. E. et al (1993). *Responsible Test Use: Case Studies For Assessing Human Behaviour*. Washington DC: American Psychological Association.
- Fremer, J., Diamond, E. E. & Camara, W. J. (1989). Developing a Code of Fair Testing Practices in Education. *American Psychologist*, **44**, 1062-1067.
- Hambleton, R. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment*, **10**, 229-244.
- Joint Committee on Testing Practices. (1988). *Code of Fair Testing Practices in Education*. Washington DC: Joint Committee on Testing Practices.

Joint Committee on Testing Practices. (2000). *Rights and Responsibilities of Test Takers: Guidelines and Expectations*. Washington DC: Joint Committee on Testing Practices.

Kendall, I., Jenkinson, J., De Lemos, M. & Clancy, D. (1997). *Supplement to Guidelines for the use of Psychological Tests*. Australian Psychological Society.

Moreland, K. L., Eyde, L. D., Robertson, G. J., Primoff, E. S. & Most, R. B. (1995). Assessment of Test User Qualifications: A Research-Based Measurement Procedure. *American Psychologist*, **50**, 14-23.

Schafer, W. D. (1992). *Responsibilities of Users of Standardized Tests: RUST Statement Revised*. Alexandria, VA: American Association for Counseling and Development.

Van de Vijver, F. & Hambleton, R. (1996). Translating tests: some practical guidelines. *European Psychologist*, **1**, 89-99.

Appendix A: Guidelines for an outline policy on testing.

The following guidelines relate to the need for organizations to consider their policy on testing in a systematic manner and to ensure that everyone involved is clear as to what the policy is. The need for an explicit policy on testing is not confined to large organisations. Small and medium-sized enterprises that use testing, as well as large ones, should pay regard to testing policy in the same way as they do to health and safety, equal opportunities, disability and other areas relating to good practice in the management, treatment and care of personnel.

While the following considerations or requirements may need to be adapted for use by individual test users operating as sole professional practitioners, it remains important that they have a clear understanding of their own policy and can communicate it to others.

A policy on testing is produced in order to:

- ensure personal and organisational aims are met;
- ensure that potential misuse is avoided;
- demonstrate commitment to good practice;
- ensure test use is appropriate for its purpose;
- ensure tests do not discriminate unfairly;
- ensure evaluations are based on comprehensive, relevant information;

- *ensure tests are only used by qualified staff.*

A policy on testing will need to cover most if not all the following issues:

- *proper test use;*
- *security of materials and scores;*
- *who can administer tests, score and interpret tests;*
- *qualification requirements for those who will use the tests;*
- *test user training;*
- *test taker preparation;*
- *access to materials and security;*
- *access to test results and test score confidentiality issues;*
- *feedback of results to test takers;*
- *responsibility to test takers before, during and after test session;*
- *responsibilities & accountability of each individual user.*

Any policy needs to be regularly reviewed and updated as advances in testing, or changes in practice occur.

Relevant parties need to have access to and be informed about the policy on testing.

Responsibility for any organisation's testing policy should reside with a qualified test user who has the authority to ensure implementation of and adherence to the policy.

Appendix B: Guidelines for developing contracts between parties involved in the testing process.

Contracts between the test user and test takers should be consistent with good practice, legislation and the test user's policy on testing. The following is provided as an example of the sort of matters such a contract might cover. The details will vary as a function of the assessment context (e.g., occupational, educational, clinical, forensic) and local or national regulations and laws.

Contracts between test user, test takers and other parties are often implicit and unspoken (at least in part). Making clear the expectations, roles and responsibilities of all parties can help to avoid misunderstanding, harm, and litigation.

For their part, the test user will endeavour to:

ДОДАТКИ

- b.1 inform test takers of their rights regarding how their test scores will be used and their rights of access to them⁷¹;*
- b.2 give adequate prior warning of any financial charges that may be entailed by the testing process, who will be responsible for their payment, and when payment will be due;*
- b.3 treat test takers with courtesy, respect and impartiality regardless of race, gender, age, disability, etc.;*
- b.4 use tests of proven quality, appropriate for the test takers, and appropriate for the assessment purpose;*
- b.5 inform test takers prior to testing about the purpose of the assessment, the nature of the test, to whom test results will be reported and the planned use of the results;*
- b.6 give advance notice of when the test will be administered, and when results will be available, and whether or not test takers or others may obtain copies of the test, their completed answer sheets, or their scores⁷²;*
- b.7 have a trained person administer the test and have the results interpreted by a qualified person;*
- b.8 ensure test takers know if a test is optional and, when it is, the consequences of taking or not taking the test;*
- b.9 ensure test takers understand the conditions, if any, under which they may re-take tests, have tests re-scored, or have their scores cancelled;*
- b.10 ensure test takers know that they will have their results explained to them as soon as possible after taking the test in easily understood terms;*
- b.11 ensure test takers understand that their results are confidential to the extent allowed by law and best practice;*
- b.12 inform test takers who will have access to their results, and the conditions which scores will be released;*

⁷¹ Legislation varies between countries on this issue. For example, the current UK Data Protection Act provides rights of access to data stored on computer different from those for data written on paper.

⁷² While tests and answer sheets are not normally passed on to others, there is some variation between countries in practice relating to what test takers or others are permitted to have. However, there is much greater variation in the expectations of test takers concerning what information they will be given. It is important that contracts make clear what they will not be given as well as what they will.

- b.13 ensure that test takers are aware of the procedures for making complaints or notifying problems;*

The test user will inform test-takers that they are expected to:

- b.14 treat others with courtesy and respect during the testing process;*
- b.15 ask questions prior to testing if uncertain about why the test is to be administered, how it will be administered, what they will be required to do and what will be done with the results;*
- b.16 inform an appropriate person about any condition that they believe might invalidate the test results or which they would wish to have taken into consideration;*
- b.17 follow the instructions of the test administrator;*
- b.18 be aware of the consequences of not taking a test if they choose not to take it, and be prepared to accept those consequences;*
- b.19 ensure that, if required to pay for any the testing service(s), payment is made by the agreed date.*

Appendix C: Points to consider when making arrangements for testing people with disabilities or impairments

Considerable care and expertise is needed when the mode of administration of a test has to be changed to accommodate the needs of people with disabilities. As always, local and national law and practice⁷³ need to be considered, and the individual's rights to privacy must be respected. In seeking information regarding types and levels of disability, inquiries should only seek information relating to each person's ability to undertake the activities required to complete the test. Particular care needs to be exercised in relation to employment testing⁷⁴.

There is no simple rule of thumb that can be used to ensure that a test is administered fairly for people with all types of disability. It is a matter of professional judgement as to whether it is better to use some alternative form of assessment, or to modify the test or its mode of administration. In

⁷³ In the United States, for example, attention must be paid to the provisions of the Americans with Disabilities Act (1990). In the UK, the Disability Discrimination Act (1995), Employment Code of Practice states that "employers are required to revise tests — or the way the results of such tests are assessed — to take account of specific disabled candidates."

⁷⁴ For detailed guidance on this in the United States, see Eyde, Nestor, Heaton and Nelson (1994).

practice, it is rarely possible to norm modified tests on sufficient samples of people with equivalent disability in order to ensure comparability of the test with the standardised version. However, where data exist on, for example, the effects of changing time limits, use of Braille or audiotape spoken versions of tests, such data should guide the user in making the necessary accommodations. While full standardization of a modified version may not be possible, pilot testing on small samples of individuals should be carried out whenever practical.

Given the dearth of information about the performance of people with disabilities on tests (whether modified or not), it is often more appropriate for test result to be used in a more qualitative manner. They can be used to give an indication of the characteristic being assessed (ability, motivation, personality, etc.), which can be supplemented and supported by information gathered using other methods.

For individual assessment, the assessor can usually tailor the assessment procedures to the capabilities of the person who is being assessed. However, particular issues arise in group testing (e.g., for selection into employment). Here there may be practical difficulties involved in varying the mode of administration for particular individuals within a group administration setting. Furthermore, all parties may see differences in treatment as being unfair. For example if more time is given for test completion, those with the disability may be conscious that they are being treated 'differently', and those without the disability may feel that the extra time provides an unfair advantage.

Advice on special needs can usually be obtained from relevant disability organisations as well as the individual test takers. It is generally helpful (where the law permits) to ask the individual directly in a non-threatening and supportive way if there are any considerations that need to be taken into account⁷⁵. In many cases such consultation will enable suitable modifications to be made to the test taking environment without requiring changes to the test itself.

The following outline protocol provides a general guide to the process of deciding whether to modify testing and how to carry out the modification. Essentially, disability may contribute no variance to test scores, contribute construct relevant variance or construct irrelevant variance. In the first case, no modifications are necessary. In the final case, modifications should be aimed at removing the irrelevant source of variance (by suitable

⁷⁵ In the UK, the Disability Discrimination Act (1995) also places some obligation on the individual to raise awareness of their needs.

modification of the test conditions or substitution of a more suitable test). For the second case (construct relevant variance), however, modification to the test will affect the relevance of the test scores.

Is the disability likely to have an effect upon test performance? Many people have disabilities that would not affect test performance. In such cases, it would be inappropriate to make accommodations for them.

If the disability is likely to affect test performance, then is the effect on performance incidental to the construct being measured? For example, a person with an arthritic hand may have trouble with a speeded test which involves writing. If the ability to perform manual tasks rapidly is part of the construct being measured, then the test should not be changed. However, if the purpose is to assess visual checking speed, then an alternative mode of response would be appropriate.

When the particular disability is incidental to the construct being measured but is likely to affect the individual's performance on the test, then modification of the procedure may be considered.

Users should always consult the test manual and the publisher for guidance on modification and for information regarded alternative formats and procedures.

Users should also consult relevant disability organisations for advice and guidance on the possible implications of a specific disability, relevant literature or documentation, and the sort of adaptations or accommodations that may prove helpful.

Any modifications made to the test or test administration procedures should be carefully documented along with the rationale behind the modification.

Appendix D. Conditions governing the translation of the ITC Guidelines on Test Use.

The following conditions apply to official versions of the International Test Commission (ITC) Guidelines. The conditions delegate authority for checking the quality and accuracy of translation to the local national Psychological Association.

There is only to be one official translation in each country.

Copyright of the original version remains vested in the ITC. The ITC will give permission for an official version to be produced under the auspices of the local national Psychological Association, subject to a

copy being lodged with the ITC and the Psychological Association taking responsibility for the accuracy of the translation.

A letter formally approving the accuracy of the translation, from the relevant officer of the local national Psychological Association, must be lodged with the ITC.

The ITC will reserve the right to distribute copies of that translation itself without payment to the Psychological Association, and will want to have the guidelines available through the ITC website.

The official version should be referred to as:

«International Test Commission (ITC) Guidelines on Test Use: [Language] Version. Translation authorised by the [full name of the Psychological Association]»

The official version should have both the logo of the Psychological Association and that of the ITC clearly displayed.

The Guidelines themselves should be made available either free of charge or on a not-for-profit basis. Locally developed supporting documents, applications, qualification procedures, etc that build on the Guidelines may be charged for on a commercial basis.

Normal copyright rules apply, and permission will need to be sought by people wishing to publish extracts. In relation to the locally translated version, the ITC delegates the giving of permission for this to the local Psychological Association responsible for the translation. A notice to this effect should appear on the Guidelines document.

ITC International Guidelines on Computer-Based and Internet Delivered Testing

Version 2005

The Council of the International Test Commission formally adopted the Guidelines [7] at its July 2005 meeting in Granada, Spain. This document is copyright International Test Commission (ITC) (c) 2005. No part of it may be reproduced, translated or quoted without the prior written agreement of the ITC.

Acknowledgements

The Council of the International Test Commission formally adopted the guidelines at its July 2005 meeting in Granada, Spain. This document is copyright International Test Commission (ITC) (c) 2005. No part of it may be reproduced, translated or quoted without the prior written agreement of the ITC. For permission to translate, quote or reproduce the International Guidelines on Computer-based and Internet-delivered testing please contact the ITC Secretary, Cheryl Foxcroft (Cheryl.Foxcroft@nmmu.ac.za). A web-based version of the guidelines can be found at <http://www.intestcom.org/guidelines>.

The guidelines were prepared for the ITC Council by Professor Dave Bartram and Dr Iain Coyne. The authors would like to thank the ITC Council for their help and support. Additionally, the authors are grateful for the assistance of the following individuals who provided such valuable input to the development of the present guidelines.

Dr Eugene Aidman, Defence Science and Technology Organisation, Australia;

Ms Dusica Boben, Produktivnost, Slovenia;

Dr Marise Born, Erasmus University Rotterdam, The Netherlands;

Ms Martyne Boutot, CSI Global Education Inc, Canada;

Prof Bruce Bracken, The College of William and Mary, USA;

Mr Patrick Coates, Promissor, England;

Dr Robert Feltham, Cubiks, England;

Dr Cyndy Fitzgerald, Caveon, USA;

Mr Ian Florance, NFER-NELSON, England;

Prof Cheryl Foxcroft, Test Commission of South Africa, South Africa;

Mr John Hackston, OPP Ltd, England;

Mr John Kleeman, Questionmark, England;

Prof Kruno Matešić, Naklada Slap, Croatia;

The National Board of Medical Examiners (NBME), USA

Mr Ian Newcombe, PSL, England;

Prof Tom Oakland, The University of Florida, USA;

Mr Richard Sale, EDAC Ltd, Cyprus;

The Society for Industrial and Organizational Psychology (SIOP), USA;

Dr Kaivo Thomson, Tallinn Pedagogical University, Estonia;

Dr Jac Zaal, Rijks Psychologische Dienst, The Netherlands.

The authors are also grateful to the many other individuals and organisations who provided feedback during the various stages of consultation and in conference presentations.

ITC Guidelines on Computer/Internet based Testing

Introduction

Over the past few years the International Test Commission (ITC) has adopted a policy of promoting good practice in testing issues where international coordination of effort is most important. For example, the ITC has devised guidelines to promote good practice in test adaptations (Hambleton, 1994; Van de Vijver & Hambleton, 1996) and good practice in test use (ITC, 2001). In recent years substantial and rapid developments have occurred in the provision of stand-alone and Internet-delivered computer based testing. These developments raise a number of issues in relation to standards of administration, security of the tests and test results and control over the testing process. Therefore, as the market for such testing increases and as the technological sophistication of the products increases issues associated with ensuring those developing, distributing, using and taking such tests and assessment tools follow good practice will increase in importance. In response to this, the ITC Council decided to invest in a program of research, consultation, and conferences designed to develop internationally agreed guidelines specifically aimed at computer/Internet based testing.

Aims and Objectives

The ultimate aims of this project were

- *to produce a set of internationally developed and recognised guidelines that highlight good practice issues in computer-based (CBT) and Internet-delivered testing*
- *to raise awareness among all stakeholders in the testing process of what constitutes good practice.*

The aim was not to 'invent' new guidelines but to draw together common themes that run through existing guidelines, codes of practice, standards, research papers and other sources, and to create a coherent structure within which these guidelines can be used and understood. Contributions to the guidelines have been made by psychological and educational testing specialists, including test designers, test developers, test publishers and test users drawn from a number of countries.

Further, the aim is to focus on the development of guidelines specific to CBT/Internet based testing, not to reiterate good practice issues in testing in general. Clearly, any form of testing and assessment should conform to good practice issues regardless of the method of presentation. These guidelines are intended to complement the ITC Guidelines on Test Use (2001), with a specific focus on CBT/Internet testing.

Development of the Guidelines

As with previous ITC guidelines, the present guidelines can be seen as a benchmark against which existing local standards can be compared or as a basis for the development of locally applicable standards or codes of practice. The advantage of these guidelines is that local standards can be compared to these set guidelines for coverage and international consistency in order to promote consistency across national boundaries and for benchmarking purposes.

The project commenced with an initial literature search and review of existing references and guidelines on computer-based testing and Internet testing from a number of different countries (see Appendix). A number of these sources were particularly influential in the development of the guidelines:

- *Bartram, D. (2001). The impact of the Internet on testing: Issues that need to be addressed by a Code of Good Practice. Internal report for SHL Group plc.*
- *British Psychological Society Psychological Testing Centre (2002) Guidelines for the Development and Use of Computer-based Assessments.*
- *European Federation of Psychologists' Associations (EFPA). Review model for the description and evaluation of psychological tests (Bartram, 2002).*
- *British Standards' Institute (BSI). BS 7988 (2001). A code of practice for the use of information technology for the delivery of assessments.*
- *Association of Test Publishers (ATP). Guidelines for Computer Based Testing*

The next stage involved a small scale survey of United Kingdom test publishers, examining good practice issues in Internet-delivered personality tests in the UK. Further examples of good practice were highlighted from this survey.

As a third method of obtaining relevant information, the ITC organised a conference in Winchester, England in June 2002 on Computer-based

Testing and the Internet. The goal of this conference was to bring together people working in the field of computer/Internet testing (e.g., practitioners, scholars, industry leaders and others) from around the world and to extract common issues and themes that would inform the guidelines. In total 254 delegates from 21 countries attended the conference. The conference was composed of workshops, keynote presentations and themed papers, posters and symposia on a number of topics concerning computer/Internet testing. A review of the material from this conference coupled with the small survey data and literature review provided the basis for the development of the draft guidelines for initial consultation (version 0.3).

Four general issues emerged from the information gathering process and these formed the basis of the development of an initial draft version. The four issues were:

- *Technology — ensuring that the technical aspects of CBT/Internet testing are considered, especially in relation to the hardware and software required to run the testing.*
- *Quality — ensuring and assuring the quality of testing and test materials and ensuring good practice throughout the testing process.*
- *Control — controlling the delivery of tests, test-taker authentication and prior practice.*
- *Security — security of the testing materials, privacy, data protection and confidentiality.*

These four issues were considered high level issues and were further broken down into second-level specific guidelines. A third-level set of accompanying examples is provided to the relevant stakeholder. The guidelines are primarily written to provide advice to test developers, test publishers and test users; however, these guidelines also provide a useful source of reference for test-takers. Given these intended applications, the guidelines are structured in a three (main stakeholders) by three (level of guideline) matrix.

Following development of the initial draft by the two authors, a consultation process was undertaken. This involved circulation of the draft to all those who attended the ITC Conference in Winchester and all those on the ITC circulation list for Testing International. A copy also was placed on the ITC web site. Comments on the draft guidelines were received and version 0.4 produced. In addition the report of the APA Internet Task force was published (Naglieri et al, 2004). This was reviewed in detail and elements from the report were included in version 0.5 of the draft guidelines.

Another cycle of consultation was implemented including those people previously contact in the first consultation process. The revisions and edits from this process were completed and version 0.6 of the draft guidelines was produced. Final revisions were produced and the final draft version was devised (1.0). The current guidelines (version 2005) were officially launched in July 2005 after approval by the ITC Council.

Timeline

The following shows the timeline in the design and development of the guidelines.

1. *Completion of first draft and first consultation initiated: March 2003*
2. *End of first Consultation period: June 2003.*
3. *Revisions completed and second consultation initiated: February 2004*
4. *End of second Consultation period: April 2004*
5. *A symposium on CBT and Internet testing at the International Congress of Psychology in Beijing, August 2004.*
6. *Final version for approval: January 2005*
7. *Development of final version and design of web-based version: March 2005*
8. *Approval by ITC Council and formal launch: July 2005*

Scope

As with the International Guidelines of Test Use (2001), the current guidelines use the terms 'test' and 'testing' in their broadest sense and include psychological and educational tests used in clinical, health, educational and work and organisational assessment settings. CBT/Internet tests should be supported by evidence of their technical adequacy for their intended purpose. These guidelines are aimed at tests conducted both online and onscreen (offline), which can include testing via the use of a CD ROM or a download executable. The document includes guidance for fully computerised testing and for part-computerised testing and the reader can refer to the most appropriate elements. For example, only the sending and scoring of assessment papers may be computerised (the rest paper and pencil). Given this, the guidelines dealing with security and confidentiality of data are important.

In general, the guidelines can apply to both high stakes and low stakes assessment. As an example, high stakes assessments are those where a third party requires the results of the test for use in the process of making an important decision about a test-taker (high stakes testing

may also include those that are used to make decisions about groups of test-takers, such as a school class). By contrast, an example of low stakes assessment would be where the test-taker obtains the information for his or her own interest. That some guidelines apply only to high stakes testing environments is made clear within the text itself.

Again, unless otherwise specified in the text, the guidelines presented here should be considered as applying to a number of modes of supervision and across a number of testing scenarios. Four modes of test administration are considered:

- *Open mode* — Where there is no direct human supervision of the assessment session and hence there is no means of authenticating the identity of the test-taker. Internet-based tests without any requirement for registration can be considered an example of this mode of administration.
- *Controlled mode* — No direct human supervision of the assessment session is involved but the test is made available only to known test-takers. Internet tests will require test-takers to obtain a logon username and password. These often are designed to operate on a one-time-only basis⁷⁶.*
- *Supervised (Proctored) mode* — Where there is a level of direct human supervision over test-taking conditions. In this mode test-taker identity can be authenticated. For Internet testing this would require an administrator to log-in a candidate and confirm that the test had been properly administered and completed.
- *Managed mode* — Where there is a high level of human supervision and control over the test-taking environment. In CBT testing this is normally achieved by the use of dedicated testing centres, where there is a high level of control over access, security, the qualification of test administration staff and the quality and technical specifications of the test equipment⁷⁷.**

Application of these guidelines needs to be considered in terms of their relevance for a range of different testing scenarios (e.g., guidelines are more appropriate for the more high stakes forms of scenarios). For example, in relation to testing in work and organisational settings, four main scenarios can be identified:

⁷⁶ Standardisation of the testing environment is not possible with open mode testing, and often not possible in the controlled mode of testing

⁷⁷ Standardisation is possible with supervised mode and managed mode

- *Guidance* — personnel development or career guidance, where the test-taker requires the information for his/her own interest.
- *Pre-screening recruitment* — covers assessment carried out on people up to the point at which they are sifted to form a short-list.
- *Post sift selection* — assessments on a known set of applicants who have been previously short-listed
- *Post-hire assessment* — assessments carried out on employees of an organisation by or on behalf of the organisation. This may be either high or low stakes assessment.

Additionally in clinical/counseling settings, four scenarios could be:

- *Development and decision-making purposes* — where the information is used by the client and therapist/counselor to identify aspects of functioning that require development or to make decisions (e.g., career assessment).
- *Screening* — to get a global picture of the client's functioning.
- *Diagnostic purposes* — to identify specific strengths and weaknesses which can guide intervention planning.
- *Planning and evaluating intervention/therapy.*

Each of these raises different issues regarding control and security.

- * *Standardisation of the testing environment is not possible with open mode testing, and often not possible in the controlled mode of testing.*
- ** *Standardisation is possible with supervised mode and managed mode.*

Who are the Guidelines for?

The guidelines apply to the use of CBT and Internet tests in professional practice. Thus they are directed towards test users who:

- *purchase and use CBT/Internet tests;*
- *are responsible for selecting tests and determining the use to which tests will be put;*
- *administer, score, or interpret tests (invigilators/proctors);*
- *provide advice to others on the basis of test results (e.g., recruitment consultants, educational and career counsellors, educational and school psychologists, trainers, succession planners, organisational development consultants);*
- *are concerned with the process of reporting test results and providing feedback to people who have been tested.*

ДОДАТКИ

These guidelines also specifically address three other main stakeholders in the testing process:

- *developers of CBT and Internet tests,*
- *publishers of CBT and Internet tests (who also may be involved in test development), and*
- *consultants to developers and publishers of CBT and Internet tests.*

The guidelines are relevant to others involved in the use of CBT and Internet tests. These include:

- *those involved in the training of test users,*
- *those who take tests and their relevant others (e.g., parents, spouses, partners),*
- *professional bodies and other associations with an interest in the use of psychological and educational testing, and*
- *policy makers and legislators.*

Contextual factors

The guidelines are intended to be applicable internationally. Many factors may affect how standards may be managed and realised in practice. These contextual factors have to be considered at the local level when interpreting these guidelines and defining what they would mean in practice within any particular setting.

The factors that need to be considered for turning the guidelines into specific standards include:

- *social, political, institutional, linguistic, and cultural differences between assessment settings;*
- *laws, statutes, policy and other legal documentation that addresses testing issues;*
- *laws applying to the various countries through which test data may pass or be stored;*
- *existing national guidelines and performance standards set by professional psychological societies and associations;*
- *differences relating to individual versus group assessment;*
- *differences related to the test setting (educational, clinical, work-related and other assessment);*
- *who the primary recipients of the test results are (e.g., the test-takers, their parents or guardian, the test developer, an employer or other third party);*

- *differences relating to the use of test results (e.g., for decision-making, as in selection screening, or for providing information to support guidance or counselling); and*
- *variations in the degree to which opportunity exists for the accuracy of interpretations to be checked in light of subsequent information and amended if needed.*

1. Give due regard to technological issues in Computer-based (CBT) and Internet Testing

a. Give consideration to hardware and software requirements

1. Test Developers

1. *Provide a clear description of the minimum hardware and software requirements of the CBT. For Internet testing specify browsers which will support the test.*
2. *Conduct adequate usability testing of the system requirements using the appropriate delivery platforms to ensure consistency of appearance and delivery.*
3. *Use appropriate technological features to enhance usability and follow established graphical user interface (GUI) design standards. For example, complex graphics and interactive features may reduce software running speed or increase download time. Items should be designed to fit the test purpose and objectives of assessment, and advanced multimedia features should be used only where justified by validity.*
4. *Design the system to accommodate likely advances in technology.*
5. *Design the Internet-delivered testing system to take account of the possibility of fluctuations in demand at different times.*
6. *Ensure applications of technology advances are tested, documented, and explained to users.*
7. *Minimise the number of updates and version changes that are issued.*
8. *Take account of the widely differing connection speeds that apply globally.*

2. Test Publishers

1. *Verify the documented minimum hardware, software or browser requirements to ensure that they are communicated clearly to the user. Ensure that other technical and operational requirements for the test are explained to the user.*

ДОДАТКИ

2. *Confirm that adequate testing of the system has been completed and documented on the appropriate delivery platforms stated to be suitable.*
3. *Use only software or hardware features that are essential for measuring the construct and that are likely to be available on systems used by the intended test users and test-takers.*
4. *Ensure that the test will be as easy as possible to support and maintain in light of likely developments in hardware and software (operating systems etc).*
5. *Test and document any new features added to the program after publication.*

3. Test Users

1. *Ensure that you have sufficient understanding of the technical and operational requirements of the test (i.e. hardware and software), as well as the necessary hardware, software and human resources to obtain, use, and maintain the CBT on an on-going basis.*
2. *Confirm that the system the test-taker is using is documented as being suitable.*
3. *Ensure there is a good justification for the use of complex software, graphics, and technical IT features in the CBT/Internet test.*
4. *Monitor supplier for information on future changes to the hardware requirements, test system, or software.*
5. *Ensure understanding of the implications of changes and their impact on the testing process.*

b. Take account of the robustness of the CBT/Internet test**4. Test Developers**

1. *Test the system to confirm that it is sufficiently robust and capable of dealing with likely system failures and user error.*
2. *Ensure that the CBT/Internet test is as 'fail-safe' as possible in order to minimise problems arising while the test-taker is responding. Where possible and appropriate:*
 - *treat upper and lower case fonts as equivalent,*
 - *prevent operation of keys or controls that have no function in the test,*
 - *eliminate auto-repeat functions of keys,*
 - *prevent a test-taker from exiting the test by accident,*

- *provide timely and helpful error feedback,*
 - *follow GUI standards regarding features such as colour, layout, and design, and*
 - *if standardization is not important, allow the user multiple ways to navigate through the system, or allow the user to modify the interface to their liking.*
3. *When the CBT/Internet test is timed, design the system to respond promptly so that commands have an immediate effect on the screen (e.g., GUI design standards would indicate no more than a 2 second delay onscreen).*
 4. *When the CBT/Internet test is timed, design features so that the time required to move between questions and for the system to record the answer is not part of the timed element (e.g., the test software should deduct these times from the timing of the test or the timing clock should stop during access transitions).*
 5. *For Internet testing, minimise the impact of hang-ups , lost Internet connections and slow downloading (e.g., the system should ensure that no information is lost when the Internet connection is lost).*
 6. *Provide documentation that specifies what to do in the event of routine problems with hardware and/or software.*

5. Test Publishers

1. *Confirm the robustness of the system has been checked across a range of suitable platforms.*
2. *Provide sufficient redundancy on all systems throughout the testing site (including incoming and outgoing communications) to allow the site to operate even if one of its components fails.*
3. *Check the degree to which the test prevents user errors from causing administration problems. Provide users with guidance on what to do in the event that 'bugs' occur during testing (e.g. a test user should be able to report bugs and problems that may be experienced during the testing process).*
4. *Provide users with contact details (e.g., telephone number, internet address) for technical support.*
5. *Confirm that the CBT/Internet test responds in a timely manner when taking the test. Where this does not occur, inform test developers and discontinue use of the test until the problem is solved.*
6. *For Internet testing, put procedures in place to deal fairly with the impact of hang-ups , lost connections and slow downloads. Where dow-*

ДОДАТКИ

nload or other technical problems occur, advise the test user/taker of alternatives (e.g., using alternative media or an alternative venue).

- 7. Document and disseminate relevant technical support to test users. Where appropriate, offer technical support services with trained staff.*

6. Test Users

- 1. Before beginning a test, verify that its robustness has been adequately tested (e.g. documentation provides supporting evidence).*
- 2. Ensure processes are in place to log and resolve problems that may arise during testing.*
- 3. Check availability of the information necessary for contacting the provider of technical support and use technical support services as necessary.*
- 4. Inform test publishers/developers where problems occur with the responsiveness of the computer to the test-taker input.*
- 5. For Internet testing, know the recommended procedures for dealing with hang-ups , lost connections and slow downloads, and advise test-takers accordingly.*
- 6. Provide the test-taker with the technical support specified in the test documentation if any routine problems occur.*

c. Consider human factors issues in the presentation of material via computer or the Internet**7. Test Developers**

- 1. Design systems to follow GUI design standards that have been established by groups such as Human Factors International, including but not limited to:*
 - ensuring screens have adequate resolution and colour,*
 - using consistent screen locations and colour for instructional text and prompts,*
 - using consistent screen design, layout and colours,*
 - differentiating between test items and test instructions,*
 - displaying only relevant information on-screen and ensuring the screen is not overfilled,*
 - placing critical information at the start of the text,*
 - providing instruction screens with clear fonts and avoiding distracting logos/images,*
 - allowing test-takers to review or return to the instruction screen(s) where appropriate, and*

- *ensuring representation of status change of display entities (e.g., dimming, highlighting) is consistent in appearance, and logical and meaningful.*
- 2. *Display test name, item number, and test prompts or directions at the same location on the screen for each test page.*
- 3. *Produce non-alarming, clear and concise error messages that inform how to proceed. Following an error alert, allow the test-taker to correct any errors and continue the test in the most efficient manner possible.*

8. Test Publishers

1. *Verify that screen design issues have been taken into account in the development of the CBT/Internet test. Where problems are noticed, provide clear and detailed information about the problems to the test developer.*
2. *Verify that item presentation is consistent throughout the test.*
3. *Verify that appropriate and informative error messages are presented when necessary.*

9. Test Users

1. *Be familiar with the screen design requirements of the test and ensure that such features are compatible with the systems being used.*
2. *Ensure that test-takers are informed of screen design conventions, including where instructional text and prompts are placed, and how instructions can be accessed once testing begins.*
3. *Be familiar with how items are presented and how the test-taker is required to respond.*
4. *Verify that error messages are non-alarming and inform how to proceed.*

d. Consider reasonable adjustments to the technical features of the test for candidates with disabilities

10. Test Developers

1. *Design CBT/Internet tests with hardware/software (e.g., response format) that facilitates the participation of test-takers with disabilities and special needs.*
2. *Design CBT/Internet tests with hardware and software that can be modified to allow for appropriate test accommodations (e.g., increased font size).*

11. Test Publishers

ДОДАТКИ

1. *Confirm that the hardware/software features of the CBT/Internet test facilitate the participation of test-takers with disabilities and those with special needs (e.g., those who need larger page font).*
2. *Inform test users about the types of accommodations and modifications that can be made for test-takers with disabilities and those with special needs.*
3. *Inform test users of the acceptable limits to which tests can be modified or accommodations provided to test-takers.*
4. *Ensure that test modification and accommodations provided to test users are consistent with legislation regarding individuals with disabilities and special needs.*

12. Test Users

1. *Check that the hardware/software features facilitate the participation of test-takers with disabilities and those with special needs.*
2. *Follow best practice as in other modes of testing [see ITC Guidelines on Test Use].*
3. *Ensure that any necessary test modifications specifically address the test-taker's special needs and are within acceptable limits so as to not adversely affect score validity.*
4. *Be aware of the impact these modifications may have on the test-taker's score.*
5. *Consider the use of alternative assessment procedures, rather than modifications to CBT/Internet tests, (e.g., paper and pencil test or alternative structured forms of assessment).*

e. Provide help, information, and practice items within the CBT/Internet test**13. Test Developers**

1. *Provide clear, accurate, and appropriate technical support documentation in both electronic and paper formats. Ensure that such documentation is written at an appropriate level for its target audience.*
2. *Provide clear instructions on how to load and set up the testing system. For Internet testing, information should be provided on how to log test-takers on and off the system.*
3. *Provide sufficient and easily available on-screen instructions and help for test-takers. This should include, at a minimum, information about the test (number of items, timing, and types of items) and the testing procedure (how to navigate through the system and how to exit).*

4. *Where appropriate, develop tutorials or practice tests/items that provide test-takers the opportunity to familiarise themselves with the CBT/Internet test.*

14. Test Publishers

1. *Provide technical support documentation at a level appropriate for test users. Where appropriate, provide additional customer support services.*
2. *Disseminate instructions on how to set-up the system to test users. For Internet testing, inform, where appropriate, test users on how to log a test-taker on and off the system.*
3. *Provide clear and sufficient on-screen instructions.*
4. *Where appropriate, verify that suitable practice items and tutorials are available. For Internet testing, provide procedures to verify whether a test-taker has accessed practice items and tutorials. Often a test cannot be started until certain practice items have been completed.*

15. Test Users

1. *Understand the technical support documentation provided with the test and how to access additional technical support when needed.*
2. *Know how to set up, load and log onto the system.*
3. *Ensure the test-taker has access to information on the test and the testing process before beginning the test and is able to access on-screen help while completing the test.*
4. *For Internet testing, provide clear information to the test-taker on how to log-in to and off from the system (e.g., the use of passwords).*
5. *Provide sufficient opportunity for the test-taker to become familiar with the testing software and the required hardware.*
6. *Where appropriate, direct test-takers to appropriate Internet testing practice sites.*
7. *Where appropriate, inform the test-taker about available practice tests. Make it clear that it is the test-taker's responsibility to practice any embedded tutorials and responses to test items (e.g., use of the input device).*
8. *Where appropriate, collect data on test-taker reactions towards Internet-delivered testing and provide feedback to test developers to help them ensure a more positive experience for test-takers.*

2. Attend to quality issues in CBT and Internet testing**a. Ensure knowledge, competence and appropriate use of CBT/Internet testing****16. Test Developers**

1. *Document the constructs that are intended to be measured and investigate whether CBT/Internet mode of delivery is appropriate in terms of content and technical adequacy to access the relevant constructs.*
2. *Ensure all those involved in test design and development (item writers, psychometricians, software developers etc.) have sufficient knowledge and competence to develop CBT/Internet tests.*
3. *Remain current on recent advances in CBT/Internet testing, including advances in computer hardware and software technologies and capabilities.*
4. *Adhere to legal, professional, and ethical mandates and guidelines related to CBT/Internet testing.*
5. *It is important that during the development of items and tests, the content is protected, through the use of agreements as well as sound security procedures.*

17. Test Publishers

1. *Ensure that the CBT/Internet test is suitable in terms of content and technical adequacy for its stated purpose and intended test-taker groups.*
2. *Provide test users with sufficient information about the CBT/Internet test, its modes of operation, and basic computer functions. If appropriate, provide training materials that are specific to CBT/Internet tests and testing.*
3. *Provide test users with 'best practice' testing policies.*
4. *Provide test users with clear instructions on how to correctly access and administer Internet tests, including how to log test-takers onto the system.*
5. *Maintain and regularly update documentation relating to CBT/Internet testing, including pertinent changes in legislation and policy.*
6. *Adhere to legal, professional, and ethical mandates related to CBT/Internet testing.*
7. *For Internet testing, document the limitations of the test in terms of the professional context in which it operates:*

- *provide a statement indicating the limitations of the relationships between test user and test-taker that can be achieved through this mode (e.g. the Internet is a impersonal medium and a test user may provide only limited advice)*
- *provide a statement stating that there are limitations to the conclusions that can be reached just using the Internet test scores.*

18. Test Users

1. *Assess the appropriateness of the content and technical adequacy of CBT/Internet testing relative to alternative testing methods for each client. Inform test-takers of the purpose of the testing so they are able to make an informed decision on the appropriateness of the test for their situation.*
2. *Have adequate knowledge of the CBT/Internet test and its modes of operation. When required, attend appropriate training events and read and have knowledge of relevant training materials.*
3. *Follow best practices in the use of CBT/Internet testing and, where appropriate, create 'best practice' testing policies.*
4. *Verify that test-takers know how to interact with an Internet testing system (e.g., basic browser operation, use of access passwords).*
5. *Maintain and regularly update knowledge about CBT/Internet testing, including pertinent changes in legislation and policy.*
6. *Adhere to legal, professional, and ethical mandates related to CBT/Internet testing.*
7. *Inform test-takers of the limitations of the Internet test in terms of the professional relationship expected from this medium.*
8. *For Internet testing, provide a contact point (e.g., email or phone) for those who do not understand the purpose of the test*

b. Consider the psychometric qualities of the CBT/Internet test

19. Test Developers

1. *Document and disseminate information on the validity, reliability, and fairness of the CBT/Internet testing process.*
2. *Ensure that current psychometric standards (test reliability, validity, etc) apply even though the way in which the tests are developed and delivered may differ.*
3. *Take care that the CBT/Internet test does not require knowledge, skills, or abilities (e.g., computer skills) that are irrelevant to or might impede the test-taker's ability to perform the test.*

ДОДАТКИ

4. *Describe the theoretical and practical applications of algorithms used in test-item selection and/or controlling item or test order (as in adaptive testing).*
5. *Where test-item content changes, retest and evaluate the changes.*

20. Test Publishers

1. *Provide appropriate documentation for the psychometric properties of the CBT/Internet test.*
2. *Ensure that current psychometric standards (test reliability, validity etc.) have been met even though the way in which the tests are developed and delivered may differ.*
3. *Publish and offer online only those tests that have appropriate psychometric evidence to support their use.*
4. *When offering assessments online, give advice to test users as to what to look for in order to help them distinguish between tests with and without documented psychometric properties.*
5. *Verify that the CBT/Internet test does not require knowledge, skills or abilities that are irrelevant to the construct being assessed.*
6. *Provide documentation that describes the algorithms and measurement models used and present evidence showing that the test has been validated using these algorithms or models.*
7. *For tests based on models that may be unfamiliar to test users, provide explanations of the relevant concepts for the user.*
8. *Verify that psychometric model fit has been re-evaluated when changes are made to the test content.*

21. Test Users

1. *Ensure that documentation of the appropriate psychometric evidence is supplied with the CBT/Internet test.*
2. *Ensure that current psychometric standards (test reliability, validity etc.) have been met even though the way in which the tests are developed and delivered may differ.*
3. *Be able to distinguish between tests with and without documented psychometric properties. Those with documented evidence ensure that the evidence is appropriate for the intended use of the test.*
4. *For Internet testing, use only those websites supported by publishers who offer validated psychometric tests.*
5. *Check that the CBT/Internet test does not require knowledge, skills or abilities that are irrelevant to the construct being assessed.*

6. *Where appropriate, review and understand the documentation that describes how the CBT/Internet test uses algorithms for item generation, selection, or test construction, for controlling the order of testing, and the model underlying the development of the test.*
7. *When necessary, access appropriate training to ensure continuing professional development.*
8. *Document information provided about changes to test items or parameters and their impact on the test properties.*

c. *Where the CBT/Internet test has been developed from a paper and pencil version, ensure that there is evidence of equivalence*

22. Test Developers

1. *Provide clear documented evidence of the equivalence between the CBT/Internet test and non-computer versions (if the CBT/Internet version is a parallel form). Specifically, to show that the two versions:*
 - *have comparable reliabilities,*
 - *correlate with each other at the expected level from the reliability estimates,*
 - *correlate comparably with other tests and external criteria, and*
 - *produce comparable means and standard deviations or have been appropriately calibrated to render comparable scores.*
2. *When designing a CBT/Internet version of a non-computerised test, ensure that:*
 - *there is equivalent test-taker control (such as the ability to skip or review items) as on the manual version,*
 - *the method of item presentation ensures that the results from the CBT/Internet test are equivalent to the manual version, and*
 - *the format for responding is equivalent.*
3. *For Internet-based tests, studies of test equivalence and norming should be conducted over the Internet with participants completing the test under conditions that represent those that the intended target population will experience (e.g., unproctored or unstandardised testing conditions).*

23. Test Publishers

1. *Evaluate the documented evidence of the equivalence of the CBT/Internet test, especially if norms from manual versions are to be used by test users to interpret scores on a computerised version of the test.*

ДОДАТКИ

2. *If the developer does not provide evidence of equivalence (e.g., comparable reliabilities, etc.), conduct appropriate equivalence studies.*
3. *If the developer does not provide evidence relating to the use of the test under conditions that represent those that the intended target population will experience (e.g., unproctored, unstandardised testing), additional studies of test equivalence and norming should be conducted.*
4. *Verify that the technical features of the CBT/Internet test (e.g., test-taker control and item presentation) allow the results from the CBT/Internet test to be equivalent to the manual version.*

24. Test Users

1. *Confirm that the evidence regarding the equivalence of the CBT/Internet test to the manual version is sufficient.*
2. *If norms are based on manual versions of the test, confirm that evidence has been obtained to show equivalence of test means and SDs across versions and for appropriate subpopulations.*
3. *Verify that the technical features of the CBT/Internet test (e.g., test-taker control and item presentation) allow the results from the CBT/Internet test to be equivalent to the manual version.*
4. *Only use the test in those modes of administration for which it has been designed (e.g., do not use a test in an unproctored mode when it is specified for use only in proctored modes).*

d. Score and analyse CBT/Internet testing results accurately**25. Test Developers**

1. *Ensure the accuracy of rules/algorithms underlying the scoring of the CBT/Internet test.*
2. *Provide appropriate documentation of the use and validity of scoring rules.*
3. *Where reports classify test respondents into categories, such as 'Introverted type' or 'High sales potential', provide information in the test manual that specifies the accuracy of the classification system used to generate computer-based test interpretations (CBTI).*
4. *Describe the rationale for CBTI statements and how statements are derived from particular scores or score patterns.*
5. *When test data are hand-entered into a computer, devise procedures to allow for data to be checked for accuracy.*

26. Test Publishers

1. *Confirm that the accuracy of scoring rules has been adequately evaluated prior to test use.*
2. *Inform test users about the scoring rules employed within the CBT/Internet test (e.g., use of non-scored items, penalties for guessing).*
3. *Inform test users how CBTI statements are derived and the validity of that methodology.*
4. *Stress to test users the importance of carefully checking data input by hand into a computer for scoring.*

27. Test Users

1. *Review and understand the rules underlying the scoring of the CBT/Internet test.*
2. *Inform test-takers, when appropriate, about how scores are generated.*
3. *Know how the statements in the CBTI are derived and be aware of the limitations such methods may have.*
4. *Ensure the accuracy of test data that are hand-entered into the computer.*

e. Interpret results appropriately and provide appropriate feedback**28. Test Developers**

1. *Illustrate potential limitations of the computer-based test interpretations (CBTI) specific to the current CBT/Internet test.*
2. *Design and embed individual CBTI report templates for all stakeholders in the testing process.*
3. *Illustrate how to obtain these various reports and what is contained within each report. In particular consider the:*
 - *media (e.g., text, graphics, etc.),*
 - *complexity of the report,*
 - *report structure,*
 - *purposes of testing,*
 - *degree of modifiability,*
 - *style and tone of report, and*
 - *intended recipients.*
4. *Provide appropriate guidance on giving feedback, including necessary training requirements for interpreting the CBTI.*

29. Test Publishers

1. *Inform test users of the potential limitations of interpreting results using CBTI. Specifically:*
 - *statements in a report may be general and not directed towards the specific purpose of the assessment (or specific individuals);*
 - *interpretation is based only on scores of those tests whose data were used as input; therefore, other ancillary data which may be important cannot be taken into account (e.g., scores on other, non-computerised, forms of assessment);*
 - *for open or controlled modes of Internet testing, test-takers may have been tested in non-standardised, unproctored, or variable conditions, whereas score interpretations are based on administration in proctored, standardised conditions;*
 - *some tests are completed in an administration mode that makes it impossible to guarantee the true identity of the test-taker.*
2. *Assess the suitability of the CBTI provided within the CBT/Internet test system. In particular, take note of*
 - *evidence of the validity and utility of reports,*
 - *the coverage of the reports,*
 - *the consistency of the reports based on similar sets of data,*
 - *the acceptability of the report to intended audiences,*
 - *time, cost and length implications for a test user, and*
 - *freedom from systematic bias.*
3. *Advise test users on how best to share CBTI with test-takers and other relevant stakeholders.*
4. *Inform test users of ethical and other accepted practice issues related to providing CBTI feedback to test-takers.*

30. Test Users

1. *When interpreting the CBTI results, be aware of potential limitations, general and specific, to the reports being used. For example:*
 - *Score interpretations are based on administration in proctored, standardised conditions and the test has been administered under open or controlled modes and there is no evidence provided to support the validity of the report under such conditions.*
 - *Tests are completed in an administration mode that makes it impossible to guarantee the true identity of the test-taker.*

- *Tests alone, however administered, may not provide a complete assessment of an individual, as other confirmatory or ancillary information is not considered.*
- 2. *Select and use the most appropriate CBTI template for the client or intended audience.*
- 3. *Ensure that the language and information given in the CBTI fit the needs of the intended stakeholder (e.g., test-taker, organisation, and client).*
- 4. *Confirm that there is a sound basis for the CBTI and that its rationale is well-documented.*
- 5. *Where possible, edit CBTI reports to include information obtained from other sources to ensure a comprehensive treatment of the test-taker's background, behaviour, ability, aptitude, and personality.*
- 6. *Ensure appropriate, relevant, and timely feedback is provided to the test-taker and other relevant stakeholders.*
- 7. *Ensure that Internet testing presents test interpretations in a comprehensible and meaningful form.*
- 8. *Provide client test interpretations that are appropriate for the context and intended use of the test (e.g., high or low stakes testing, corporate versus individual applications).*
- 9. *Take account of ethical issues surrounding the provision of feedback using the Internet (e.g. the difficulty of knowing the effect of providing negative feedback to a test-taker, the lack of knowledge of the emotional state of the test-taker, or the difficulty of providing immediate support to a test-taker when feedback has a negative impact). Where appropriate, feedback should include directions on how to access support and other information.*

f. Consider equality of access for all groups

31. Test Developers

1. *Document the methods used to enhance psychometric fairness and equality of access.*
2. *Assess Differential Item Functioning (DIF) and, where DIF might be a problem for one or more groups, identify where this problem occurs and attempt to modify the test to overcome such problems.*
3. *When developing CBT/Internet tests that may be used internationally, take into account the fact that countries differ in their access to computer technology or the Internet.*
4. *For tests that are to be used internationally:*

ДОДАТКИ

- *avoid the use of language, drawings, content, graphics (etc.,) that are country or culture specific.*
 - *where culture specific tests may be more suitable than culturally-neutral ones, ensure that there is construct equivalence across the different forms.*
5. *If developing adapted versions of an Internet test for use in specific countries ensure the equivalence of the adapted version and that the adaptation conforms to the ITC Guidelines on Test Adaptation.*

32. Test Publishers

1. *Where possible, encourage test users to collect biographical data on test-takers in order to monitor the number of people from protected/minority groups who take any CBT/Internet test.*
2. *Where unequal access to CBT/Internet tests may occur, recommend that test users make alternative forms of assessment available.*
3. *Inform test users of any evidence regarding DIF for different test-taker groups.*
4. *When tests are published internationally, provide test users with advice on how to ensure equivalent access to computer technology or the Internet for geographically-diverse groups of test-takers.*
5. *Where an adapted version of a test is available, provide documentation specifying the equivalence of the adaptation to the original assessment.*

33. Test Users

1. *To monitor for possible adverse impact, collect data on the number of individuals accessing the CBT/Internet test from protected/minority groups.*
 - *For most countries such groups may be legally defined in terms of one or more of the following: ethnicity, gender, age, disability, religion, and sexual orientation.*
2. *Where there is evidence of possible inequality of access, offer the use of alternative methods of testing.*
3. *Where possible, collect data to monitor group differences in test scores.*
4. *Consider the appropriateness and feasibility of Internet testing if testing in locations with limited access to computer technology or the Internet.*
5. *If testing internationally, use the country-specific adapted versions of the test, if available.*

3. Provide appropriate levels of control over CBT and Internet testing

a. Detail the level of control over the test conditions

34. Test Developers

1. *Document the hardware, software, and procedural requirements for administration of a CBT/Internet test.*
2. *Provide a description of the test-taking conditions required for appropriate CBT/Internet test administration.*
3. *Design the CBT/Internet test to be compatible with country-specific health and safety, legal, and union regulations and rules (e.g., time on task).*

35. Test Publishers

1. *Provide sufficient details to test users on hardware, software, and procedural requirements for administering the CBT/Internet test.*
2. *Describe the test taking conditions candidates should consider when undertaking an Internet-based test.*
3. *Inform test users of the need to consider health and safety rules during CBT/Internet testing. For example, identify whether an Internet test has the facility for breaks if the testing process is lengthy.*

36. Test Users

1. *When administering the test, adhere to the standard hardware, software, and procedural requirements specified in the test manual. Before testing, ensure that software and hardware are working properly.*
2. *When testing at a specific test centre, ensure that the test-taker is comfortable with the workstation and work surface (e.g., the ergonomics are suitable). For example, test-takers should:*
 - *be encouraged to maintain proper seating posture,*
 - *be able to easily reach and manipulate all keys and controls,*
 - *have sufficient leg room, and*
 - *not be required to sit in one position for too long.*
3. *When testing via the Internet, provide instructions to test-takers that specify the best methods of taking the test.*
4. *Ensure that the facilities, conditions, and requirements of the testing conform to national health and safety, and union rules. For example, there may be rules governing the length of time a person should work at a monitor before having a break, or rules as to adequate lighting,*

heating, and ventilation. When testing over the Internet, inform test-takers of such rules and regulations.

b. Detail the appropriate control over the supervision of the testing

37. Test Developers

1. Document the level of supervision required for the CBT/Internet test.
 - Open mode — No direct human supervision required
 - Controlled mode — Although no direct human supervision is required, the test is made available only to known test-takers
 - Supervised mode — Test users are required to log on a candidate and confirm that the testing was administered and completed correctly
 - Managed mode — A high level of human supervision and control over test-taking conditions is required (as in a dedicated test centre)
2. Provide documentation for the testing scenarios for which the CBT/Internet test has been designed.

38. Test Publishers

1. Document the level of supervision expected for the CBT/Internet test.
2. Specify and restrict the use of specific CBT/Internet tests for particular testing scenarios. For example, psychometric tests for use in post-sift selection testing and/or post-hire assessment normally would not be available in open mode.

39. Test Users

1. Identify the level of supervision required to administer the CBT/Internet test.
2. Use the CBT/Internet test only in the appropriate testing scenarios for which it was designed.

c. Give due consideration to controlling prior practice and item exposure

40. Test Developers

1. For high-stakes Internet-based tests, use software that tries to equate item exposure rates for items drawn from item banks.
2. Limit pilot testing of items on live tests, to minimize unnecessary exposure.
3. Make sure item banks are sufficiently large to permit making multiple parallel forms secure and to manage item exposure rates in adaptive testing.

4. *When parallel forms of a test are created, undertake appropriate psychometric analysis to document their equivalence.*
5. *Contemplate delivery strategies that deter memorization of test content (e.g. by generation of unique tests for each candidate from item banks; or by use of computer adaptive testing).*
6. *Control exposure of fixed forms in geographies where cheating is more prevalent by restricted administration to supervised or managed modes.*

41. Test Publishers

1. *Verify that Internet-based maximum performance tests have appropriate controls to reduce item exposure .*
2. *Provide test users with sufficient information on and training in how to control item exposure .*
3. *Where appropriate, provide test-takers with practice without compromising the security of the test items.*

42. Test Users

1. *Document for test-takers the equivalence of parallel or multiple forms of a test.*
2. *Protect the CBT/Internet test from previous item exposure by not coaching test-takers with actual test content.*
3. *Where appropriate, provide test-takers with practice without compromising the security of the actual test items themselves.*

d. Give consideration to control over test-taker's authenticity and cheating

43. Test Developers

1. *Design features within the system (e.g., the facility for passwords and username access) that enables test publishers/users to have a level of control over access to various parts of the assessment system.*

44. Test Publishers

1. *Detail the level of authentication required to access various parts of the assessment system, based on the mode of operation used. Exercise control by requiring test users (in the Supervised and Managed modes) and test-takers (in the Controlled mode) to use a username and password when accessing the test.*
2. *For moderate or high stakes assessment involving multiple stages, provide information on how test users can reduce the risk of test-taker*

ДОДАТКИ

cheating (e.g., having another person to take the test as a proxy). Where an assessment is carried out in open or controlled mode, checks against cheating can be carried out by requiring the test-taker to undertake a subsequent validation assessment in proctored conditions (i.e. supervised or managed conditions) and a comparison of scores made.

3. *Identify the threats to test validity that exist if test control is not maintained properly.*
4. *Provide advice on the design and implementation of 'honesty (honor) policies' in assessment procedures if one or more stages of the process are to be carried out without direct human supervision.*

45. Test Users

1. *Ensure test-takers provide the appropriate level of authentication before testing begins. Remind test-takers (in the Controlled mode) of the need to obtain a password and username to access the test. In supervised and managed testing conditions, test-takers should be required to provide authentic, government approved picture identification.*
2. *For moderate or high stakes testing confirm that procedures are in place to reduce the opportunity for cheating. Technological features may be used where appropriate and feasible (e.g., Closed Circuit Television, CCTV) but it is likely that such testing will require the presence of a test administrator, a follow-up supervised assessment, or a face to face feedback session (e.g., for post-sift assessment in job selection situations).*
3. *For moderate and high stakes assessment (e.g., job recruitment and selection), where individuals are permitted to take a test in controlled mode (i.e. at their convenience in non-secure locations), those obtaining qualifying scores should be required to take a supervised test to confirm their scores.*
 - *Procedures should be used to check whether the test-taker's original responses are consistent with the responses from the confirmation test.*
 - *Test-takers should be informed in advance of these procedures and asked to confirm that they will complete the tests according to instructions given (e.g. not seek assistance, not collude with others etc).*
 - *This agreement may be represented in the form of an explicit honesty policy which the test-taker is required to accept.*
4. *Provide test-takers with a list of expectations and consequences for fraudulent test taking practices, and require test-takers to accept or sign the agreement form indicating their commitment.*

4. Make appropriate provision for security and safeguarding privacy in CBT and Internet testing

a. Take account of the security of test materials

46. Test Developers

1. *Design features into the CBT/Internet system that minimise the risk of test items, scoring keys, and interpretation algorithms being illegitimately printed, downloaded, copied, or sent electronically to another computer. For example, software can be developed that controls browser function by disabling access to menu selections (such as copy, paste).*
2. *Design features into the system (e.g., firewalls) that protects the CBT/Internet test system and associated databases from illegal hacking and computer viruses.*

47. Test Publishers

1. *Protect sensitive features of the test from illegitimate disclosure. For Internet testing, all important intellectual property (e.g., scoring rules, norms, interpretation algorithms) associated with a test should remain on the host server. Only test items and the outputs from report generators usually should appear on the test user's or test-taker's screens.*
2. *Where appropriate, develop a policy that limits test material access to qualified and authorised test users and testing centres. For example, when testing over the Internet, test users would need to obtain and use a password before they were able to access test materials or set up an assessment for a test-taker.*
3. *Passwords should be issued only to users qualified to use the Internet test.*
4. *Verify and check that the CBT/Internet test has features to protect it from illegal hacking and computer viruses. Confirm for Internet testing that reasonable steps have been taken to prevent servers from being accessed by unauthorised or illegal means.*
5. *For Internet testing, maintain control over the sensitive features of the test and report copyright violations on the Internet. Monitor the web for illegal versions, old/outdated versions and part versions of the Internet test and take steps (e.g., enforcing copyright law) to eliminate these violations.*
6. *Take steps to secure protection of test content under existing laws.*
7. *Take appropriate measures to identify stolen test material on the Internet and to estimate its impact of its distribution on the testing program.*

ДОДАТКИ

8. *Take appropriate measures to control the distribution of stolen test material on the Internet including notification of appropriate legal authorities.*
9. *Maintain a process for the adjudication of security breach allegations and specify appropriate sanctions.*

48. Test Users

1. *Know the features that have been developed to ensure the security of test materials, and develop procedures that reduce unauthorised access to such materials.*
2. *Respect the sensitive nature of test materials and intellectual property rights of test publishers/developers.*
3. *Protect test materials from being copied, printed, or otherwise reproduced without the prior written permission of the holder of the copyright.*
4. *Protect passwords and usernames from becoming known to others who are not authorised or qualified to have them.*
5. *Inform the service provider/publisher of any breach in security.*

b. Consider the security of test-taker's data transferred over the Internet**49. Test Developers**

1. *When designing an Internet test, build in features that safeguard test-taker data and maintain the security of test material transferred over the Internet.*
2. *Make use of proxy servers, where appropriate, and embed transactions within secure socket layers .*
3. *Design data management systems to enable users to access, check, and/or delete data from the server in accordance with local data protection and privacy legislation.*
4. *Design features that ensure regular and frequent backups of all collected data and that allow for recovery of data when problems emerge.*

50. Test Publishers

1. *Maintain the security of test-taker data transmitted over the Internet (e.g. by encryption).*
2. *Ensure that test users and test-takers are informed that the host server has correctly received their data.*
3. *Inform test users of their rights and obligations in relation to local data protection and privacy legislation.*

4. *Conduct regular and frequent backups of all collected data and provide test users with a detailed disaster recovery plan should problems emerge.*

51. Test Users

1. *Prior to test administration, have knowledge of and inform test-takers of the security procedures used to safeguard data transmitted over the internet.*
2. *Confirm with the service provider that they frequently back up data.*
3. *Verify that the service provider is able to allow test users and authorised others to discharge their responsibilities as data controllers under local data protection and privacy legislation (e.g. the European Union's Directive on Data Protection).*

c. Maintain the confidentiality of test-taker results

52. Test Developers

1. *Design features to allow secure storage of CBT/Internet test data on computer, disks or server .*
2. *Maintain the integrity of CBT/Internet test data by providing technology that does not allow unauthorised altering of information and that can detect unauthorised changes to information.*
3. *Devise encryption devices and password protection that restrict access to test data.*

53. Test Publishers

1. *When test data must be stored with publishers, specify the procedures and systems to maintain the confidentiality and security of data.*
2. *Inform test users of who has access to test data, for what purposes, and how long the data will be stored electronically.*
3. *Adhere to country-specific data protection laws/regulations governing the storage of personal data.*
4. *Restrict access to personal data stored on the host server to those who are qualified and authorised.*
5. *Protect all sensitive personal material held on computer, disk, or a server with robust (non-trivial) encryption devices or passwords.*
6. *Confirm the security and confidentiality of the backup data when used to store sensitive personal data.*

54. Test Users

ДОДАТКИ

1. *Know how confidentiality will be maintained when data are stored electronically.*
2. *Adhere to country-specific data protection laws/regulations governing the collection, use, storage and security of personal data.*
3. *Protect all material via the use of encryption or passwords when storing sensitive personal data electronically on test centre facilities.*
4. *Apply the same levels of security and confidentiality to backup data as to the data on the live system when backups are used to store personal data.*

References

- Association of Test Publishers (ATP: 2002). Guidelines for computer-based testing: ATP.
- Bartram, D. (2001). The impact of the Internet on testing: Issues that need to be addressed by a Code of Good Practice. Internal report for SHL Group plc.
- Bartram, D. (2002). Review model for the description and evaluation of psychological tests. European Federation of Psychologists' Associations (EFPA): <http://www.efpa.be>
- British Psychological Society Psychological Testing Centre (2002) Guidelines for the Development and Use of Computer-based Assessments. Leicester: British Psychological Society.
- British Standards' Institute (BSI). BS 7988 (2001). A code of practice for the use of information technology for the delivery of assessments. British Standards' Institute
- Hambleton, R. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment*, 10, 229-244.
- International Test Commission (2001). International guidelines for test use. *International Journal of Testing*, 1, 93-114.
- Naglieri, J. A., Drasgow, F., Schmit, M., Handler, L., Prifitera, A., Margolis, A., and Velasquez, R. (2004). Psychological Testing on the Internet: New Problems, Old Issues. *American Psychologist*, 59, 3, 150-162.
- Van de Vijver, F., & Hambleton, R. (1996). Translating tests: Some practical guidelines. *European Psychologist*, 1, 89-99.

Appendix: Papers consulted in drafting the ITC guidelines on Computer-based and Internet-delivered Testing

Association of Test Publishers (ATP: 2002). Guidelines for computer-based testing: ATP.

Bartram, D. (1985). The automation of psychological testing procedures: Towards some guidelines for management and operation. Paper presented at the Conference on the Management and Operation of Computer-based Testing Procedures: London.

Bartram, D. (1989). Computer-based assessment. In P. Herriot, (Ed.), *Handbook of Assessment in Organisations* (pp. 369-390). Chichester: John Wiley & Sons.

Bartram, D. (1999) Testing and the Internet: Current realities, issues and future possibilities. Keynote paper for the 1999 Test User Conference.

Bartram, D. (2000) Internet recruitment and selection: Kissing frogs to find princes. *International Journal of Selection and Assessment*, 8, 261 — 274.

Bartram, D. (2001). The impact of the Internet on testing: Issues that need to be addressed by a Code of Good Practice. Internal report for SHL Group plc.

Bartram, D. (2002). Review model for the description and evaluation of psychological tests. European Federation of Psychologists' Associations (EFPA): <http://www.efpa.be>

Bartram, D. & Bayliss, R. (1984). Automated testing: Past, present and future. *Journal of Occupational Psychology*, 57, 221 — 237

Bartram, D., Beaumont, J. G., Cornford, T., Dann, P. L., & Wilson, S. L. (1987). Recommendations for the design of software for computer-based assessment — summary statement. *Bulletin of the British Psychological Society*, 40, 86-87

British Psychological Society Psychological Testing Centre (2002). *Guidelines for the Development and Use of Computer-based Assessments*. Leicester: British Psychological Society.

British Standards' Institute (BSI). BS 7988 (2001). A code of practice for the use of information technology for the delivery of assessments. British Standards' Institute

Burke, M. J. & Normand, J. (1987). Computerized psychological testing: Overview and critique. *Professional Psychology: Research and Practice*, 18, 42 — 51.

Farrell, A. D. (1989). Impact of standards for computer-based tests on practice: Consequences of the information gap. *Computers in Human Behavior*, 5, 1 — 11

Fremer, J. (1996). Promoting high standards for test use: Developments in the United States. *European Journal of Psychological Assessment*, 12, 160 — 168

Hofer, P. J. (1986). Developing standards for computerized psychological testing. *Computers in Human Behaviour*, 1, 301 –315.

HUSAT Research Centre. (1988). *Human Factors Guidelines for the design of Computer-Based Systems*. Volume 4. Loughborough University of Technology.

Copyright © ITC 2005 27ITC Computer-based and Internet delivered testing Guidelines (2005)

Kratochwill, T. R., Doll, E. J. & Dickson, W. P. (1985). Microcomputers in behavioral assessment: Recent advances and remaining issues. *Computers in Human Behavior*, 1, 277 — 291.

Matarazzo, J. D. (1985). Clinical psychological test interpretations by computer: Hardware outpaces software. *Computers in Human Behavior*, 1, 235 — 253

Mitchell, J. V. & Kramer, J. J. (1986). Computer-based assessment and the public interest: An examination of the issues and introduction to the special issue. *Computers in Human Behavior*, 1, 203 — 205.

Naglieri, J. A., Drasgow, F., Schmit, M., Handler, L., Prifitera, A., Margolis, A., and Velasquez, R. (2004). Psychological Testing on the Internet: New Problems, Old Issues. *American Psychologist*, 59, 150-162.

Rolls, S., & Feltham, R. (1993). Practical and professional issues in computer-based assessment and interpretation. *International Review of Professional Issues in Selection*, 1, 135 — 146.

Schoenfeldt, L. F. (1989). Guidelines for computer-based psychological tests and interpretations. *Computers in Human Behavior*, 5, 13 — 21.

Glossary

Browser	A program (e.g. Netscape or Internet Explorer) which allows an individual to view pages downloaded from the Internet
Connection speed	The time taken for the browser to connect with the Internet test and download material.
Controlled mode	This is a mode of test administration in which control is exercised over who can access a test on the internet and how often they can access it. It may also include controls over the location they can access it from and the time or date it is available.

Computer-based test interpretations (CBTI)	Reports that are generated using algorithms to detail a test-takers profile from computer/Internet testing. Such interpretations can be tailored to specific stakeholders and can differ in format.
Data management systems	Computer systems that handle file maintenance and disk space allocation, including data storage, record access and the physical location of the files.
Delivery platforms	Combination of hardware and operating system for which the test has been developed. For example, designing a test to work on the Windows platform.
Differential item functioning (DIF)	A difference in responses to a test item (or items) as a result of group membership (e.g. ethnicity) and not as a result of actual differences on the trait being measured.
Embedded tutorial	Help feature that is built into the computer software, which can be accessed by the test taker and which provides a tutorial on how to use the test software.
Encryption device	Feature that allows data to be encrypted during transmission over the Internet
Firewall	A server that protects data from outside or malicious access whilst allowing a reduced level of security precautions for internal communications (e.g. communications within an organisation)
Graphical User Interface (GUI) design standards	Industry standards for the effective design of the interactive outer layer (e.g. operating systems) of computer programmes.
Hang-ups	A failure that results in the computer or internet test not responding. Normally, a reset is required before testing can begin.
Input device	The hardware used to input information (data or commands) into the computer (e.g. keyboard, mouse etc).
Item bank	A large set of items, generally calibrated using Item Response Theory, which can be used for computer adaptive testing or to build up multiple unique Internet tests with known properties.
Item exposure	Repeated presentation of the same item. High levels of item exposure increase the likelihood that item security may be compromised.
Managed mode	A mode of administration in which there is both direct supervision and control over the equipment being used, and other conditions. Typically managed mode administration refers to the use of dedicated testing centres.
Multimedia	A combination of moving and still pictures as well as the use of sound by the computer software.

ДОДАТКИ

Open mode	A mode of administration where the test taker has direct access to the test materials and there is no involvement of a test user or test administrator. Such tests include the books of tests you might buy in the local book shop or the tests you can find on the Internet that are directly accessible to everyone. Often the only requirement is that you pay some money before you can access the test. However, no qualifications are required from you either in terms of test use or test administration.
Proctored	Testing in a supervised or managed format where there is direct human supervision over test-taking conditions.
Proxy servers	A server that acts as an intermediary between a user's computer and the computer they want to access. If a user makes a request for a resource from computer «A,» this request is directed to the proxy server, which makes the request, gets the response from computer «A,» and then forwards the response to the client.
Secure socket layers (SSL)	A set of rules that enable encrypted communications to be made between devices (e.g. computers) over the Internet in a secure manner
Servers	A Server is a computer, or a software application that provides a specific kind of service to client software running on other computers. The term can refer to a particular piece of software, such as a Web Server, or to the computer on which the software is running. A single computer may have several different server software applications running on it, thus providing many different servers to clients on a network.
Supervised mode	This is the mode in which the test administrator has direct face-to-face involvement with the test taker. The test takers will come to a location where the test administrator is able to supervise them taking the test. However, the test distributor has no means of directly controlling the nature of the location or the type of equipment being used.
Test accommodations	Changes or adaptations made to the test (software, input devices, hardware etc) to accommodate individuals with disabilities.
Test Developers	Those involved in designing and creating computer and Internet tests.
Test Publishers	Those involved in selling and marketing computer and Internet tests. Publishers may also be involved in developing tests.
Test-takers	People who complete tests on computer or over the Internet.

Test Users	A broad category of stakeholders who are involved in actually using the computer and Internet tests. This could involve purchasing tests, selecting appropriate tests, administering and scoring the tests, providing feedback on test results to test takers and others.
Unproctored	Testing in an open or controlled format where there is no direct human supervision over test-taking conditions.

Code of Professional Responsibilities in Educational Measurement

BUROS
INSTITUTE

Buros Institute of Mental Measurements

Prepared by the NCME Ad Hoc Committee on the Development of a Code of Ethics:

Cynthia B. Schmeiser, ACT--Chair
Kurt F. Geisinger, State University of New York
Sharon Johnson-Lewis, Detroit Public Schools
Edward D. Roeber, Council of Chief State School Officers
William D. Schafer, University of Maryland
copyright 1995 National Council on Measurement in Education

Any portion of this Code may be reproduced and disseminated for educational purposes.

As an organization dedicated to the improvement of measurement and evaluation practice in education, the National Council on Measurement in Education (NCME) has adopted this Code to promote professionally responsible practice in educational measurement. Professionally responsible practice is conduct that arises from either the professional standards of the field, general ethical principles, or both.

The purpose of the Code of Professional Responsibilities in Educational Measurement, hereinafter referred to as the Code, is to guide the conduct of NCME members who are involved in any tinge of assessment activity in education. NCME is also providing this Code as a public service for all

individuals who are engaged in educational assessment activities in the hope that these activities will be conducted in a professionally responsible manner. Persons who engage in these activities include local educators such as classroom teachers, principals, and superintendents; professionals such as school psychologists and counselors; state and national technical, legislative, and policy staff in education; staff of research, evaluation, and testing organizations; providers of test preparation services; college and university faculty and administrators; and professionals in business and industry who design and implement educational and training programs.

This Code applies to any type of assessment that occurs as part of the educational process, including formal and informal, traditional and alternative techniques for gathering information used in making educational decisions at all levels. These techniques include, but are not limited to, large-scale assessments at the school, district, state, national, and international levels; standardized tests; observational measures; teacher-conducted assessments; assessment support materials; and other achievement, aptitude, interest, and personality measures used in and for education.

Although NCME is promulgating this Code for its members, it strongly encourages other organizations and individuals who engage in educational assessment activities to endorse and abide by the responsibilities relevant to their professions. Because the Code pertains only to uses of assessment in education, it is recognized that uses of assessments outside of educational contexts, such as for employment, certification, or licensure, may involve additional professional responsibilities beyond those detailed in this Code.

The Code is intended to serve an educational function: to inform and remind those involved in educational assessment of their obligations to uphold the integrity of the manner in which assessments are developed, used, evaluated, and marketed. Moreover, it is expected that the Code will stimulate thoughtful discussion of what constitutes professionally responsible assessment practice at all levels in education.

The Code enumerates professional responsibilities in eight major areas of assessment activity. Specifically, the Code presents the professional responsibilities of those who:

- 1) *Develop Assessments*
- 2) *Market and Sell Assessments*
- 3) *Select Assessments*
- 4) *Administer Assessments*
- 5) *Score Assessments*
- 6) *Interpret, Use, and Communicate Assessment Results*

- 7) *Educate About Assessment*
- 8) *Evaluate Programs and Conduct Research on Assessments*

Although the organization of the Code is based on the differentiation of these activities, they are viewed as highly interrelated, and those who use this Code are urged to consider the Code in its entirety. The index following this Code provides a listing of some of the critical interest topics within educational measurement that focus on one or more of the assessment activities.

General Responsibilities

The professional responsibilities promulgated in this Code in eight major areas of assessment activity are based on expectations that NCME members involved in educational assessment will:

- 1) *protect the health and safety of all examinees;*
- 2) *be knowledgeable about, and behave in compliance with, state and federal laws relevant to the conduct of professional activities;*
- 3) *maintain and improve their professional competence in educational assessment;*
- 4) *provide assessment services only in areas of their competence and experience, affording full disclosure of their professional qualifications;*
- 5) *promote the understanding of sound assessment practices in education;*
- 6) *adhere to the highest standards of conduct and promote professionally responsible conduct within educational institutions and agencies that provide educational services; and*
- 7) *perform all professional responsibilities with honesty, integrity, due care, and fairness.*

Responsible professional practice includes being informed about and acting in accordance with the **Code of Fair Testing Practices in Education** (Joint Committee on Testing Practices, 1988), the **Standards for Educational and Psychological Testing** (American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 1985), or subsequent revisions, as well as all applicable state and federal laws that may govern the development, administration, and use of assessments. Both the **Standards for Educational and Psychological Testing** and the **Code of Fair Testing Practices in Education** are intended to establish criteria for judging the technical adequacy of tests and the appropriate uses of tests and test results. The purpose of this Code is to describe the professional responsibilities of those individuals who are engaged in assessment activities. As would be expected, there is a strong relationship between professionally responsible practice and sound

educational assessments, and this Code is intended to be consistent with the relevant parts of both of these documents.

It is not the intention of NCME to enforce the professional responsibilities stated in the Code or to investigate allegations of violations to the Code.

Since the Code provides a frame of reference for the evaluation of the appropriateness of behavior, NCME recognizes that the Code may be used in legal or other similar proceedings.

Section 1: Responsibilities of Those Who Develop Assessment Products and Services

Those who develop assessment products and services, such as classroom teachers and other assessment specialists, have a professional responsibility to strive to produce assessments that are of the highest quality. Persons who develop assessments have a professional responsibility to:

- 1.1 Ensure that assessment products and services are developed to meet applicable professional, technical, and legal standards.*
- 1.2 Develop assessment products and services that are as free as possible from bias due to characteristics irrelevant to the construct being measured, such as gender, ethnicity, race, socioeconomic status, disability, religion, age, or national origin.*
- 1.3 Plan accommodations for groups of test takers with disabilities and other special needs when developing assessments.*
- 1.4 Disclose to appropriate parties any actual or potential conflicts of interest that might influence the developers' judgment or performance.*
- 1.5 Use copyrighted materials in assessment products and services in accordance with state and federal law.*
- 1.6 Make information available to appropriate persons about the steps taken to develop and score the assessment, including up-to-date information used to support the reliability, validity, scoring and reporting processes, and other relevant characteristics of the assessment.*
- 1.7 Protect the rights to privacy of those who are assessed as part of the assessment development process.*
- 1.8 Caution users, in clear and prominent language, against the most likely misinterpretations and misuses of data that arise out of the assessment development process.*
- 1.9 Avoid false or unsubstantiated claims in test preparation and program support materials and services about an assessment or its use and interpretation.*

- 1.10 *Correct any substantive inaccuracies in assessments or their support materials as soon as feasible.*
- 1.11 *Develop score reports and support materials that promote the understanding of assessment results.*

Section 2: Responsibilities of Those Who Market and Sell Assessment Products and Services

The marketing of assessment products and services, such as tests and other instruments, scoring services, test preparation services, consulting, and test interpretive services, should be based on information that is accurate, complete, and relevant to those considering their use. Persons who market and sell assessment products and services have a professional responsibility to:

- 2.1 *Provide accurate information to potential purchasers about assessment products and services and their recommended uses and limitations.*
- 2.2 *Not knowingly withhold relevant information about assessment products and services that might affect an appropriate selection decision.*
- 2.3 *Base all claims about assessment products and services on valid interpretations of publicly available information.*
- 2.4 *Allow qualified users equal opportunity to purchase assessment products and services.*
- 2.5 *Establish reasonable fees for assessment products and services.*
- 2.6 *Communicate to potential users, in advance of any purchase or use, all applicable fees associated with assessment products and services.*
- 2.7 *Strive to ensure that no individuals are denied access to opportunities because of their inability to pay the fees for assessment products and services.*
- 2.8 *Establish criteria for the sale of assessment products and services, such as limiting the sale of assessment products and services to those individuals who are qualified for recommended uses and from whom proper uses and interpretations are anticipated.*
- 2.9 *Inform potential users of known inappropriate uses of assessment products and services and provide recommendations about how to avoid such misuses.*
- 2.10 *Maintain a current understanding about assessment products and services and their appropriate uses in education.*
- 2.11 *Release information implying endorsement by users of assessment products and services only with the users' permission.*

ДОДАТКИ

- 2.12 *Avoid making claims that assessment products and services have been endorsed by another organization unless an official endorsement has been obtained.*
- 2.13 *Avoid marketing test preparation products and services that may cause individuals to receive scores that misrepresent their actual levels of attainment.*

Section 3: Responsibilities of Those Who Select Assessment Products and Services

Those who select assessment products and services for use in educational settings, or help others do so, have important professional responsibilities to make sure that the assessments are appropriate for their intended use. Persons who select assessment products and services have a professional responsibility to:

- 3.1 *Conduct a thorough review and evaluation of available assessment strategies and instruments that might be valid for the intended uses.*
- 3.2 *Recommend and/or select assessments based on publicly available documented evidence of their technical quality and utility rather than on unsubstantiated claims or statements.*
- 3.3 *Disclose any associations or affiliations that they have with the authors, test publishers, or others involved with the assessments under consideration for purchase and refrain from participation if such associations might affect the objectivity of the selection process.*
- 3.4 *Inform decision makers and prospective users of the appropriateness of the assessment for the intended uses, likely consequences of use, protection of examinee rights, relative costs, materials and services needed to conduct or use the assessment, and known limitations of the assessment, including potential misuses and misinterpretations of assessment information.*
- 3.5 *Recommend against the use of any prospective assessment that is likely to be administered, scored, and used in an invalid manner for members of various groups in our society for reasons of race, ethnicity, gender, age, disability, language background, socioeconomic status, religion, or national origin.*
- 3.6 *Comply with all security precautions that may accompany assessments being reviewed.*
- 3.7 *Immediately disclose any attempts by others to exert undue influence on the assessment selection process.*

- 3.8 Avoid recommending, purchasing, or using test preparation products and services that may cause individuals to receive scores that misrepresent their actual levels of attainment.*

Section 4: Responsibilities of Those Who Administer Assessments

Those who prepare individuals to take assessments and those who are directly or indirectly involved in the administration of assessments as part of the educational process, including teachers, administrators, and assessment personnel, have an important role in making sure that the assessments are administered in a fair and accurate manner. Persons who prepare others for, and those who administer, assessments have a professional responsibility to:

- 4.1 Inform the examinees about the assessment prior to its administration, including its purposes, uses, and consequences; how the assessment information will be judged or scored; how the results will be kept on file; who will have access to the results; how the results will be distributed; and examinees' rights before, during, and after the assessment.*
- 4.2 Administer only those assessments for which they are qualified by education, training, licensure, or certification.*
- 4.3 Take appropriate security precautions before, during, and after the administration of the assessment.*
- 4.4 Understand the procedures needed to administer the assessment prior to administration.*
- 4.5 Administer standardized assessments according to prescribed procedures and conditions and notify appropriate persons if any nonstandard or delimiting conditions occur.*
- 4.6 Not exclude any eligible student from the assessment.*
- 4.7 Avoid any conditions in the conduct of the assessment that might invalidate the results.*
- 4.8 Provide for and document all reasonable and allowable accommodations for the administration of the assessment to persons with disabilities or special needs.*
- 4.9 Provide reasonable opportunities for individuals to ask questions about the assessment procedures or directions prior to and at prescribed times during the administration of the assessment.*
- 4.10 Protect the rights to privacy and due process of those who are assessed.*
- 4.11 Avoid actions or conditions that would permit or encourage individuals or groups to receive scores that misrepresent their actual levels of attainment.*

Section 5: Responsibilities of Those Who Score Assessments

The scoring of educational assessments should be conducted properly and efficiently so that the results are reported accurately and in a timely manner. Persons who score and prepare reports of assessments have a professional responsibility to:

- 5.1 Provide complete and accurate information to users about how the assessment is scored, such as the reporting schedule, scoring process to be used, rationale for the scoring approach, technical characteristics, quality control procedures, reporting formats, and the fees, if any, for these services.*
- 5.2 Ensure the accuracy of the assessment results by conducting reasonable quality control procedures before, during, and after scoring.*
- 5.3 Minimize the effect on scoring of factors irrelevant to the purposes of the assessment.*
- 5.4 Inform users promptly of any deviation in the planned scoring and reporting service or schedule and negotiate a solution with users.*
- 5.5 Provide corrected score results to the examinee or the client as quickly as practicable should errors be found that may affect the inferences made on the basis of the scores.*
- 5.6 Protect the confidentiality of information that identifies individuals as prescribed by state and federal laws.*
- 5.7 Release summary results of the assessment only to those persons entitled to such information by state or federal law or those who are designated by the party contracting for the scoring services.*
- 5.8 Establish, where feasible, a fair and reasonable process for appeal and rescoring the assessment.*

Section 6: Responsibilities of Those Who Interpret, Use, and Communicate Assessment Results

The interpretation, use, and communication of assessment results should promote valid inferences and minimize invalid ones. Persons who interpret, use, and communicate assessment results have a professional responsibility to:

- 6.1 Conduct these activities in an informed, objective, and fair manner within the context of the assessment's limitations and with an understanding of the potential consequences of use.*
- 6.2 Provide to those who receive assessment results information about the assessment, its purposes, its limitations, and its uses necessary for the proper interpretation of the results.*

- 6.3 *Provide to those who receive score reports an understandable written description of all reported scores, including proper interpretations and likely misinterpretations.*
- 6.4 *Communicate to appropriate audiences the results of the assessment in an understandable and timely manner, including proper interpretations and likely misinterpretations.*
- 6.5 *Evaluate and communicate the adequacy and appropriateness of any norms or standards used in the interpretation of assessment results.*
- 6.6 *Inform parties involved in the assessment process how assessment results may affect them.*
- 6.7 *Use multiple sources and types of relevant information about persons or programs whenever possible in making educational decisions.*
- 6.8 *Avoid making, and actively discourage others from making, inaccurate reports, unsubstantiated claims, inappropriate interpretations, or otherwise false and misleading statements about assessment results.*
- 6.9 *Disclose to examinees and others whether and how long the results of the assessment will be kept on file, procedures for appeal and rescoring, rights examinees and others have to the assessment information, and how those rights may be exercised.*
- 6.10 *Report any apparent misuses of assessment information to those responsible for the assessment process.*
- 6.11 *Protect the rights to privacy of individuals and institutions involved in the assessment process.*

Section 7: Responsibilities of Those Who Educate Others About Assessment

The process of educating others about educational assessment, whether as part of higher education, professional development, public policy discussions, or job training, should prepare individuals to understand and engage in sound measurement practice and to become discerning users of tests and test results. Persons who educate or inform others about assessment have a professional responsibility to:

- 7.1 *Remain competent and current in the areas in which they teach and reflect that in their instruction.*
- 7.2 *Provide fair and balanced perspectives when teaching about assessment.*
- 7.3 *Differentiate clearly between expressions of opinion and substantiated knowledge when educating others about any specific assessment method, product, or service.*

ДОДАТКИ

- 7.4 Disclose any financial interests that might be perceived to influence the evaluation of a particular assessment product or service that is the subject of instruction.*
- 7.5 Avoid administering any assessment that is not part of the evaluation of student performance in a course if the administration of that assessment is likely to harm any student.*
- 7.6 Avoid using or reporting the results of any assessment that is not part of the evaluation of student performance in a course if the use or reporting of results is likely to harm any student.*
- 7.7 Protect all secure assessments and materials used in the instructional process.*
- 7.8 Model responsible assessment practice and help those receiving instruction to learn about their professional responsibilities in educational measurement.*
- 7.9 Provide fair and balanced perspectives on assessment issues being discussed by policymakers, parents, and other citizens.*

Section 8: Responsibilities of Those Who Evaluate Educational Programs and Conduct Research on Assessments

Conducting research on or about assessments or educational programs is a key activity in helping to improve the understanding and use of assessments and educational programs. Persons who engage in the evaluation of educational programs or conduct research on assessments have a professional responsibility to:

- 8.1 Conduct evaluation and research activities in an informed, objective, and fair manner.*
- 8.2 Disclose any associations that they have with authors, test publishers, or others involved with the assessment and refrain from participation if such associations might affect the objectivity of the research or evaluation.*
- 8.3 Preserve the security of all assessments throughout the research process as appropriate.*
- 8.4 Take appropriate steps to minimize potential sources of invalidity in the research and disclose known factors that may bias the results of the study.*
- 8.5 Present the results of research, both intended and unintended, in a fair, complete, and objective manner.*
- 8.6 Attribute completely and appropriately the work and ideas of others.*
- 8.7 Qualify the conclusions of the research within the limitations of the study.*

8.8 Use multiple sources of relevant information in conducting evaluation and research activities whenever possible.

8.9 Comply with applicable standards for protecting the rights of participants in an evaluation or research study, including the rights to privacy and informed consent.

Afterword

As stated at the outset, the purpose of the **Code of Professional Responsibilities in Educational Measurement** is to serve as a guide to the conduct of NCME members who are engaged in any type of assessment activity in education. Given the broad scope of the field of educational assessment as well as the variety of activities in which professionals may engage, it is unlikely that any code will cover the professional responsibilities involved in every situation or activity in which assessment is used in education. Ultimately, it is hoped that this Code will serve as the basis for ongoing discussions about what constitutes professionally responsible practice. Moreover, these discussions will undoubtedly identify areas of practice that need further analysis and clarification in subsequent editions of the Code. To the extent that these discussions occur, the Code will have served its purpose.

To assist in the ongoing refinement of the Code, comments on this document are most welcome. Please send your comments and inquiries to:

Dr. William J. Russell, Executive Officer, National Council on Measurement in Education, 1230 Seventeenth Street, NW, Washington, DC 20036-3078

Supplementary Resources

The following list of resources is provided for those who want to seek additional information about codes of professional responsibility that have been developed and adopted by organizations having an interest in various aspects of educational assessment.

American Association for Counseling and Development (now American Counseling Association). (1988). *Ethical standards of the American Counseling Association*. Alexandria, VA: Author.

American Association for Counseling and Development (now American Counseling Association) & Association for Measurement and Evaluation in Counseling and Development (now Association for Assessment in Counseling). (1989). *Responsibilities of users of standardized tests: RUST statement revised*. Alexandria, VA: Author.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1985).

Standards for educational and psychological testing. Washington, DC: Author. American Educational Research Association. (1992). Ethical standards of the American Educational Research Association. *Educational Researcher*, **21(7)**, 23-26.

American Federation of Teachers, National Council on Measurement in Education, & National Education Association. (1990). *Standards for teacher competence in educational assessment of students.* Washington, DC: Author.

American Psychological Association. (1992). *Ethical principles of psychologists and code of conduct.* Washington, DC: Author.

American Psychological Association President's Task Force on Psychology in Education. (in press). *Learner-centered psychological principles: Guidelines for school redesign and reform.* Washington, DC: Author.

Joint Advisory Committee. (1993). *Principles for fair assessment practices for education in Canada.* Edmonton, Alberta: Author.

Joint Committee on Testing Practices. (1988). *Code of fair testing practices in education.* Washington, DC: Author.

Joint Committee on Standards for Educational Evaluation. (1988). *The personnel evaluation standards: How to assess systems for evaluating educators.* Newberry Park, CA: Sage.

Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards: How to assess evaluations of educational programs.* Thousand Oaks, CA: Sage.

National Association of College Admission Counselors. (1988). *Statement of principles of good practice.* Alexandria, VA: Author.

Index to the Code of Professional Responsibilities in Educational Measurement

This index provides a list of major topics and issues addressed by the responsibilities in each of the eight sections of the Code. Although this list is not intended to be exhaustive, it is intended to serve as a reference source for those who use this code.

Advertising: 1.9, 1.10, 2.3, 2.11, 2.12

Bias: 1.2, 3.5, 4.5, 4.7, 5.3, 8.4

Cheating: 4.5, 4.6, 4.11

Coaching and Test Preparation: 2.13, 3.8, 4.11

Competence: 2.10, 4.2, 4.4, 4.5, 5.2, 5.5, 7.1, 7.8, 7.9, 8.1, 8.7

Conflict of Interest: 1.4, 3.3, 7.4, 8.2

Consequences of Test Use: 3.4, 6.1, 6.6, 7.5, 7.6

Copyrighted Materials, Use of: 1.5, 8.6

Disabled Examinees, Rights of: 1.3, 4.8

Disclosure: 1.6, 2.1, 2.2, 2.6, 3.3, 3.7, 4.1, 5.1, 5.4, 6.2, 6.3, 6.4, 6.6, 6.9, 8.2, 8.4, 8.5

Due Process: 4.10, 5.8, 6.9

Equity: 1.2, 2.4, 2.7, 3.5, 4.6

Fees: 2.5, 2.6, 2.7

Inappropriate Test Use: 1.8, 2.8, 2.9, 3.4, 6.8, 6.10

Objectivity: 3.1, 3.2, 3.3, 6.1, 6.5, 7.2, 7.3, 7.9, 8.1, 8.2, 8.5, 8.7

Rights to Privacy: 1.7, 3.4, 4.10, 5.6, 5.7, 6.11, 8.9

Security: 3.6, 4.3, 7.7, 8.3

Truthfulness: 1.10, 2.1, 2.2, 2.3, 2.11, 2.12, 3.2, 4.6, 7.3

Undue Influence: 3.7

Unsubstantiated Claims: 1.9, 3.2, 6.8

Standards for Teacher Competence in Educational Assessment of Students

Developed by the:

American Federation of Teachers

National Council on Measurement in Education

National Education Association

This is not copyrighted material. Reproduction and dissemination are encouraged. 1990

The professional education associations began working in 1987 to develop standards for teacher competence in student assessment out of concern that the potential educational benefits of student assessments be fully realized. The Committee⁷⁸ appointed to this project completed its work in 1990 following

⁷⁸ The Committee that developed this statement was appointed by the collaborating professional associations: James R. Sanders (Western Michigan University) chaired the Committee and represented NCME along with John R. Hills (Florida State University) and Anthony J. Nitko (University of Pittsburgh). Jack C. Merwin (University of Minnesota) represented the American Association of Colleges for Teacher Education, Carolyn Trice represented the American Federation of Teachers, and Marcella Dianda and Jeffrey Schneider represented the National Education Association.

reviews of earlier drafts by members of the measurement, teaching, and teacher preparation and certification communities. Parallel committees of affected associations are encouraged to develop similar statements of qualifications for school administrators, counselors, testing directors, supervisors, and other educators in the near future. These statements are intended to guide the preservice and inservice preparation of educators, the accreditation of preparation programs, and the future certification of all educators.

A standard is defined here as a principle generally accepted by the professional associations responsible for this document. Assessment is defined as the process of obtaining information that is used to make educational decisions about students, to give feedback to the student about his or her progress, strengths, and weaknesses, to judge instructional effectiveness and curricular adequacy, and to inform policy. The various assessment techniques include, but are not limited to, formal and informal observation, qualitative analysis of pupil performance and products, paper-and-pencil tests, oral questioning, and analysis of student records. The assessment competencies included here are the knowledge and skills critical to a teacher's role as educator. It is understood that there are many competencies beyond assessment competencies which teachers must possess.

By establishing standards for teacher competence in student assessment, the associations subscribe to the view that student assessment is an essential part of teaching and that good teaching cannot exist without good student assessment. Training to develop the competencies covered in the standards should be an integral part of preservice preparation. Further, such assessment training should be widely available to practicing teachers through staff development programs at the district and building levels.

The standards are intended for use as:

a guide for teacher educators as they design and approve programs for teacher preparation

a self-assessment guide for teachers in identifying their needs for professional development in student assessment

a guide for workshop instructors as they design professional development experiences for in-service teachers

an impetus for educational measurement specialists and teacher trainers to conceptualize student assessment and teacher training in student assessment more broadly than has been the case in the past.

The standards should be incorporated into future teacher training and certification programs. Teachers who have not had the preparation these standards imply should have the opportunity and support to develop

these competencies before the standards enter into the evaluation of these teachers.

The Approach Used To Develop The Standards

The members of the associations that supported this work are professional educators involved in teaching, teacher education, and student assessment. Members of these associations are concerned about the inadequacy with which teachers are prepared for assessing the educational progress of their students, and thus sought to address this concern effectively. A committee named by the associations first met in September 1987 and affirmed its commitment to defining standards for teacher preparation in student assessment. The committee then undertook a review of the research literature to identify needs in student assessment, current levels of teacher training in student assessment, areas of teacher activities requiring competence in using assessments, and current levels of teacher competence in student assessment.

The members of the committee used their collective experience and expertise to formulate and then revise statements of important assessment competencies. Drafts of these competencies went through several revisions by the Committee before the standards were released for public review. Comments by reviewers from each of the associations were then used to prepare a final statement.

The Scope of a Teacher's Professional Role and Responsibilities for Student Assessment

There are seven standards in this document. In recognizing the critical need to revitalize classroom assessment, some standards focus on classroom-based competencies. Because of teachers' growing roles in education and policy decisions beyond the classroom, other standards address assessment competencies underlying teacher participation in decisions related to assessment at the school, district, state, and national levels.

The scope of a teacher's professional role and responsibilities for student assessment may be described in terms of the following activities. These activities imply that teachers need competence in student assessment and sufficient time and resources to complete them in a professional manner.

Activities Occurring Prior to Instruction

- a) *Understanding students' cultural backgrounds, interests, skills, and abilities as they apply across a range of learning domains and/or subject areas;*
- b) *understanding students' motivations and their interests in specific class content;*

ДОДАТКИ

- c) *clarifying and articulating the performance outcomes expected of pupils; and*
- d) *planning instruction for individuals or groups of students.*

Activities Occurring During Instruction

- a) *Monitoring pupil progress toward instructional goals;*
- b) *identifying gains and difficulties pupils are experiencing in learning and performing;*
- c) *adjusting instruction;*
- d) *giving contingent, specific, and credible praise and feedback;*
- e) *motivating students to learn; and*
- f) *judging the extent of pupil attainment of instructional outcomes.*

Activities Occurring After The Appropriate Instructional Segment (e.g. lesson, class, semester, grade)

- a) *Describing the extent to which each pupil has attained both short- and long-term instructional goals;*
- b) *communicating strengths and weaknesses based on assessment results to students, and parents or guardians;*
- c) *recording and reporting assessment results for school-level analysis, evaluation, and decision-making;*
- d) *analyzing assessment information gathered before and during instruction to understand each students' progress to date and to inform future instructional planning;*
- e) *evaluating the effectiveness of instruction; and*
- f) *evaluating the effectiveness of the curriculum and materials in use.*

Activities Associated With a Teacher's Involvement in School Building and School District Decision-Making

- a) *Serving on a school or district committee examining the school's and district's strengths and weaknesses in the development of its students;*
- b) *working on the development or selection of assessment methods for school building or school district use;*
- c) *evaluating school district curriculum; and*
- d) *other related activities.*

Activities Associated With a Teacher's Involvement in a Wider Community of Educators

- a) *Serving on a state committee asked to develop learning goals and associated assessment methods;*

- b) participating in reviews of the appropriateness of district, state, or national student goals and associated assessment methods; and*
- c) interpreting the results of state and national student assessment programs.*

Each standard that follows is an expectation for assessment knowledge or skill that a teacher should possess in order to perform well in the five areas just described. As a set, the standards call on teachers to demonstrate skill at selecting, developing, applying, using, communicating, and evaluating student assessment information and student assessment practices. A brief rationale and illustrative behaviors follow each standard.

The standards represent a conceptual framework or scaffolding from which specific skills can be derived. Work to make these standards operational will be needed even after they have been published. It is also expected that experience in the application of these standards should lead to their improvement and further development.

Standards for Teacher Competence in Educational Assessment of Students

1. Teachers should be skilled in choosing assessment methods appropriate for instructional decisions.

Skills in choosing appropriate, useful, administratively convenient, technically adequate, and fair assessment methods are prerequisite to good use of information to support instructional decisions. Teachers need to be well-acquainted with the kinds of information provided by a broad range of assessment alternatives and their strengths and weaknesses. In particular, they should be familiar with criteria for evaluating and selecting assessment methods in light of instructional plans.

Teachers who meet this standard will have the conceptual and application skills that follow. They will be able to use the concepts of assessment error and validity when developing or selecting their approaches to classroom assessment of students. They will understand how valid assessment data can support instructional activities such as providing appropriate feedback to students, diagnosing group and individual learning needs, planning for individualized educational programs, motivating students, and evaluating instructional procedures. They will understand how invalid information can affect instructional decisions about students. They will also be able to use

and evaluate assessment options available to them, considering among other things, the cultural, social, economic, and language backgrounds of students. They will be aware that different assessment approaches can be incompatible with certain instructional goals and may impact quite differently on their teaching.

Teachers will know, for each assessment approach they use, its appropriateness for making decisions about their pupils. Moreover, teachers will know of where to find information about and/or reviews of various assessment methods. Assessment options are diverse and include text- and curriculum-embedded questions and tests, standardized criterion-referenced and norm-referenced tests, oral questioning, spontaneous and structured performance assessments, portfolios, exhibitions, demonstrations, rating scales, writing samples, paper-and-pencil tests, seatwork and homework, peer- and self-assessments, student records, observations, questionnaires, interviews, projects, products, and others' opinions.

2. Teachers should be skilled in developing assessment methods appropriate for instructional decisions.

While teachers often use published or other external assessment tools, the bulk of the assessment information they use for decision-making comes from approaches they create and implement. Indeed, the assessment demands of the classroom go well beyond readily available instruments.

Teachers who meet this standard will have the conceptual and application skills that follow. Teachers will be skilled in planning the collection of information that facilitates the decisions they will make. They will know and follow appropriate principles for developing and using assessment methods in their teaching, avoiding common pitfalls in student assessment. Such techniques may include several of the options listed at the end of the first standard. The teacher will select the techniques which are appropriate to the intent of the teacher's instruction.

Teachers meeting this standard will also be skilled in using student data to analyze the quality of each assessment technique they use. Since most teachers do not have access to assessment specialists, they must be prepared to do these analyses themselves.

3. The teacher should be skilled in administering, scoring and interpreting the results of both externally-produced and teacher-produced assessment methods.

It is not enough that teachers are able to select and develop good assessment methods; they must also be able to apply them properly.

Teachers should be skilled in administering, scoring, and interpreting results from diverse assessment methods.

Teachers who meet this standard will have the conceptual and application skills that follow. They will be skilled in interpreting informal and formal teacher-produced assessment results, including pupils' performances in class and on homework assignments. Teachers will be able to use guides for scoring essay questions and projects, stencils for scoring response-choice questions, and scales for rating performance assessments. They will be able to use these in ways that produce consistent results.

Teachers will be able to administer standardized achievement tests and be able to interpret the commonly reported scores: percentile ranks, percentile band scores, standard scores, and grade equivalents. They will have a conceptual understanding of the summary indexes commonly reported with assessment results: measures of central tendency, dispersion, relationships, reliability, and errors of measurement.

Teachers will be able to apply these concepts of score and summary indices in ways that enhance their use of the assessments that they develop. They will be able to analyze assessment results to identify pupils' strengths and errors. If they get inconsistent results, they will seek other explanations for the discrepancy or other data to attempt to resolve the uncertainty before arriving at a decision. They will be able to use assessment methods in ways that encourage students' educational development and that do not inappropriately increase students' anxiety levels.

4. Teachers should be skilled in using assessment results when making decisions about individual students, planning teaching, developing curriculum, and school improvement.

Assessment results are used to make educational decisions at several levels: in the classroom about students, in the community about a school and a school district, and in society, generally, about the purposes and outcomes of the educational enterprise. Teachers play a vital role when participating in decision-making at each of these levels and must be able to use assessment results effectively.

Teachers who meet this standard will have the conceptual and application skills that follow. They will be able to use accumulated assessment information to organize a sound instructional plan for facilitating students' educational development. When using assessment results to plan and/or evaluate instruction and curriculum, teachers will interpret the results correctly and avoid common misinterpretations, such as basing decisions on scores that lack curriculum validity. They will be informed about the results of local, regional,

state, and national assessments and about their appropriate use for pupil, classroom, school, district, state, and national educational improvement.

5. Teachers should be skilled in developing valid pupil grading procedures which use pupil assessments.

Grading students is an important part of professional practice for teachers. Grading is defined as indicating both a student's level of performance and a teacher's valuing of that performance. The principles for using assessments to obtain valid grades are known and teachers should employ them.

Teachers who meet this standard will have the conceptual and application skills that follow. They will be able to devise, implement, and explain a procedure for developing grades composed of marks from various assignments, projects, inclass activities, quizzes, tests, and/or other assessments that they may use. Teachers will understand and be able to articulate why the grades they assign are rational, justified, and fair, acknowledging that such grades reflect their preferences and judgments. Teachers will be able to recognize and to avoid faulty grading procedures such as using grades as punishment. They will be able to evaluate and to modify their grading procedures in order to improve the validity of the interpretations made from them about students' attainments.

6. Teachers should be skilled in communicating assessment results to students, parents, other lay audiences, and other educators.

Teachers must routinely report assessment results to students and to parents or guardians. In addition, they are frequently asked to report or to discuss assessment results with other educators and with diverse lay audiences. If the results are not communicated effectively, they may be misused or not used. To communicate effectively with others on matters of student assessment, teachers must be able to use assessment terminology appropriately and must be able to articulate the meaning, limitations, and implications of assessment results. Furthermore, teachers will sometimes be in a position that will require them to defend their own assessment procedures and their interpretations of them. At other times, teachers may need to help the public to interpret assessment results appropriately.

Teachers who meet this standard will have the conceptual and application skills that follow. Teachers will understand and be able to give appropriate explanations of how the interpretation of student assessments must be moderated by the student's socio-economic, cultural, language, and other background factors. Teachers will be able to explain that assessment results do not imply that such background factors limit a student's ultimate educational development. They will be able to communicate to students

and to their parents or guardians how they may assess the student's educational progress. Teachers will understand and be able to explain the importance of taking measurement errors into account when using assessments to make decisions about individual students. Teachers will be able to explain the limitations of different informal and formal assessment methods. They will be able to explain printed reports of the results of pupil assessments at the classroom, school district, state, and national levels.

7. Teachers should be skilled in recognizing unethical, illegal, and otherwise inappropriate assessment methods and uses of assessment information.

Fairness, the rights of all concerned, and professional ethical behavior must undergird all student assessment activities, from the initial planning for and gathering of information to the interpretation, use, and communication of the results. Teachers must be well-versed in their own ethical and legal responsibilities in assessment. In addition, they should also attempt to have the inappropriate assessment practices of others discontinued whenever they are encountered. Teachers should also participate with the wider educational community in defining the limits of appropriate professional behavior in assessment.

Teachers who meet this standard will have the conceptual and application skills that follow. They will know those laws and case decisions which affect their classroom, school district, and state assessment practices. Teachers will be aware that various assessment procedures can be misused or overused resulting in harmful consequences such as embarrassing students, violating a student's right to confidentiality, and inappropriately using students' standardized achievement test scores to measure teaching effectiveness.

Invitation to Users

The associations invite comments from users that may be used for improvement of this document. Comments may be sent to:

Teacher Standards in Student Assessment
American Federation of Teachers
555 New Jersey Avenue, NW, Washington, DC 20001
Teacher Standards in Student Assessment
National Council on Measurement in Education
1230 Seventeenth Street, NW, Washington, DC 20036
Teacher Standards in Student Assessment
Instruction and Professional Development
National Education Association
1201 Sixteenth Street, NW, Washington, DC 20036

Competency Standards in Student Assessment for Educational Administrators

Developed by the:

American Association of School Administrators

National Association of Elementary School Principals

National Association of Secondary School Principals

*National Council on Measurement in Education*⁷⁹

This is not copyrighted material. Reproduction and dissemination are encouraged. 1997.

Background

Student assessment is pervasive in schools, yet few preservice programs for educational professionals require substantive course work in assessment. These standards are intended to guide the preservice and in-service preparation of educational administrators and to inform future standards for the accreditation of administrator preparation programs and the certification of educational administrators. They are intended for both building-level and central office personnel who assist, supervise, or evaluate individuals directly instructing students.

The committee that formulated these competency standards for educational administrators first met in 1991. Draft standards went through several revisions, informed by a sample survey of members of participating organizations⁸⁰, by feedback received in a forum at a professional meeting, and by responses to a draft circulated for public comment.

In these standards the term **assessment** refers to the process of obtaining information about student learning outcomes to guide educational decisions about students; to inform students, their parents, teachers, or other appropriate audiences about their progress, strengths, and weaknesses; to

⁷⁹ In 1990, the American Federation of Teachers, the National Council on Measurement in Education, and the National Education Association published the *Standards for Teacher Competence on Educational Assessment of Students*. The joint committee recommended those standards as a framework for preservice and in-service training for teachers. The committee also recommended that standards be developed for other categories of educational professionals. This document is intended to complement the *Standards for Teacher Competence*.

⁸⁰ More information about this survey can be found in: Impara, J. C., Plake, B. S., & Merwin, J. C. (1994). Student assessment tasks and knowledge: Comparing superintendents and elementary and secondary principals. *Journal of School Leadership*, 4 (5), 517-528.

judge instructional effectiveness and curricular adequacy; and to inform policy. **Assessment methods or techniques** include, but are not limited to, formal and informal observation of student performance, student demonstration of skills and knowledge, qualitative analysis of pupil performance and products, portfolio-based assessments, commercially developed and teacher-developed paper-and-pencil tests, oral questioning, computer-managed or adaptive tests, and analysis of student records.

Overview of the Standards

These standards represent assessment competencies in the administrative contexts in which assessment is relevant to educational administrators. These contexts include: (a) assisting teachers in creating and using assessment effectively; (b) providing leadership in the creation and implementation of building- or district-level assessments policies; and (c) using assessment results in their capacity as administrators in making decisions about students, teachers, and instruction, and in reporting on assessment results to a variety of stakeholders and constituencies.

Educational administrators should have the same student assessment competencies as teachers when they are performing in the context of **assisting teachers** in creating and using assessment. Because of that role the first standard is a summary of the **Standards for Teacher Competence in Educational Assessment of Students**. Educational administrators should have the competencies described in the teachers' standards consistent with the level of direct interaction, supervision, and evaluation of teachers. Administrators should possess, at a minimum, a working knowledge of these competencies. This working level is needed for two reasons. First, administrators must understand why teachers need these competencies; and second, if appropriate, administrators may be required to observe, recognize, record, monitor, and evaluate them in teachers. These competencies are reflected in Standards 1 and 2.

Educational administrators require a somewhat different mix of knowledge, skills, and abilities than teachers when **providing leadership** in the creation and development of assessment policies. Because the field of assessment is changing rapidly, administrators should be able to judge the reasonableness of new assessment techniques proposed for use in their schools and districts. They ought to be able to evaluate the appropriateness of the proposed uses and interpretations of both tradition and new assessment strategies. Standards 3, 4, and 5 are associated with the fundamentals skills of student assessment: the sound principles as assessment, the uses of assessment, and the language of assessment.

Educational administrators need a variety of competencies when **using and reporting** assessment results for decision making. Administrators often communicate assessment results to others and use assessment results to make decisions about individual students, groups of students, teachers and other personnel, curricula, and educational programs. The remaining Standards (6 through 12) apply to various contexts within which administrators and the principal users of assessment results.

Standards

Competencies associated with assisting teachers:

1. Have a working level of competence in the Standards for Teacher Competence in Educational Assessment of Students. These standards are:

- a) choosing assessment methods appropriate for instructional decisions.*
- b) developing assessment methods appropriate for instructional decisions.*
- c) administering, scoring, and interpreting the results of both externally produced and teacher-produced assessment methods.*
- d) using assessment results when making decisions about individual students, planning, teaching, developing curriculum, and school improvement.*
- e) developing valid pupil grading procedures that use pupil assessments.*
- f) communicating assessment results to students, parents, other lay audiences, and other educators.*
- g) recognizing unethical, illegal, and otherwise inappropriate assessment methods and uses of assessment information.*

2. Know the appropriate and useful mechanics of constructing various assessments.

Locally built assessments are pervasive. Teachers construct daily, weekly, and term assessments for their classrooms. Administrators who have supervisory responsibility must be able to determine the quality of the assessment procedures their staff members use in making decisions about students. The growing use of performance assessment, portfolio-based assessment, and computerized testing requires administrators to play a critical role in the proper development and use of these approaches.

Competencies associated with providing leadership in developing and implementing assessment policies:

3. Understand and be able to apply basic measurement principles to assessments conducted in school settings.

Administrators ought to be aware of how these principles apply to school settings so that in their supervisory roles they can support the proper use and interpretation of assessment results. They should understand the distinction between criterion-referenced and norm-referenced test interpretations, and should be able to judge the appropriateness of each kind of interpretation in specific decision contexts. They must understand the principle that validity inheres in the use or interpretation made of a test score, not in the score itself, and should be prepared to exercise leadership both in supporting appropriate uses and in discouraging inappropriate uses of assessment results.

4. Understand the purposes (e.g., description, diagnosis, placement) of different kinds of assessment (e.g., achievement, aptitude, attitude) and the appropriate assessment strategies to obtain the assessment data needed for the intended purpose.

This standard extends Standards 1 and 2 in the *Standards for Teacher Competence in Educational Assessment of Students*, which focus only on making instruction decisions. Most assessment techniques are best used for a single purpose. Therefore, decisions should be based on assessment results consistent with the purpose for which the technique was constructed. Administrators have a responsibility to use assessment results appropriately in each decision context.

5. Understand the need for clear and consistent building- and district-level policies on student assessment.

This extends Standard 5 in the *Standards for Teacher Competence in Educational Assessment of Students*, which focuses only on grading. School- and district-level policies should incorporate grading, final examinations, and other aspects of assessment. Administrators should work to ensure that teachers share a common understanding of the appropriate bases on which to assign marks and grades, and share common standards for the quality of work that merits a given grade. Similarly, administrators should provide opportunities for teachers and others to understand the proper means by which final examinations and other assessments are used in making decisions about students.

Competencies needed in using assessments in making decisions and in communicating assessment results.

6. Understand and express technical assessment concepts and terminology to others in nontechnical but correct ways.

Administrators should understand technical concepts in order to make

decisions about what assessments to use or how to present assessment results to others. As administrators interpret assessment results to various audiences and stakeholders (e.g., teachers, other administrators, parents, school boards, the media), they need to be able to express technical concepts in nontechnical but correct ways. Moreover, decisions such as selecting a standardized test or other assessment procedure often require reading manuals, technical reports, or other technical literature on assessment that use technical concepts.

7. Understand and follow ethical and technical guidelines for assessment.

Because assessment results may be used to make important and sometimes difficult-to-reverse decisions about individual pupils, school personnel, and educational programs, interpretations of results must be valid and consistent with technically sound assessment principles and good educational practice. Administrators should be familiar with available guidelines and standards that describe how to develop, use, and report assessments properly.

The major educational assessment guidelines and standards applicable to administrators' assessment uses at the time these competency standards were developed in 1995 are:

American Federation of Teachers, National Council on Measurement in Education, and the National Education Association. (1990). *Standards for Teacher Competence in Educational Assessments of Students*. Washington, DC: Author.

American Psychological Association, American Educational Research Association, and the National Council on Measurement in Education. (1985). *Standards for Educational and Psychological Testing*. Washington, DC: Author. (Currently under revision.)

Joint Committee on Testing Practices. (1988). *Code of Fair Testing Practices in Education*. Washington, DC: Joint Committee on Testing Practices, American Psychological Association.

8. Reconcile conflicting assessment results appropriately.

Sometimes assessment results appear contradictory, and these apparent contradictions need to be resolved to make appropriate decisions. For example, results from a curriculum-based assessment may indicate strong performance in reading, yet scores from a standardized test in reading may be low; measures of school climate may indicate a positive perception of the school by the students, yet parental complaints about safety may be frequent. Administrators should be able to recognize apparent contradictions in assessment results in light of contextual factors and make appropriate interpretations.

- 9. Recognize the importance, appropriateness, and complexity of interpreting assessment results in light of students' linguistic and cultural backgrounds and other out-of-school factors in light of making accommodations for individual differences, including disabilities, to help ensure the validity of assessment results for all students.**

Assessment results may be influenced by a number of social, cultural, and other factors. These factors may or may not directly cause poor or good performance, but knowledge of these factors often helps in the interpretation of performance on assessment tasks. Administrators should be aware of the kinds of accommodations for administering assessments that may be appropriate in inappropriate for different students under different circumstances.

- 10. Ensure the assessment and information technology are employed appropriately to conduct student assessment.**

Technology includes computer-based assessment tools such as computerized-adaptive testing, computer-managed testing, computerized test item pools, computerized assessment records, and computerized databases. As technology becomes available to more and more schools, administrators will need to make decisions about the use of this technology to conduct student assessments and to store, organize, use, and safeguard the results obtained.

- 11. Use available technology appropriately to integrate assessment results and other student data to facilitate students' learning, instruction, and performance.**

Administrators should be able to organize a wide array of student assessment information with other student data and to make inferences about the quality of performance of students, their teachers, and school programs. Administrators should recognize appropriate and inappropriate uses of assessment results in this context.

- 12. Judge the quality of an assessment strategy or program used for decision making within their jurisdiction.**

Teachers' classroom assessments, school-based testing programs, system-wide testing, and state assessment potentially have an impact on school operations. Administrators should be able to assist teachers and others to make informed judgments about assessment strategies. Administrators ought to be able to undertake or coordinate systematic investigations that correctly inform decisions about the quality of existing and proposed assessment programs. The changing technology of assessment procedures makes this responsibility especially critical. For example, administrators should establish and monitor assessment procedures so that performance

and portfolio assessments are undertaken in technically sound and educationally appropriate ways.

Synthesis of Competency Standards in Student Assessment for Educational Administrators

Competencies associated with assisting teachers:

1. *Have a working level of competence in the Standards for Teacher Competence in Educational Assessment of Students.*
2. *Know the appropriate and useful mechanics of constructing various assessments.*

Competencies associated with providing leadership in developing and implementing assessment policies:

3. *Understand and be able to apply basic measurement principles to assessments conducted in school settings.*
4. *Understand the purposes (e.g., description, diagnosis, placement) of different kinds of assessment (e.g., achievement, aptitude, attitude) and the appropriate assessment strategies to obtain the assessment data needed for the intended purpose.*
5. *Understand the need for clear and consistent building- and district-level policies on student assessment.*

Competencies needed in using assessments in making decisions and in communicating assessment results:

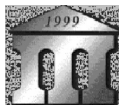
6. *Understand and express technical assessment concepts and terminology to others in nontechnical but correct ways.*
7. *Understand and follow ethical and technical guidelines for assessment.*
8. *Reconcile conflicting assessment results appropriately.*
9. *Recognize the importance, appropriateness, and complexity of interpreting assessment results in light of students' linguistic and cultural backgrounds and other out-of-school factors in light of making accommodations for individual differences, including disabilities, to help ensure the validity of assessment results for all students.*
10. *Ensure the assessment and information technology are employed appropriately to conduct student assessment.*
11. *Use available technology appropriately to integrate assessment results and other student data to facilitate students' learning, instruction, and performance.*
12. *Judge the quality of an assessment strategy or program used for decision making within their jurisdiction.*

ЛІТЕРАТУРА

1. Стандарти освітнього та психологічного тестування — <http://www.apa.org/science/standards.html>.
2. ETS Стандарти якості та справедливості — <http://www.europe.etseurope.org/home-corpora-default/about-ets/quality/>
3. Рекомендації з освітнього та психологічного тестування — <http://www.cpa.ca/documents/PsyTest.html>
4. Кодекс справедливого тестування в освіті — <http://www.apa.org/science/fairtestcode.html>
5. Міжнародні Рекомендації ITC з використання тестів — <http://www.psychtesting.org.uk/files/getfile1.asp?id=8>
6. Міжнародні Рекомендації ITC з адаптації тестів — http://www.intestcom.org/itc_projects.htm#ITC Guidelines on Adapting Tests
7. Міжнародні Рекомендації ITC з комп'ютерного та інтернет-тестування — <http://www.intestcom.org/guidelines/index.html>
8. EFPA Модель аналізу та оцінювання психологічних тестів -- <http://www.efpa.be/doc/EFPA TEST REVIEW FORM and NOTES Verson 3-41-EU.doc>.
9. Кодекс професійної відповідальності при педагогічних вимірюваннях — http://www.natd.org/Code_of_Professional_Responsibilities.html.
10. Стандарти компетентності вчителів з оцінювання учнів — <http://www.unl.edu/buros/bimm/html/article3.html>
11. Стандарти компетентності освітніх адміністраторів з оцінювання учнів — <http://www.unl.edu/buros/bimm/html/article4.html>
12. Mehrens W.A., Lehman I.J. Measurement and evaluation in education and psychology, (3rd edition). - New York: Holt; Rinehart & Winston, 1991. — 592 p.
13. James C. Impara. Licensure testing: purposes, procedures and practices. Buros institute of mental measurements, 1995.
14. Булах І.Є., Філончук І.В. Проблема стандартизації і засоби педагогічної діагностики. Післядипломна освіта в Україні, 2004, № 2, — С. 90-93.
15. Харрисон К. Роль учителей, учебных методик, школьной учебной программы. // Совершенствование систем образования: Стратегические приоритеты в подготовке и повышении квалификации учителей, в осуществлении контроля качества их работы. Материалы конференции, Будапешт, 1999. С. 18-29.
16. Юнавичюте Л. Начальная подготовка учителей. // Совершенствование систем образования: Стратегические приоритеты в подготовке и повышении квалификации учителей, в осуществлении контроля качества их работы. Материалы конференции, Будапешт, 1999. С. 57-59.

17. Ладерьер П. Начальная подготовка учителей. // Совершенствование систем образования: Стратегические приоритеты в подготовке и повышении квалификации учителей, в осуществлении контроля качества их работы. Материалы конференции, Будапешт, 1999. С. 45-56.
18. Булах І.Є., Філончук І.В. Особливості діагностування професійної компетентності викладачів. Вісник післядипломної освіти: Збірник наукових праць. Випуск 1 / Ред. кол.: В.В.Олійник (гол. ред.) та ін.. — К.: Міленіум, 2005, С. 137-154.
19. Interstate New Teacher Assessment and Support Consortium (INTASC). (1992). Model standards for beginning teacher licensing and development: Washington, DC: Council of Chief State School Officers.
20. National Council for Accreditation of Teacher Education. (2000b). NCATE 2000 unit standards. Available: <www.ncate.org/standard/m_stds.htm>
21. National Board for Professional Teaching Standards. (1994). What teachers should know and be able to do. Detroit, MI.
22. Келли Дж. Роль образования и подготовки учителей в удовлетворении индивидуальных, общественных и экономических нужд. // Совершенствование систем образования: Стратегические приоритеты в подготовке и повышении квалификации учителей, в осуществлении контроля качества их работы. Материалы конференции, Будапешт, 1999. С. 10-17.
23. Karen J. Mitchell, David Z. Robinson, Barbara S. Plake, Kaeli T. Knowles. Testing Teacher Candidates. The role of licensure tests in improving teacher quality. // Committee on Assessment and Teacher Quality. National Research Council. National Academy Press. Washington, DC. — 2003. 460 p.
24. From National Association of State Directors of Teacher Education and Certification. 2000.
25. Quality Counts. From Education Week. 2000, №52.
26. Schmitt, K. (2000). An overview of credentialing. Paper commissioned by the Committee on Assessment and Teacher Quality, Center for Education, National Research Council. Unpublished paper; to obtain a copy contact the author at karaschmitt@hotmail.com.
27. Andrews T. Andrews L. (Eds.). The NAS-DTEC Manual on the preparation and certification of educational personnel 5th ed. // National Association of State Directors of Teacher Education and Certification. Dubuque, IA: Kendall/Hunt. 2000.
28. Removing the barriers to improving teaching. U.S. Department of Education. Washington, DC. 2001.
29. Van de Vijver, F. & Hambleton, R. Translating tests: some practical guidelines. European Psychologist, 1996, 1, 89-99.
30. Hambleton, R. Guidelines for adapting educational and psychological tests: A progress report. European Journal of Psychological Assessment, 1994, 10, 229-244.

31. Kendall, I., Jenkinson, J., De Lemos, M. & Clancy, D. (1997). Supplement to Guidelines for the use of Psychological Tests. Australian Psychological Society.
32. Bartram, D. (1995). The Development of Standards for the Use of Psychological Tests in Occupational Settings: The Competence Approach. *The Psychologist*, May, 219-223.
33. Bartram, D. (1996). Test Qualifications and Test Use in the UK: The Competence Approach. *European Journal of Psychological Assessment*, 12, 62-71.
34. Schafer, W. D. (1992). Responsibilities of Users of Standardized Tests: RUST Statement Revised. Alexandria, VA: American Association for Counseling and Development.
35. Eyde, L. D., Moreland, K. L. & Robertson, G. J. (1988). Test User Qualifications: A Data-based Approach to Promoting Good Test Use. Report for the Test User Qualifications Working Group. Washington DC: American Psychological Association.
36. Eyde, L. D., Robertson, G. J., Krug, S. E. et al (1993). Responsible Test Use: Case Studies For Assessing Human Behaviour. Washington DC: American Psychological Association.
37. Moreland, K. L., Eyde, L. D., Robertson, G. J., Primoff, E. S. & Most, R. B. (1995). Assessment of Test User Qualifications: A Research-Based Measurement Procedure. *American Psychologist*, 50, 14-23.
38. Fremer, J., Diamond, E. E. & Camara, W. J. (1989). Developing a Code of Fair Testing Practices in Education. *American Psychologist*, 44, 1062-1067.
39. Van de Vijver, F. & Hambleton, R. Translating tests: some practical guidelines. *European Psychologist*, 1996, 1, 89-99.
40. Hambleton, R. Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment*, 1994, 10, 229-244.



Центр тестування при МОЗ України



Видавництво «Майстер-клас»,
інформаційна агенція «Вища освіта»

Наукове видання

**Педагогічне оцінювання і тестування.
Правила, стандарти, відповідальність.**

Авторський колектив:
Ярослав Якович Болюбаш

Ірина Євгенівна Булах

Марина Рашидівна Мруга

Ігор Владиславович Філончук

Директор Департаменту вищої освіти
Міністерства освіти і науки України

Директор Центру тестування при МОЗ України,
професор

Заступник директора Центру тестування при
МОЗ України

Доцент кафедри медичної інформатики та
комп'ютерних технологій навчання Національного
медичного університету ім. О. О. Богомольця

Продюсер проекту
Головний редактор
Літературний редактор
Коректори

Комп'ютерне макетування
Дизайн обкладинки

О. В. Симоненко

О. В. Вікторова

Г. П. Лукова

Л. Г. Дончевська,

В. М. Нестеренко

А. В. П'янков

М. К. Линник

Формат
Друк
Гарнітура
Папір
Обл.-вид. арк
Ум. друк. арк.
Наклад
Зам. №

70x100/16

офсетний

Myriad

офсетний

14,24

22,10

1000 прим.

Видавництво

МК головний офіс

**ІА ВО (інформаційна агенція
«Вища освіта»)**

МК-Харків

**НМО (навчально-
методичний відділ)**

МАЙСТЕР-КЛАС

оф. 320, вул. Івана Кудрі, 5, м. Київ, 01042
kiev@masterclassbook.com

а/с 492, м. Харків, 61103

оф. 147, вул. Сумська, 39, м. Харків, 61022
iavo@masterclassbook.com

оф. 3-16, вул. Бакуліна, 11, м. Харків, 61166
kharkov@masterclassbook.com

оф. 3-10 А, вул. Бакуліна, 11, м. Харків, 61166
nmo@masterclassbook.com

Свідоцтво про внесення суб'єкта видавничої
справи до державного реєстру видавців,
виготівників та розповсюджувачів видавничої
продукції ДК № 2678