

Міщенко В. В.¹, Міщенко М. І.²

¹*Буковинський державний медичний університет, Україна*

²*Чернівецька ЗОШ I-III ступенів №28, , Україна*

Система оцінювання в Україні.

В Україні впроваджено дві загальнонаціональні системи тестувань, ліцензійний іспит Крок та зовнішнє незалежне оцінювання. З часу впровадження ЗНО змінилась система перетворення сирого тестового результату в рейтинговий, а питання валідності отриманих балів не зникають. Актуальним є питання співставності рейтингового балу з фактичним рівнем знань. Гострим є питання коректного вибору критеріїв оцінювання. Проаналізовано системи оцінювання в цілому та ЗНО зокрема.

Ключові слова: оцінювання, критерій, зовнішнє незалежне оцінювання, державна підсумкова атестація.

Вступ. Стан при якому б набір критеріїв якісного оцінювання у повній мірі відповідав усім завданням, контекстам та задовільняв всіх учасників процесу є практично недосяжним. Так підсумкове оцінювання у різних формах, як от державна підсумкова атестація (ДПА), зовнішнє незалежне оцінювання (ЗНО) перш за все провадяться задля контролю та звітування. Покращення успішності екзаменованого при цьому не є у пріоритеті. В Україні починаючи з 2008 року відбір абітурієнтів закладами вищої освіти (ЗВО) ґрунтується на результатах ЗНО). Природньо, що для ЗВО результати ЗНО та їх валідність є значимими. В той же час для закладів освіти, які забезпечували здобуття загальної середньої освіти інтерпретація результатів ЗНО також є хвилюючою, так як впливає на оцінку діяльності закладу загалом та педагога зокрема. Безумовно, якісне оцінювання є критичним і для абітурієнтів.

Як бачимо, якість системи оцінювання залежить не тільки від об'єктивних характеристик її функціонування, а й від врахування потреб всіх учасників процему.

Основна частина. Оцінювання повинно задовільняти низку критеріїв:

валідність – оцінювання має досягати чітко визначені цілі та мету;

відтворюваність – за незмінних умов результати мають повторюватись, що забезпечується стандартизацією та процедурою проведення;

еквівалентність – коректне перетворення сирих тестових результатів;

практичність – оцінювання є реалістичним, процедура проведення зрозумілою, результати використовуються відповідно до реалій та контексту;

освітній вплив – справляється мотивуючий вплив на тих, чиї знання оцінюються;

каталітичний вплив – аналіз отриманих результатів оцінювання сприяє прийняттю коректних управлінських рішень в напрямку прогресу навчального процесу;

прийнятність – зрозумілість оцінювання та його мети, прозорість процесу забезпечують формування довіри всіх учасників до отриманих результатів.

За метою є 2 типи оцінювань, формувальне та підсумкове.

Для формувального, або ж навчального оцінювання пріоритетним буде освітній та каталітичний вплив, а також практичність. Підсумкове оцінювання – оцінювання для контролю, перш за все повинно відповідати валідності, надійності, еквівалентності та прийнятності. Критеріям, що так би мовити гарантують справедливий відбір. Ці критерії важливі для абітурієнтів наприклад, в той же час каталітичний вплив важливий для закладів освіти.

Інтерпретація результатів є певною угодою між розробниками оцінювання та всіма зацікавленими сторонами. Інтерпретація результатів має задовільняти низку вимог, зокрема: форма представлення результатів (шкала) обрана відповідно до мети оцінювання; перетворення сирих тестових результатів, вирівнювання та переведення результатів з однієї шкали в іншу повинні бути прозорими, чіткими, стандартизованими; тим хто оцінювався

мають бути надані всі результати оцінювання: сирі тестові результати та бали за різними шкалами з необхідною інформацією для інтерпретації отриманих результатів та сфери їх використання тощо [1].

Підрахунок первинних тестових результатів можна здійснити в один з способів:

в підрахунку враховуються вірні відповіді, хибні до уваги не приймаються;

в підрахунку враховуються вірні відповіді, а за хибні нараховується штрафний бал (коефіцієнт 0,25–1);

в підрахунок введено поправку на вгадування.

Який би спосіб не було обрано, її методика має бути відома, особам, що оцінюються заздалегідь, так як це впливає на тактику проходження тесту. Наприклад, підрахунок здійснюється з нарахуванням штрафних балів за невірні відповіді, в такому випадку особі не вигідно вказувати відповіді, в яких сумнівається.

Відповідно до мети дослідження розрізняють нормо-орієнтовані вимірювання та вимірювання досягнень. Для першого основним є розрізнити учасників між собою, наприклад на тих, що можуть навчатись у вишах та не можуть. Тому доменом оцінювання буде не тільки зміст дисципліни, що оцінюється, а й відповідність певним кваліфікаційним вимогам (н-д на здатність особи здобувати освіту у виші). Тестові завдання в такому вимірюванні, як правило, використовують середньої складності [2].

Нормо-орієнтовані вимірювання виходять з теоретичних уявлень про здібності людини, розподіл людей за здібностями відповідно до нормального розподілу.

Вимірювання досягнень орієнтоване ж перш за все на оцінку рівня досягнень екзаменуємого. Домен оцінювання буде вужчим ніж у першому випадку і стосуватиметься змісту дисципліни. Результати вимірювання досягнень на відміну від нормо-орієнтованих вимірювань можуть не підлягати нормальному розподілу.

Для поділу осіб, що екзаменуються використовують 2 типи критеріїв, абсолютний чи відносний. Критерій встановлюють експертним чином, для абсолютного критерію визначають граничний бал, що відповідатиме мінімально достатньому обсягу знань, а для відносного – відсоток тих, хто не складе випробування. Критерії встановлюються та оприлюднюються заздалегідь (в Україні з 2015 року критерії ЗНО оголошуються після самого випробування). Відносний критерій використовують для відбору та рейтингового шкалювання, граничний бал (абсолютний критерій) при цьому вже значення не матиме.

З метою встановлення абсолютного критерію застосовують різні методики які ґрунтуються на уявленнях про:

- екзаменаційні завдання (методи Недельського, Ангоффа, Ебеля);
- окремих екзаменованих (методи Лівінгстона, Берка);
- групи екзаменованих.

Вибір критеріїв та коректне шкалювання є одним з найскладніших завдань у тестуванні.

Обговорення та результати. З 2008 року в Україні сертифікати ЗНО стали обов'язковими у вступних випробуваннях до вишів, у 2008 та 2009 році результати ЗНО також використовували і для ДПА, до ДПА у формі ЗНО повернулись у 2015 році і з того часу результати ЗНО з декількох предметів враховують як оцінки ДПА [3].

Як бачимо ЗНО в Україні виконує одразу два завдання, вимірювання з метою відбору та вимірювання досягнень. Таке поєднання призвело до ускладнення вибору критеріїв оцінювання, конвертації ще на початку впровадження системи. Сирі тестові результати до 2015 року перетворювали в рейтинговий бал (100-200) за допомогою еквіпроцентильного підходу. При цьому було обрано в якості абсолютного критерію 124 бали або ж 4 бали за 12-бальною шкалою (що пов'язало рейтингову шкалу з шкалою ДПА). Абсолютний критерій в 124 бали діяв з 2008 по 2014 рік. Однак при цьому робились додаткові обмеження, як от встановлення додаткового абсолютного критерію з

профільних предметів для деяких спеціальностей в 150 балів. Недоліком конвертації тестового результату в рейтинговий було те, що приріст тестових балів не призводив до пропорційної зміни рейтингового балу. Крім того, еквіпроцентильний підхід частіше використовується в практиці як метод вирівнювання. В 2015 році на зміну еквіпроцентильного методу впроваджено процентно-процентильний метод конвертації тестових балів у рейтинговий бал (за шкалою 100-200). В новій гібридній технології переведення сирого балу у шкальований поріг «склав / не склав» вже не статистичний, а розраховується щороку експертною групою на основі методик Ангоффа-Беука. Відбулось поєднання різних підходів, протилежних за своєю філософією.



Рис.1 Гістограма розподілу сирих тестових результатів ЗНО з хімії у 2016 -2020 роках [4].

З рис.1 видно, що тесту не вистачає розподільчої здатності, що може бути зумовлено надто складними завданнями, або ж надто низьким рівнем підготовки учасників, що утруднює розподіл осіб з достатнім та початковим рівнем підготовки. В той же час бачимо деяке покращення розподілу у 2019 та 2020 роках.

Висновки: Хоча проблему вольового, а не наукового рубіжного балу (124 бали) із запровадженням нового гібридного підходу й було усунуто, з'явилися

нові. Так значення критеріїв встановлюються щороку, а їх величина залишається загадкою для учасників на момент випробування.

Зміна статусу предмету з обов'язкового на вибірковий теоретично має демонструвати зсув у бік кращої успішності чи це враховується під час визначення критеріїв залишається загадкою, що може становити ризик зниження результатів ДПА після конвертації.

Поєднання в одному випробуванні двох протилежних за філософією завдань ускладнює процес оцінювання. Так розподільча функція для вступу до ЗВО потребує використання нормо-орієнтованого вимірювання, тоді як ДПА - вимірювання досягнень. Через це досягнення цілей стає компромісним.

Література:

- [1] Болюбаш Я.Я., Булах І.Є., Мруга М.Р., Філончук І.В. Педагогічне оцінювання і тестування: правила, стандарти, відповідальність: Наук. вид. – К.: Майстер-клас, 2007. – 272 с.
- [2] Булах І.Є., Мруга М.Р. / Створюємо якісний тест: Навч. посіб. —К.: Майстер-клас, 2006 — 160 с.
- [3] <https://testportal.gov.ua/dpa>
- [4] <https://zno.testportal.com.ua/opendata>