**Визначення результатів зовнішнього незалежного оцінювання: старі проблеми та нові перспективи**

|  |  |
| --- | --- |
| http://education-ua.org/images/users/avatars/900/1136/%D0%91%D0%BE%D0%B9%D0%BA%D0%BE%20%D1%84%D0%BE%D1%82%D0%BE.JPG | Автор: Валерій БойкоОпубліковано 17.10.2014 |

Зовнішнє незалежне оцінювання, увійшовши дев’ять років тому в освітній простір України, посіло в ньому чільне місце та довело суспільству свою життєздатність. Завдячуючи ЗНО, успіх вступу абітурієнта до ВНЗ став визначатися лише знаннями. Двері провідних вишів відкрилися для усіх талановитих випускників шкіл не залежно від того, наскільки далеко вони мешкають від столиці або обласних центрів. Безперечними перевагами ЗНО, на наш погляд, є: об’єктивна оцінка знань; автоматизація системи визначення результатів; створення об’єктивного механізму оцінки якості освіти; зменшення поля для корупції; транспарентність системи; підвищення мобільності абітурієнтів (отримання доступу на навчання в кращих вишах, незалежно від місця проживання, можливість подавати документи одразу до кількох університетів).

Проте, як будь-яка система, створена людиною для людини, система ЗНО не є ідеальною й потребує постійного ухвалення управлінських рішень, які, з одного боку, були б дієвими запобіжниками проти її маргіналізації, а, з іншого боку,– надавали б системі потужних імпульсів для її розвитку та удосконалення.

На сьогоднішній день система перебуває на порозі прийняття одного з таких рішень. Рішення, яке має стосуватися найбільш суспільно чутливого елементу системи – визначення результатів ЗНО. Метою написання даної статті є спроба пояснити виправданість такого рішення й механізм його реалізації.

Із перших років запровадження зовнішнього незалежного оцінювання для визначення його результатів Український центр оцінювання якості освіти застосовує шкалу 100–200 балів, яка будується автоматично на основі методу еквіпроцентильної нормалізації з середнім значенням 150 балів і середньоквадратичним відхиленням 20. Використання такої процедури шкалювання має низку суттєвих переваг. По-перше, шкала 100–200 балів є наближеною до еталонної шкали з середнім значенням результатів із кожного предмета. По-друге, шкала є стандартизованою, тобто розподіл абітурієнтів за цією шкалою є приблизно однаковим незалежно від предмета тестування. По-третє, шкала є рейтинговою, оскільки вказує на місце результату зовнішнього оцінювання абітурієнта серед результатів інших абітурієнтів із певного предмета. По-четверте, шкала є справедливою, тому що абітурієнти, які набрали однакову кількість тестових балів із певного предмета, отримують за цією шкалою однакову оцінку. Водночас абітурієнт, який має більшу кількість тестових балів, отримає за шкалою 100–200 балів оцінку не нижчу, ніж той, хто набрав меншу кількість балів.

При цьому, з метою відсіювання абітурієнтів, рівень знань і умінь яких є фрагментарним і занизьким для продовження навчання у ВНЗ, було запроваджено своєрідний «поріг» у 124 бали. Чому вибрано саме 124 бали, а не, наприклад, 125 або 130, пояснити й обґрунтувати з позиції теорії педагогічних вимірювань неможливо. Як неможливо пояснити, чому результат у 123,5 бала є ЩЕ «непрохідним», а результат у 124 бали є ВЖЕ «прохідним». А якщо врахувати ту обставину, що цей кількісний показник є «універсальним», оскільки застосовується до результатів тестування з усіх предметів, парадоксальність ситуації стає ще більш очевидною.

Крім того, іншим «вузьким» місцем «порогу» в 124 бали є те, що його автоматичне встановлення не гарантує відсікання бала «сліпого вгадування» (рис. 1).



Аби зрозуміти наскільки такий стан речей є властивим для системи зовнішнього незалежного оцінювання в цілому розглянемо результати конвертації тестових балів зовнішнього оцінювання 2014 року в оцінки за шкалою 100–200 балів. У таблиці 1 подано тестовий бал з відповідного предмета та його значення в шкалі 100–200, отримання якого надавало абітурієнту право участі в конкурсному випробуванні при вступі до ВНЗ, у таблиці 2 подано бал «сліпого вгадування» з відповідного предмета та його значення в шкалі 100–200.

*Таблиця 1*

**

*Таблиця 2*

**

Аналіз кількісних даних таблиць дає змогу виокремити три групи предметів:

– предмети, де тестовий бал, необхідний для подолання «порогу» 124 балів, суттєво перевищує бал «сліпого вгадування» (українська мова і література, російська мова, біологія, географія);

– предмети, де тестовий бал, необхідний для подолання «порогу» 124 балів, мінімально перевищує бал «сліпого вгадування» (хімія, іноземна мова) або є тотожнім (історія України);

– предмети, де тестовий бал, необхідний для подолання «порогу» 124 балів, є нижчим бала «сліпого вгадування» (математика, фізика).

Таким чином, лише в результатах з математики (рис. 2) і фізики (рис. 3) спостерігаємо картину, коли автоматично встановлений «поріг» у 124 бали, начебто відсіявши абітурієнтів, які не мають елементарних знань аби продовжувати навчання у ВНЗ, не відфільтрував тих абітурієнтів, які з високим ступенем імовірності набрали свої бали завдяки лотерейній вдачі, а не предметним знанням і вмінням.



 Отже, ситуація для системи ЗНО не є критичною, хоча і є дражливою, оскільки викликає нарікання, перш за все, з боку тих вишів, які зарахували на навчання абітурієнтів з балами вище «порогу» в 124 бали, але нижче порогу «сліпого вгадування». Насправді ж у таких новоспечених студентів за плечима немає реальних знань, а їхнє відрахування з ВНЗ є лише питанням найближчої сесії.

Яким ж є шлях вирішення проблеми співвідношення «порогу» в 124 бали та бала «сліпого вгадування»?

Найпростішим, на перший погляд, виглядає шлях збільшення кількості завдань у тестах з математики та фізики. Однак, такий крок тягне за собою збільшення видатків на підготовку та адміністрування тестів, що в нинішніх умовах є вкрай проблематичним. З іншого боку, «довантаження» й без того трудомістких тестів новими завданнями вплине на якість тестів як інструментів вимірювання, оскільки фактор втомленості абітурієнта суттєво й негативно впливатиме на результат виконання ним завдань тесту.

При цьому залишається нездоланою проблема «порогу» 124 балів з точки зору його предметного змісту. Ще раз підкреслимо, що предметно наповнити кількісний показник, отриманий статистичним шляхом, неможливо в принципі.

Таким чином, ситуація дещо нагадує цугцванг, коли будь-який крок може лише погіршувати позиції системи в цілому.

Існуючий проблемний дуалізм «порогу» (відсіяти одночасно і тих абітурієнтів, які не мають елементарних знань для подальшого навчання у виші, пояснивши який саме обсяг знань стоїть за кількісним показником у 124 бали, і тих, які з високим ступенем імовірності набрали свої бали завдяки вгадуванню), на наш погляд, може бути подоланий, якщо система перейде на нові методологічні засади його визначення.

Тобто необхідно відмовитися від ***відносного критерію*** встановлення «порогу», коли його кількісний показник залежить від того, наскільки хороші результати абітурієнт продемонстрував відносно інших учасників тестування, та запровадити ***абсолютний критерій***, коли рішення щодо «порогу» («склав»/«не склав») ухвалюється експертним шляхом.

На сьогоднішній день галузь освітніх вимірювань володіє широким спектром методик експертного визначення порогу «склав»/«не склав», проте метою даної статті не є огляд усіх існуючих методів [1]. Розглянемо лише метод, який, на наш погляд, доцільно використати в системі зовнішнього незалежного оцінювання. Мова йде про метод, запропонований і описаний В. Ангоффом[2].

Чому саме метод Ангоффа? Серед чинників, що свідчать на його користь, визначальними є такі:

– пройшов ґрунтовну емпіричну перевірку й довів свою ефективність;

– широко представлений  в науковій літературі й високо оцінений фахівцями з педагогічних вимірювань;

– не потребує залучення значних матеріальних і людських ресурсів;

– потребує мінімального навчання та інструктажу залучених експертів;

– простота, легкість та прозорість процедури реалізації (адміністрування);

– висока надійність результатів його застосування.

Сутність метода Ангоффа полягає в наступному. Після завершення тестування з кожного предмета збирається група експертів (спеціальна експертна комісія), яка, аналізуючи завдання тесту, визначає (оцінює) вірогідність правильної відповіді на кожне завдання мінімально підготовленого абітурієнта. Експертні оцінки усереднюються, а ці середні значення використовуються для визначення порогу «склав/не склав» або так званого «порогового бала» (рис. 4).

 *Рисунок 4*

**

Таким чином, **«пороговий бал»** – мінімальна кількість (сума) тестових балів зовнішнього незалежного оцінювання з певного предмета, отримання яких надає абітурієнту право участі в конкурсному випробуванні при вступі на навчання до вищого навчального закладу України.

                     Як вже зазначалося, процедура адміністрування методу є доволі простою й прозорою та складається з кількох етапів.

**1 етап – організаційний.**

На цьому етапі відбувається формування кола експертів–предметників. Їх має бути не менше 10. При цьому, експерти мають бути різної статі, віку, сфери діяльності (вчителі загальноосвітніх навчальних закладів, гімназій і ліцеїв; викладачі вищих навчальних закладів, науковці Національної академії наук, Академії педагогічних наук). У такий спосіб створюється прийнятний «освітній фон», коли експертна комісія в цілому репрезентуватиме освітнє середовище України. Варто підкреслити, що чим більше буде залучено експертів, тим кращою буде дисперсія експертних оцінок. Крім того, різність думок (оцінок) зменшуватиме ефект «яструба/голубки». Як наслідок – більш точним виявиться середнє значення критеріального «порогового бала».

**2 етап – підготовчо-методологічний.**

На цьому етапі відбувається загальне знайомство експертів з тестом, його структурою, формами завдань тощо. З метою забезпечення спільності розуміння підходу до оцінювання завдань експерти, ознайомившись із тестом, у своєму колі мають обговорити та визначитися з поняттям «мінімально підготовлений абітурієнт». У якості ілюстрації наведемо можливі визначення загального характеру.

«Мінімально підготовлений абітурієнт»:

– той, хто володіє найменшим обсягом знань, необхідних для виконання завдання;

– той, хто своїми знаннями й уміннями ледь відповідає освітньому предметному стандарту, чиї знання й уміння є фрагментарними, граничними, але прийнятними.

Зрозуміло, що експерти під час обговорення мають наповнити поняття предметним змістом. Наприклад, «мінімально підготовлений абітурієнт з іноземної мови – це той, хто читає короткі прості тексти з метою отримання основної інформації; знаходить необхідну інформацію у простих текстах повсякденного характеру (реклама, оголошення, меню, розклад руху поїздів, електронні листи тощо); вживає прості граматичні структури із завченими фразами, групами з кількох слів та стійких виразів для опису простих щоденних ситуацій; пише короткі прості записки та звернення, пов’язані з нагальними потребами».

        Після завершення обговорення модератор роботи комісії роздає Листи оцінювання завдань тесту (рис. 5) та здійснює загальний інструктаж експертів щодо оцінювання завдань: *«Визначте (оцініть) кількість мінімально підготовлених абітурієнтів зі 100, які правильно виконають завдання. Наприклад, якщо на Вашу думку це завдання виконає 60 зі 100, то в графі «% правильної відповіді» Листа оцінювання запишіть значення 60%»*.

Рисунок 5



          Під час інструктажу модератор групи обов’язково наголошує на таких моментах:

1) оцінка не повинна бути вищою, ніж 95 та нижчою, ніж 25 (для завдань з чотирма варіантами відповіді), оскільки навіть дуже добре підготовлений абітурієнт може помилитися й не виконати завдання, а мінімально підготовлений абітурієнт може вгадати правильну відповідь;

2) під час роботи експертам не надаються «ключі» відповідей до завдань, оскільки це може необґрунтовано впливати на експерта, нівелюючи його розуміння, «відчуття» складності завдання.

**3 етап – експертна оцінка тестових завдань.**

На цьому етапі експерти самостійно працюють із завданнями тесту,заповнюючи Лист оцінювання. Варто підкреслити, що обговорення експертами завдань та можливих ймовірностей правильних відповідей не допускається.

**4 етап – узагальнення експертних оцінок.**

Модератор групи збирає Листи оцінювання та вводить значення в таблицю Excel для кожного завдання по кожному експерту й обчислює стандартне відхилення (рис. 6).

 *Рисунок 6*

**

Стандартне відхилення є «дзеркалом», яке відображає рівень розбіжностей між експертами по кожному завданню. Низький його показник свідчить про високий рівень згоди між експертами. Високий показник є підґрунтям для колективного обговорення експертами причин розбіжностей в оцінках, яке організовує та координує модератор. Мета обговорення – підвищити рівень згоди між експертами шляхом перегляду ними своїх оцінок.

Варто підкреслити, що будь-яке завдання, для якого стандартне відхилення експертних оцінок перевищує 10, колективно обговорюється. За результатами таких обговорень експерти можуть змінити свою оцінку. У такому разі модератор групи вносить відповідні зміни в узагальнену таблицю Excel.

Слід зауважити, що на цьому етапі роботи комісії експерти мають бути ознайомлені зі статистичними характеристиками результатів тестування (*мода, медіана, середнє значення, стандартне відхилення, асиметрія, ексцес*) та психометричними характеристиками завдань тесту (*складність, розподільна здатність, кореляція*). Урахування цих даних під час колективного обговорення експертних оцінок дозволить підвищити інформативність і обґрунтованість встановленого «порогового бала».

**5 етап – визначення «порогового бала».**

Після обговорення, коригування експертами своїх оцінок модератор в таблиці Excel здійснює перетворення відсотка ймовірності правильної відповіді по кожному завданню у відповідний коефіцієнт. Наприклад, 45% – 0,45, 30% – 0,30 і т. д.

Далі модератор обчислює суму всіх коефіцієнтів завдань тесту по кожному експерту. Ця сума є «пороговим балом», який встановив певний експерт.

На завершення обчислюється середнє значення «порогових балів» усіх експертів (рис. 7). Це значення і є підсумковим «пороговим балом» для тесту, заради визначення якого й працювала експертна комісія.

*Рисунок 7*

**

Після проведення усіх процедур обчислення результат роботи експертної комісії документально оформлюється Протоколом засідання комісії та Протокольним рішенням (ухвалою) щодо встановленого «порогового бала» з відповідного предмета.

Далі відбувається стандартна процедура переведення тестових балів у оцінки за шкалою 100–200 балів. При цьому точкою відліку на шкалі буде не «**0»** тестових балів, а значення «порогового балу». Таким чином, результати абітурієнтів, які подолали «поріг», займуть відповідне місце на рейтинговій шкалі в діапазоні від 100 до 200 балів. У результаті навіть абітурієнти, які матимуть рейтинговий результат 100,5 балів, не кажучи вже про тих, чий результат становитиме 123,5 бала, отримають можливість брати участь у конкурсному випробуванні під час вступу на навчання до вищого навчального закладу України.

Таким чином, перехід до використання в практиці визначення результатів зовнішнього незалежного оцінювання критеріального «порогового бала» («склав»/ «не склав») дозволить подолати дуалізм проблеми «порогу» в 124 бали та розширити можливості тих абітурієнтів, які мають відповідний багаж знань і умінь, аби стати студентами омріяних університетів.

***Валерій Бойко****, начальник відділу змісту суспільно-гуманітарних навчальних предметів Українського центру оцінювання якості освіти*