

HISTORY AND THEORY OF EDUCATION

Віталія Готинян-Журавльова, к. філос. н.

Одеський національний університет імені І.І. Мечникова, Україна

ОСОБЛИВОСТІ ВИМІРЮВАННЯ ЯКОСТІ ЗНАНЬ: ФІЛОСОФСЬКИЙ АСПЕКТ

Vitaliya Gotynyan-Zhuravlyova, PhD in Philosophy

Odesa I. I. Mechnikov National University, Ukraine

PECULIARITIES OF KNOWLEDGE QUALITY MEASUREMENT: THE PHILOSOPHICAL ASPECT

The methodological aspects of quality measurement of students' knowledge, skills, and abilities are analyzed in the article. It is difficult to follow the methodological requirements of reference measurement within such measurements. The results of such measurements aren't objective. The testing is one of the most famous methods of quality measurement. It is assumed that the teacher can provide an objective assessment of students' knowledge and skills using tests. However, tests do not meet the requirements for the reference measurement, and therefore automatically become a non-reliable method of the measurement. The possibility of using the standardless measurement is mentioned in the article. The standard (as an international standard of measured value) is not used at standardless measurement. Then the question of the objectivity of the results is quite different. The methodological requirements of the standardless measurement allow considering testing as a scientific method of measurement. It should be emphasized that two types of testing according to their mechanisms correspond to two types of standardless measurement.

Keywords: measurement, standardless measurement, testing, objectivity, methodological aspects.

Будь-яка людська діяльність потребує контролю якості виконання і перевірки результатів. Сутність контролю зазвичай полягає в співвідношенні мети з результатами, які були отримали. Найчастіше контроль проводиться за допомогою певних нормативів, певної системи вимірювань, одиниць вимірювань. Вимірювання, яке є невід'ємною частиною методології багатьох наук, зокрема використовується в педагогіці. Саме завдяки вимірюванню можна визначити рівень досягнень, вмій учнів відповідно до мети навчання і відповідно до освітнього стандарту. Той факт, що в педагогіці використовується поняття «освітнього стандарту», вказує на те, що методологи розробляють засоби вимірювання, які схожі на еталонне вимірювання, а «освітній стандарт» виконує функцію еталону.

По суті освітнім стандартом є особливі нормативні документи, які встановлюють обов'язкові мінімальні вимоги до навчання і виховання. А далі на основі стандарту розробляються освітні програми, посібники, підручники, які деталізують зміст стандартів. Але чи можливо створити стандарти, які були б однаково вдалим як для математиків, технарів, так і для гуманітаріїв? Чи можливо створити стандарти, які б запропонували програму, що була б однаково цікавою для відмінників, і для тих, хто вчиться «на одні трійки»? Крім того, однією з найскладніших проблем стандартизації освіти є співвідношення в педагогічній діяльності стандарту з елементами творчості, які неодмінно повинні бути присутні в освітньому процесі. Чи можливо в загалі «копіювання» методології еталонного вимірювання в педагогіці, чи може слід звернути увагу методологів, методистів, викладачів та педагогів на безеталонне вимірювання?

Еволюція поняття «педагогічне вимірювання» представлена в роботах відомих вчених, методологів, викладачів – А. М. Арсеньева, В. С. Авансова, Н. М. Розенберга, Л. Б. Смирновой, В. А. Хлебниковой, М. Б. Челишковой, В. Д. Шадрикова та інших.

На нашу думку, методологам науки, методистам, які плідно працюють задля того, щоб покращити якість, точність і об'єктивність вимірювання рівня знань учнів, слід звернути увагу на безеталонне вимірювання. Про особливості та методологію безеталонного вимірювання йшлося в роботах ак. А. І. Уймова, Г. А. Полікарпова, О. П. Пунченка, В. В. Готинян-Журавльової.

Тому **метою даної статті** є аналіз можливості застосування механізмів безеталонного вимірювання при оцінюванні якості знань учнів та студентів, а також аналіз схожості механізмів певних методів оцінювання з механізмами безеталонного вимірювання.

Найчастіше педагогічне вимірювання визначають як складну частину освітньої діяльності, сутністю якої є визначення рівня відповідності результатів учбово-виховної роботи з встановленими освітніми стандартами за допомогою тієї чи іншої системи вимірювання.

Для чіткого зазначення процедури вимірювання необхідні інструменти, які зможуть точно і об'єктивно вимірити рівень знань учнів. А для будь-якої об'єктивної оцінки потрібні надійні методи для вимірювання результатів навчання. До методів діагностики рівня і якості знань можна віднести: контрольні роботи, диктанти, реферати, твори, які необхідні для виявлення рівня знань, навиків і вмінь вирішувати типові завдання. Існують й інноваційні методи та прийоми, які орієнтовані на встановлення і оцінку рівня компетентності, та підготовки до вирішення складних творчих нестандартних завдань, таких як написання есе, участь у диспутах, турнірах, конкурсах тощо.

Одним з найефективніших, найвикористовуваних і популярних методів вимірювання рівня знань учнів та студентів є тестування. Тести використовуються для оцінки здібностей та рівня знань. Так, на думку М. Б. Челишкової, тест є сукупністю завдань, які дозволяють надати об'єктивну і навіть кількісну оцінку якості підготовки учня в певній освітній галузі¹. В. А. Хлебнікова вважає педагогічний тест як «систему спеціально підібраних перевірочних завдань специфічної форми, що дозволяють якісно оцінити учбові досягнення в одній або декількох галузях знання»². Традиційний педагогічний тест – це стандартизований метод діагностики рівня і структури підготовки учнів. Він є системою знань наростаючої складності, специфічної форми, яка дозволяє якісно і ефективно вимірити рівень і оцінити ступінь підготовки учнів та студентів. Система знань в цьому визначенні означає, що в тесті зібрані такі знання, які мають системоутворюючі властивості. Чітко виокремлюється загальна приналежність завдань до однієї і тієї ж системи знань, тобто до однієї учбової дисципліни, вказується їх взаємозв'язок і впорядкованість. Щодо специфічної форми, то в тесті завдання повинні бути сформульовані у формі висловів, які можуть бути істинними або хибними. Як правило, традиційні філософські питання істинними або хибними не можуть бути, а відповіді на них інколи бувають настільки невизначеними і багатослівними, оскільки потребують роздумів та аналізу, що для виявлення їх правильності потрібні значні витрати інтелектуальної енергії та часу викладачів.

Слід підкреслити, що критерії правильності завчасно визначаються авторами тестів, а це не може гарантувати повної об'єктивності тестів, особливо це стосується застосування тестів, як форми контролю, в суспільно-гуманітарних науках. Кожен автор буде акцентувати увагу на найбільш важливих на його думку сторонах і аспектах вивченого матеріалу, ступень засвоєння яких треба визначити в тесті. Крім того, деякі автори складають тести дуже ретельно, прописуючи питання, які стосуються усіх нюансів і подробиць вивченого матеріалу, інші – акцентують увагу на найбільш загальних положеннях. Загальна кількість балів в обох випадках може бути однаковою, а затрати часу на підготовку, необхідний рівень знань для складання тесту, зосередженість і увага при написанні можуть бути різними. Тому повної об'єктивності, на нашу думку, бути не може.

Структура знання оцінюється на основі послідовності правильних чи неправильних відповідей на завдання при їх зростаючій складності. Часто виникає питання: а скільки завдань повинно міститися у тесті (тобто яка повинна бути його довжини) для того, щоб, з одного боку, можна було охопити увесь матеріал, а, з другого боку, ті, хто тестуються, зберегли увагу і концентрацію до кінця випробування? Довжина тесту – ще одна характеристика тесту, яка саме визначає кількість завдань. Чим більшим за довжиною є тест, тим більше в ньому завдань. Від кількості завдань певним чином залежить точність педагогічного вимірювання. Але якщо при тестуванні ставити акцент не на точність, а на людину, на того, хто тестується, то всі ми знаємо, що чим довчим є тест, тим

¹ Хлебнікова, В. (2003). Основные принципы построения системы понятий и терминов педагогического тестирования. *Стандарты и мониторинг в образовании*, 2, 53-62.

² Там само, 60.

більше вірогідність того, що учень (або студент) не зможе бути зосередженим весь час написання тесту. На певному етапі проходження випробування учень може втратити концентрацію уваги і мріяти про якнайшвидше завершити проходження тесту. А тому більшість тестів «на папері» виглядають чіткими та об'єктивними, для найточнішого вимірювання якості знання учнів, а в реальному житті дуже їх важко скласти.

Крім того, тести можна розділити за їх орієнтацією на критеріально-орієнтовані і нормативно-орієнтовані. Так при критеріально-орієнтованому тесті всі, хто тестуються, відповідають на одні і ті ж завдання, за однаковий час, в однакових умовах, і, головне, оцінюються за однаковими правилами і критеріями. Вони мають назву критеріально-орієнтованих, бо створюються за схемою: завдання – відповіді – висновки про відповідність того, хто складає тест, заданому критерію. При цьому відбувається зіставлення учбових досягнень кожного учня з обсягом завдань, навичок, які він повинен був засвоїти за певний період часу за вимогами навчального стандарту.

Серед явних плюсів використання таких тестів є отримання повної інформації про досягнення учня, можливість порівняння вмінь і навичок учнів з вимогами програми, які чітко прописані за даною дисципліною. Складності, з якими зіштовхуються і вчителі, і учні, полягають в тому, як відобразити у достатньо невеликій кількості запитань зміст якомога більшого обсягу учбового матеріалу. Крім того, більшість шкіл немає професійно-орієнтованих класів, в яких мали б змогу вчитися діти, однаково зацікавлені в ретельному і різнобічному вивченні тих чи інших дисциплін, що значно спростило б складання тестів. Зазвичай викладачу або методисту, які складають тести, слід враховувати різноманіття уподобань учнів, а саме, що в одному класі вчать як «технарі», так й гумантарії, а це суттєво ускладнює складання тесту, об'єктивного для оцінки рівня знань, вмінь і навичок усіх учнів. Деяко простіше складати тести для студентів вищих навчальних закладів, оскільки спеціалізація визначає професійну орієнтацію і уподобання студентів.

При нормативно-орієнтованих тестах відбувається порівняння тих, хто тестується, за рівнем їх учбових досягнень. В результаті зіставляються бали кожного учня (або студента) з результатами інших учнів (або студентів), які виконували той самий тест при тих самих вимогах і при тих самих обставинах. При цьому встановлюється відношення порядку між тими, хто проходить тест, за рівнем виявлених при тестуванні знань. На основі отриманої інформації можна скласти рейтинг учнів (студентів), які проходили тест, за отриманими ними балами. Створюються нормативно-орієнтовані тести наступним чином: завдання – відповіді – висновки про знання тих, хто проходив тест – рейтинг, який вказує місце кожного учасника тесту. Це тести, які складаються для встановлення рейтингу між тими, хто тестується, для того, щоб з'ясувати в кого найкращі знання. Як правило, такі тести застосовуються в професійно-орієнтованих класах, де вчать учні, з однаковими вподобаннями щодо деяких дисциплін та вищих навчальних закладах для встановлення рейтингу студентів за певними дисциплінами.

Самим складним при нормативно-орієнтованих тестах залишається питання про істинність оцінки результатів кожного, хто тестується. Задля цього емпірично встановлюють норми – систему показників, що відображають результати виконання тесту великою групою тих, хто тестується.

Цікавим і таким, що потребує уваги залишається питання щодо надійності тесту. Найчастіше тест визначається надійним якщо відбувається збіг показників, що отримали ти, хто тестується, при повторному тестуванні тим самим тестом, або інколи еквівалентною його формою. Тест вважається надійним, якщо він дає одній й ті ж самі або дуже близькі показники для кожного, хто тестується, при повторному тестуванні. Але для цього необхідно щоб ті, хто тестується, не змінили свій рівень підготовки перед повторним тестом, що насправді є дуже складною вимогою. Крім того, мотивація щодо отримання найкращих результатів тих, хто тестується, не повинна змінитися. А цю умову дуже складно виконати, оскільки у більшості тих, хто тестується, може з'явитися бажання підвищити свої результати при наступному випробуванні тим самим тестом або дуже схожою його формою. Якщо такої мотивації не буде, то загальний результат тих, хто тестувався, може погіршитися, бо навіщо ж тоді тестуватися. При визначенні надійності тесту знаходять коефіцієнт між двома тестами (основним та повторним) на одній і тій самій виборці студентів.

На нашу думку, таке визначення надійності тести певною мірою дублює вимогу повторюваності будь-якого дослідження в природознавчих науках. У випадку вимірювання в педагогіці, як і в багатьох інших суспільно-гуманітарних науках, ми не можемо стовідсотково покладатися на повторюваність результатів, оскільки саме в суспільно-гуманітарних науках акцент робиться на людину, на того, хто тестується, з урахуванням його стану здоров'я, настрою, психологічного стану, мотивації. Усе це може суттєво відрізнятись при повторному написанні того

ж самого тесту.

Крім того, ступінь науковості педагогічного контролю пов'язана ще з одним поняттям – поняттям «валідності», тобто адекватності, придатності методу для досягнення визначеної мети. Педагогічний тест вважається валідним, якщо його тестові завдання знаходяться у відповідності до вивченого матеріалу і повністю відображають його зміст, та не містить у собі питання дисциплін, що вивчалися раніше, а також споріднених дисциплін. В тесті не повинні бути питання, відповіді на які можна «вгадати», або питання, які погано підібрані, нечітко сформульовані або ж такі, на які може бути декілька вірних відповідей. Цих вимог треба чітко дотримуватися, хоч зробити це важко.

На нашу думку, методологам науки, методистам, які плідно працюють задля того, щоб покращити якість, точність і об'єктивність вимірювання якості і рівня знань учнів, слід звернути увагу на безеталонне вимірювання¹.

Безеталонне вимірювання об'єднує доволі широкий клас різних за механізмом вимірювань, які відбуваються без використання еталону, як строго фіксованого міжнародного стандарту вимірюваної величини, також без обов'язкового пошуку математичного відношення «більше» ніж еталон. При безеталонному вимірюванні поняття відношення має не виключно математичний, а скоріше філософський характер, вказуючи на приписуванні тієї чи іншої властивості вимірюваному об'єкту, про що йшлося в роботах С. С. Стівенса, Н. Р. Кемпбелла та інших. Саме при безеталонному вимірюванні акцент робиться на аналізі вимірюваної якості об'єкту, зазвичай, без перетворення цієї якості в кількість, у число, якщо цього не вимагає дослідження. Зазвичай безеталонне вимірювання використовується у тих випадках, коли неможливо застосувати еталонне вимірювання, оскільки немає чітко фіксованих міжнародних стандартів – еталонів. Як еталон, точніше квазіеталон для порівняння з вимірюваною величиною, може бути використаний будь-який об'єкт або будь-яка властивість, яка, на думку дослідника, може замінити еталон в цьому дослідженні. Питання «об'єктивності отриманих результатів» перестає бути найважливішим.

В деяких випадках вимірювання в суспільно-гуманітарних науках неможливо чітко дотримуватися механізму і принципів еталонного вимірювання. Тому методологи науки пропонують альтернативні методи вимірювання, як то шкалування, вимірювання за допомогою індексів, класифікування, тощо. Більшість методологів вважають ці альтернативні види вимірювання «другорядними», щодо еталонного вимірювання. Але науковці, які працюють в галузі суспільно-гуманітарних наук, навряд чи зможуть відмовитися від використання цих методів лише тому, що вимірюванні засоби є грубими та суб'єктивними і немає бездоганих вимірюваних шкал. В деяких випадках їх просто неможливо створити. На думку Пунченка О. П., «безеталонне вимірювання може розглядатися як загальнозначений людством інтервал, в рамках якого позитивно функціонує зміст даного морального поняття або категорії. Моральні категорії, і структурні, такі як норма, принцип, оцінка, ідеал, погляди, знання, переконання, почуття, звички, звичаї, дії, і сутність, такі як – добро і лихо, совість і сором, доброзичливість, обов'язок, щастя та інші, – не можуть мати єдиного кількісного еталону, оскільки вони функціонують як міра»² Тому, на нашу думку, слід об'єднати альтернативні способи вимірювання в широкий клас безеталонних вимірювань. При використанні безеталонного вимірювання стане можливим, обійти частину труднощів, в нашому випадку труднощів, пов'язаних з написанням тестів для класів без професійної орієнтації.

Слід зазначити, що існує чотири види безеталонного вимірювання, які розрізняються за механізмом вимірювання³. Так існує вид безеталонного вимірювання, який засновано на порівнянні вимірюваної речі з іншою річчю, яка взята у якості еталона вимірюваної величини. Річ, яка розглядається як еталон, може бути еталоном лише при одному вимірюванні і не використовуватися як міжнародний стандарт. На цьому принципі базується, наприклад, вимірювання за допомогою шкал. Яскравим прикладом якого є шкала Мооса.

Інший вид безеталонного вимірювання заснований на порівнянні властивості, яка розглядається як еталон (у більшості випадках розглядається взагалі як деяка самостійна сутність і

¹ Уёмов, А., Поликарпов, Г. (1972). К проблеме безэталонного измерения в микрофизике. *Проблемы диалектико-материалистического истолкования квантовой теории. материалы IV Симпозиума по гносеологическим проблемам измерений (г. Киев, 15-16 октября 1972г.)*, 127-140.

² Пунченко, О. (2008). Безеталонне вимірювання *Актуальні проблеми духовності Кривий Ріг*, 361-365.

³ Готинян, В. (2002). Про одну логічну відмінність еталонного та безеталонного вимірювань. *Перспективи. Науковий журнал*, 2 (18), 71-78.

так само може використовуватися як еталон лише один раз), з властивістю, що належить вимірюваній речі. Так відбувається вимірювання сили вітру за допомогою шкали англійського адмірала Бофорта. Адмірал Бофорт ретельно описав дію того чи іншого вітру на суші і на воді. За потребою відбувається порівняння тих наслідків, які викликає вітер, з описами адмірала.

Третій вид безеталонного вимірювання заснований на приписуванні властивості вимірюваній речі. По суті, в результаті вимірювання ми відповідаємо на питання: чи належить властивість вимірюваній речі, чи ні? Так відбувається приписування вимірюваній речі тієї чи іншої властивості про що йдеться у визначення поняття «вимірювання», які були надані Н. Р. Кемпбеллом та С. С. Стівенсом.

Четвертий вид безеталонного вимірювання заснований на комбінації ознак в результаті яких утворюється і вимірюється річ. Так відбувається визначення і вимірювання груп крові за комбінацією параметрів – ізоантигенів та ізоантитіл.

Пригадаємо, що за своєю орієнтацією тести можна розділити на критеріально-орієнтовані і нормативно-орієнтовані. Так при критеріально-орієнтованому тесті всі, хто тестуються, відповідають на одні і ті ж завдання, за однаковий час, в однакових умовах, і, головне, оцінюються за однаковими правилами і критеріями. При цьому відбувається зіставлення учбових досягнень кожного учня з обсягом завдань, навичок, які він повинен був засвоїти за певний період часу за вимогами навчального стандарту. При використанні таких тестів ми отримаємо повну інформацію про досягнення учня, можливість порівняння вмінь і навичок учнів з вимогами програми, які чітко прописані за даною дисципліною. Такий механізм тестування відповідає механізму безеталонного вимірювання, який засновано на порівнянні властивості, яка взята як еталон (в нашому випадку це навчальний стандарт, тобто ці знання, вміння, навички, які повинен засвоїти той, хто тестується), з властивістю, що належить вимірюваній речі (тобто зі знаннями, вміннями та навичками, які засвоїв учень або студент).

При нормативно-орієнтованих тестах відбувається порівняння учнів за рівнем їх учбових досягнень. В результаті зіставляються бали кожного учня з результатами інших учнів, які виконували той самий тест при тих самих вимогах і при тих самих обставинах. При цьому встановлюється відношення порядку між тими, хто проходить тест, за рівнем виявлених ними при тестуванні знань. На основі отриманої інформації можна скласти рейтинг учнів (студентів), які проходили тест, за отриманими ними балами. Це тести, які складаються для встановлення рейтингу між учня, для того, щоб з'ясувати в кого найкращі знання. Такий механізм тестування відповідає механізму безеталонного вимірювання, який засновано на порівнянні двох речей, одна з яких виступає в якості умовного еталону для порівняння. В нашому випадку відбувається порівняння результатів, які були отримані учнями (студентами) задля встановлення рейтингу між ними.

На нашу думку, не є можливим знайти єдиний освітній «еталонний» об'єктивний стандарт, створити єдині універсальні тести для оцінювання знань, вмінь, навичок, які здобуваються учнями, студентами і, зокрема, студентами-філософами. Занадто великий елемент творчості, неординарності відповідей, яскравих думок містять роботи студентів філософських факультетів і відділень. Студент, який вивчає філософію, повинен бути мислячою, творчою особистістю, яка зазвичай критично ставиться до будь-яких рамок, стандартних відповідей, чітких безапеляційних рішень. А тому при оцінюванні якості знань слід звернутися не до еталонів-стандартів, а до безеталонного вимірювання, що і складає перспективу подальших досліджень для методологів-філософів.

References:

1. Hlebnikova, V. (2003) Osnovnye principy postroeniya sistemy ponjatij i terminov pedagogicheskogo testirovaniya [The basic principles of constructing a system of concepts and terms of pedagogical testing]. *Standarty i monitoring v obrazovanii* [The standards and monitoring at the education], 2, 53-62. [in Russian].
2. Ujomov, A., Polikarpov, G. (1972). K probleme bezjetalonnogo izmerenija v mikrofizike [To the problem of standardless measurement in microphysics]. *Problemy dialektiko-materialisticheskogo istolkovanija kvantovoj teorii. materialy IV Simpoziuma po gnoseologicheskim problemam izmerenij (g. Kiev, 15-16 oktjabrja 1972g.)* [Problems of the dialectical materialist interpretation of quantum theory. Materials of 4 Symposium of the gnoseological measurement problems], 127-140. [in Russian].
3. Puchenko, O. (2008). Bezetalonne vymiryuvannya [The standardless measurement]. *Aktualni problemy duhovnosti Kryvyj Rig* [The actual problems of spirituality. Kryvyi Rih], 2, 361-365. [in Ukrainian].
4. Gotynyan, V. (2002). Pro odnu logichnu vidminnist etalonnogo ta bezetalonnogo vymiryuvan [About one logical difference between standard and standardless measurements]. *Perspektyvy. Naukovyj zhurnal* [Prospects. The Scientific Journal], 2 (18), 71-78. [in Ukrainian].