

## **СПІВВІДНОШЕННЯ ПОНЯТЬ ВАЛІДНІСТЬ ТА НАДІЙНІСТЬ**

**Abstract.** In the Ukrainian Sociology there are two parallel "traditions" of sociological measurement as a whole and evaluation of the validity and reliability in particular – "Soviet" tradition and "Western". This article shows the ratio of the terms of "Soviet" and "Western" traditions and the only and most expedient, from the author's point of view, the terminological apparatus was suggested. Also the ratio of content of the concepts of validity and reliability is considered.

Валідність та надійність розглядаються як ключові характеристики вимірювання. Але в сучасному соціологічному дискурсі України паралельно існують дві традиції оцінки якості соціологічного вимірювання – «радянська» та «західна». Така ситуація ускладнює процес оцінки якості конкретного вимірювання, комунікацію між дослідниками всередині країни та дослідниками з інших країн, процес викладання. Метою даної статті є показати співвідношення термінів «радянської» та «західної» традицій та запропоновано єдиний, найбільш доцільний з точки зору автора, термінологічний апарат, а також окреслити співвідношення змісту понять валідність та надійність.

### **Співвідношення термінів «радянської» та «західної» традиції оцінки якості соціологічного вимірювання**

«Радянська» традиція оцінки якості соціологічного вимірювання ґрунтується на базових принципах та поняттях метрології. До «радянської» традиції оцінки ми відносимо роботи таких авторів: А.Г. Здравомислов [5], В.А. Ядов [11], В.І. Волович [1], Б.З. Докторов [4], Г.І. Саганенко [9], В.І. Паніотто [8] та ін.

Під «західною» традицією оцінки якості соціологічного вимірювання ми розуміємо роботи американських та західноєвропейських авторів починаючи від 30-х рр. ХХ ст. і до сьогодні. Всі ці роботи об'єднати в одну загальну традицію нам дозволяє їх спільність у принципових питаннях та спадкоємність цих робіт. «Західна» традиція бере свій початок з психологічних тестів, робіт з оцінки якості психологічних тестів.

У «радянській» традиції інтегральною характеристикою якості соціологічного вимірювання вважалась надійність. Такої точки зору дотримувались В.А. Ядов, Г.І. Саганенко, В.І. Волович, Б.З. Докторов, В.І. Паніотто. У 1979 р. Б.З. Докторов та у

1986 р. В.І. Паніотто провели аналіз робіт своїх колег щодо якості соціологічного вимірювання. На основі цих робіт можна виділити основні показники надійності в радянській традиції: обґрунтованість, правильність, стійкість.

**Обґрунтованість (валідність)** – це відсутність теоретичних, методологічних, логічних помилок, вірність вихідних посилань при розробці методики дослідження, відповідність числової моделі соціальної системі, що вивчається [8, с. 27]. Б. З. Докторов називав це також відсутністю квазіпомилки узгодженості – відповідність моделі досліджуваного об'єкта самому об'єкту [4, с. 28].

**Правильність** – це міра відсутності систематичних помилок вимірювання [8, с. 31].

**Стійкість** – це міра відсутності випадкових помилок вимірювання [8, с. 35].

У «західній» традиції немає інтегральної характеристики якості соціологічного вимірювання, тут виділяють два показники якості соціологічного вимірювання – валідність (validity) та надійність (reliability). Така позиція відображена у всіх Стандартах психологічного та освітнього тестування [12; 13; 14; 15.], у класичній роботі «Оцінка надійності та валідності» Дж. Кармінеса та А. Зеллера [19], у виданні Сейдж «Довідник Сейдж з методів у соціальних дослідженнях» (Sage Handbook of social research methods) [22]. У «західній» традиції валідність та надійність розуміють таким чином:

**Валідність** – це відповідність вимірювання його меті. [19, с. 12]. Валідність – це міра того, наскільки ми вимірюємо те, що заплановано виміряти [17, с. 3207]. Валідність пов'язують з відсутністю систематичних помилок – вимірювання, яке відносно вільне від систематичних помилок, вважається валідним [18, с. 14];

**Надійність** – це відтворюваність вимірювання. Надійність стосується того, в якій мірі інструмент вимірювання дає одні і ті ж самі результати при повторному вимірюванні. Надійність пов'язана з тенденцією повторюваності при вимірюванні одних й тих самих соціальних феноменів, за допомогою одного і того самого вимірювального інструменту, декілька разів. Чим стійкіші результати вимірювання при повторних вимірюваннях, тим надійніше вимірювальна процедура, і, навпаки, чим менше схожі результати вимірювання при повторному вимірюванні, тим менш надійною є процедура вимірювання [19, с. 11–12]. Надійність пов'язують з відсутністю випадкових помилок – вимірювання, яке відносно вільне від випадкових помилок, вважається надійним [18, с. 14].

З наведеного вище аналізу понять «радянської» та «західної» традицій видно, що за змістом, показники якості вимірювання у цих двох традиціях не суперечать один одному. Основні відмінності полягають у термінології. Відмінність перша – у «радянській» традиції є інтегральна характеристика якості соціологічної інформації – надійність, у «західній» традиції такої інтегральної характеристики немає. Але це не змінює сутність розуміння якості вимірювання. У «радянській» традиції показниками надійності як інтегральної характеристики якості вимірювання виступають обґрунтованість, правильність і точність. У «західній» традиції показниками якості вимірювання виступають валідність та надійність.

Виходячи із змісту вищенаведених понять, спробуємо співвіднести ці поняття (табл. 1).

Таблиця 1

**Порівняння термінології «радянської» та «західної» традицій оцінки якості вимірювання**

<b>«РАДЯНСЬКА ТРАДИЦІЯ»</b>	<b>«ЗАХІДНА ТРАДИЦІЯ»</b>
Надійність – інтегральна характеристика якості вимірювання (інформації)	Якість інформації (quality of measurement)
Правильність	Валідність (validity)
Стійкість	Надійність (Reliability)

У «західній» традиції не розмежовуються показники відсутності теоретичних та систематичних помилок. Вони об'єднані в одну категорію – валідність. У радянській традиції ці показники розмежовуються – обґрунтованість та правильність. В той же час як Б.З. Докторов, так і В.І. Паніотто відмічають, що чітко розмежувати обґрунтованість та правильність не завжди вдається, оскільки у деяких випадках, теоретичні помилки збігаються із систематичними, в інших випадках ця розбіжність виражена досить яскраво [4, с. 28; 8, с. 31;]. Як зазначив В. І. Паніотто, «розмежувати обґрунтованість і правильність – це, в кінцевому підсумку, означає визначити, де закінчується методологія і починається методика... Образно кажучи, обґрунтованість характеризує напрям, в якому ми рухаємося, а правильність – траєкторію руху в даному напрямі» [8, с. 32].

Стійкість у «радянській традиції» за змістом відповідає надійності в «західній традиції».

На наш погляд, сьогодні, коли Україна інтегрується у світовий соціологічний простір, бере участь у міжнародних порівняльних дослідженнях (наприклад, Європейське соціальне дослідження – European Social Survey (ESS) чи Міжнародна програма соціального дослідження – International Social Survey Programm (ISSP)), більш адекватним є використання саме «західної» традиції, оскільки вона вже стала стандартом для багатьох країн і її використання полегшує міжнародну дослідницьку комунікацію, оскільки «без уніфікації термінології ускладнюється взаєморозуміння між дослідниками» [8, с. 7]. В.І. Паніотто зазначав: «З двох варіантів – використовувати як ім'я поняття «надійність» або взагалі не привласнювати імені якості інформації, більш вдалим я вважаю останній варіант – використовувати обґрунтованість (валідність) та надійність як рівноправні складові якості вимірювання. Але для того, щоб не виникало розбіжностей з уже ustalеним у вітчизняній літературі слововживанням, в інтересах пришвидшення процесу стандартизації соціологічної термінології ... я буду розуміти надійність як інтегральну, узагальнену характеристику якості соціологічної інформації, як її ім'я»

[8, с. 25–26]. Ми пропонуємо, з тих же міркувань що і В.І. Паніотто, – для узгодженості вітчизняної термінології з міжнародною, розуміти валідність та надійність як ключові характеристики якості вимірювання. Тим більше, що така тенденція вже намітилась. Прикладом використання термінології «західної» традиції оцінки якості соціологічного вимірювання є роботи Н.В.Паніної [наприклад, 7], а також останні статті у журналі «Соціологія: теорія, методи, маркетинг»: «Суб'єктивна надійність: теорія і методи вимірювання (ІСН)» [10] та «Теоретична валідність вимірювальної процедури і зміщення даних у соціологічному дослідженні» [3].

### **Як співвідносяться поняття валідність та надійність?**

Валідність та надійність спрямовані на оцінку різних характеристик вимірювання.

Валідність пов'язана більше з теоретичним рівнем дослідження, в той час, як надійність пов'язана з емпіричним рівнем дослідження.

Вимірювання є валідним якщо:

1) теоретичний концепт адекватно відображає існуючі у суспільстві відносини та зв'язки. Наприклад, класова схема, яка вимірює диференціацію людей за статусом зайнятості є валідною, якщо в основу схеми закладені ті характеристики, за якими в суспільстві справді розрізняється статус зайнятості людей;

2) вимірювальний інструмент адекватно відображає теоретичний концепт. Якщо використовувати приклад класової схеми, то вимірювальний інструмент адекватно відображає теоретичний концепт у тому разі, коли запитання анкети повною мірою відображають усі характеристики, закладені в теорію цієї класової схеми і респонденти розуміють ці питання саме так, як задумав дослідник, – тобто у відповідності до теоретичних посилок.

Чим ближче теоретичний концепт до реальності і чим менша відстань від теоретичного концепту до результатів вимірювання, тим вища валідність.

Надійність стосується того, якою мірою вимірювальний інструмент дає одні і ті ж самі результати при повторному вимірюванні. Оскільки у вимірюванні завжди присутні випадкові помилки, тому два вимірювання не можуть повністю збігатися одне з одним, але вони повинні давати близькі результати. Надійність пов'язана з тенденцією повторюваності при вимірюванні одних й тих самих соціальних феноменів декілька разів. Чим стійкіші результати вимірювання при повторних спробах, тим надійніше вимірювальна процедура. І навпаки – чим менш схожі результати при повторному вимірюванні, тим менш надійною є процедура вимірювання.

Таке розуміння валідності та надійності є загальноприйнятим підходом в європейській та американській традиції [18, с. 6].

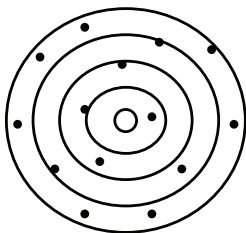
Такий підхід до валідності та надійності поділяє і російська дослідниця І. Ф. Дев'ятко: «Надійність вимірювання пов'язана з його стійкістю та відтворюваністю. Показник може вважатись надійним у тій мірі, в якій отримані

оцінки можуть бути відтворені на даній сукупності об'єктів вимірювання. Валідність вимірювання, визначається як характеристика відповідності вимірювання його меті, валідність визначає «чистоту вимірювання» теоретичного конструкту» [2, с. 167–172].

Валідність та надійність є пов'язаними характеристиками. Надійність є необхідною, але недостатньою умовою валідності. Розбіжність надійності та валідності вимірювального інструменту залежить від наявності систематичної помилки вимірювання. Якщо вимірювальний інструмент вимірює тільки заданий теоретичний концепт з точністю до випадкових помилок, то валідність буде дорівнювати надійності. Чим більша частка систематичного помилкового компонента, тим більшою буде різниця між надійністю і валідністю – надійність буде перевищувати валідність [18, с. 14].

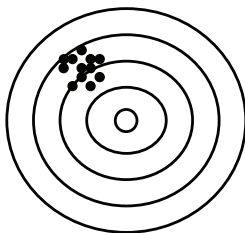
Отже, вимірювання може бути валідним тільки за умови надійності вимірювального інструменту, оскільки, у разі ненадійного інструменту кожен раз ми будемо отримувати різні результати, які не будуть відображати теоретичний концепт. У разі, якщо вимірювальний інструмент вимірює саме той теоретичний концепт, який він має вимірювати і постійно дає стійкі, повторювані результати, вимірювання є і надійним і валідним. Вищесказане можна проілюструвати на прикладі мішені – порівнявши соціологічне вимірювання з пострілами у мішень (рис. 1–3).

Рис. 1



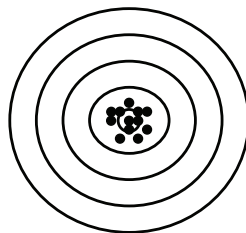
Низька надійність  
Низька валідність

Рис. 2



Висока надійність  
Низька валідність

Рис. 3



Висока надійність  
Висока валідність

На рис. 1 ми бачимо, що лучник не влучає в центр мішені, тобто не вимірюється те, що було заплановано виміряти, а це означає низьку валідність вимірювання. Також у пострілах немає ніякої системності – лучник потрапляє у різні частини мішені, тобто при повторних вимірюваннях вимірюється не одна і та сама характеристика, а різні, що означає низьку надійність вимірювання.

На рис. 2 ми бачимо, що лучник не влучає у центр мішені, що означає низьку валідність. В той же час, лучник влучає приблизно в одне і те саме місце, тобто при повторних спробах вимірюється одна і та сама, хоч і не запланована, характеристика, що означає високу надійність.

На рис. 3 лучник при повторних спробах влучає тільки в центр мішені, тобто вимірюється саме запланована характеристика і вона вимірюється у кожному вимірюванні. Це означає високу валідність і високу надійність вимірювання.

Одним з видів вимірювального інструменту в соціології виступає запитання анкети для кількісного опитування (на рівні із запитаннями для глибинних інтерв'ю, для фокус-груп, категоріями контент-аналізу, одиницями спостереження та ін.). Наведемо приклад невалідного і ненадійного запитання для кількісного опитування: *«Як Ви оцінюєте рівень соціальної напруги у Вашому регіоні?»* [6, с. 200]. У цьому запитанні присутні одразу два невизначених терміни: «соціальна напруга» і «регіон». Поняття «соціальна напруга» не має однозначного тлумачення в повсякденному житті. Дослідник може по-своєму тлумачити цю категорію, а різні респонденти по-своєму: хтось може розуміти під цим терміном мітинги, хтось незадоволеність рівнем життя і т.п. Також невизначеним є термін «регіон». Хтось може зрозуміти його як «моє село», «моє місто», хтось – як «область», а хтось – «як державу в цілому» [6, с. 200].

Отже дослідник не отримує інформації про той теоретичний концепт, який мав бути вимірний за допомогою даного запитання, тобто це запитання є невалідним. Також відповіді на це запитання при повторних вимірюваннях будуть нестійкими, оскільки кожен респондент буде сам інтерпретувати зміст цього питання, тобто це запитання не є надійним.

Для досягнення валідності та надійності вимірювальних інструментів дослідник має спочатку чітко окреслити зміст та межі теоретичного концепту, потім розробити чітку та вичерпну систему індикаторів, які відображають досліджуваний концепт, ретельно «перекласти» запитання з мови соціологічних категорій на буденну мову, провести тестування запитань на адекватність їх сприйняття респондентами.

Валідність та надійність вимірювальних інструментів досягається не одномоментно, а в багатьох різних дослідженнях, шляхом постійного уточнення та досягнення максимального наближення до теоретичного концепту, що вимірюється за допомогою конкретного вимірювального інструменту. Підвищення якості вимірювальних інструментів є колективним процесом і відповідальність за цей процес несуть як розробники вимірювальних інструментів, так і ті, хто їх використовуює [13].

**Висновки.** За змістом «радянська» та «західна» традиції оцінки якості соціологічного вимірювання є дуже схожими. В той же час у них використовується різний термінологічний апарат. На нашу думку, для уникнення плутанини при оцінці якості конкретних вимірювань та для полегшення комунікації між дослідниками, адекватним є застосування термінологічного апарату «західної» традиції. Особливо це важливо в міжнародних порівняльних дослідженнях, в яких Україна активно бере участь.

У відповідності до «західної» традиції ключовими характеристиками якості соціологічного вимірювання є валідність та надійність. Валідність та надійність спрямовані на оцінку різних характеристик вимірювання.

Валідність пов'язана більше з теоретичним рівнем дослідження, в той час як надійність пов'язана з емпіричним рівнем дослідження. Валідність це відповідність вимірювання його меті. Вимірювальний інструмент є валідним, якщо за його допомогою вимірюється те, що заплановано було виміряти. Надійність – це відтворюваність вимірювання. Надійність стосується того, в якій мірі інструмент вимірювання дає одні і ті ж самі результати при повторному вимірюванні.

Надійність є передумовою валідності. Вимірювальний інструмент може бути надійним і невалідним, але не навпаки. Підвищення якості вимірювальних інструментів є колективним процесом, і відповідальність за цей процес несуть як розробники вимірювальних інструментів, так і ті, хто їх використовує.

- 
1. *Волович В.И.* Надежность информации в социологическом исследовании : проблемы методологии и методики / В.И. Волович. – К., 1974.
  2. *Девятко И.Ф.* Методы социологического исследования / И.Ф. Девятко. – М., 2002.
  3. *Дембицький С.* Теоретична валідність вимірювальної процедури і зміщення даних у соціологічному дослідженні / С. Дембицький // Соціологія : теорія, методи, маркетинг. – 2008. – № 3. – С. 99–118.
  4. *Докторов Б.З.* О надежности измерения в социологическом исследовании / Б. З. Докторов. – Л., 1979.
  5. *Здравомыслов А.Г.* Методология и процедура социологических исследований / А. Г. Здравомыслов. – М. : Мысль, 1969.
  6. *Панина Н.В.* Избранные труды по социологии : в 3 т. / Н.В. Панина. Т. 1. Вопросы теории, методологии, технологии социологического исследования и профессиональной этики. – К. : Факт, 2008.
  7. *Панина Н.В.* Технологія соціологічного дослідження / Н.В.Паніна. – К. : Наук. думка, 1996.
  8. *Паниотто В.И.* Качество социологической информации : методы оценки и процедуры обеспечения / В.И. Паниотто. – К., 1986.
  9. *Саганенко Г.И.* Надежность результатов социологического исследования. – Л., 1983.
  10. Суб'єктивна надійність: теорія і методи вимірювання (ІСН) / Є. Головаха, А. Горбачик, Т. Любива та ін. // Соціологія: теорія, методи, маркетинг. – 2008. – № 1. – с. 166–188.
  11. *Ядов В.А.* Социологическое исследование : методология, программа, методы / В.А. Ядов. – М., 1972.
  12. American Educational Research Association, American Psychological Association and National Council on Measurement in Education. Standards for educational and psychological testing. Washington DC: American Educational Research Association, 1985.
  13. American Educational Research Association, American Psychological Association and National Council on Measurement in Education. Standards for educational and psychological testing. – Washington DC : American Educational Research Association, 1999.
  14. American Psychological Association. Standards for educational and psychological tests and manuals. – Washington, DC, 1966.
  15. American Psychological Association. Standards for educational and psychological tests and manuals. – Washington, 1974.
  16. American Psychological Association. Technical recommendations for psychological tests and diagnostic techniques // Psychological Bulletin [Supplement]. – 1954. – № 51 (2, part 2).
  17. *Bohrsted G.W.* (2000) Validity / G.W. Bohrnsted // Encyclopedia of Sociology, Second Edition, Volume 5. New York.

18. *Carmines E.G.* Measurement in the social sciences / E.G. Carmines, R.A. Zeller. – Cambridge University Press, 1980.
19. *Carmines E.G.* Reliability and Validity Assessment / E.G. Carmines, R.A. Zeller. – Beverley Hills : Sage, 1979.
20. *Cronbach L.J.* Construct Validity in Psychological Tests / L.J. Cronbach, P.E. Meehl // Psychological Bulletin. – 1955. – № 52. – P. 281–302.
21. *Cronbach L.J.* Construct Validity in Psychological Tests / L.J. Cronbach, P.E. Meehl // Psychological Bulletin. – 1955. – № 52. – P. 281–302.
22. Sage Handbook of social research methods, 2008.