

МЕТОДОЛОГІЯ І МЕТОДИ ЕМПІРИЧНИХ ДОСЛІДЖЕНЬ

УДК 303.833.4+303.712

НАДІЙНІСТЬ КОДУВАННЯ ПРОЕКТИВНИХ МАЛЮНКІВ: ПРИКЛАД ЗАСТОСУВАННЯ К КОЕНА

Н. В. Отріщенко

*Інститут соціології НАН України,
вул. Шовковична, 12, м. Київ, Україна, 01021,
trishky.natalka@gmail.com*

На прикладі роботи з проєктивними малюнками, виконаними респондентами під час інтерв'ю у 14 регіонах України, показано особливості перевірки надійності кодування бінарних категорій. Доведено проблемність використання відсотка згоди і висуває аргументи на користь к Коена. Міститься загальні рекомендації для проведення кодування візуальних матеріалів.

Ключові слова: к Коена, кодування, надійність, проєктивний малюнок.

Надійність та валідність у сучасній соціогуманітарній традиції виступають ключовими індикаторами якості даних, отриманих за допомогою тієї чи іншої методики. Робота з проєктивними методами, як вербальними, так і візуальними, передбачає особливий підхід до їхнього встановлення, адже саме критика їхньої надійності та валідності є однією з перепон для їхнього активного застосування. Існування великої кількості неоднозначностей у первинному матеріалі, недостатньо чітко прописані категорії кодування – це може стати причиною появи ненадійних даних. Усвідомлюючи цей факт, починаючи з 1991 року «Journal of Personality Assessment» почав вимагати від авторів, що використовують результати тесту Роршаха, доказів, що мінімум 80 % оцінок кодувальників по ключових критеріях сходяться між собою [1, с. 476]. Ця стаття має на меті проілюструвати особливості перевірки надійності кодування більш структурованого матеріалу – проєктивних малюнків «свого простору», виконаних у межах проєкту «Регіон, нація та інше»¹ – а також показати, що не завжди варто орієнтуватися виключно на відсоток згоди, а треба застосовувати більш гнучкі показники.

Ф. Белл, описуючи роботу з масивами візуальних матеріалів, розглядає надійність як міру узгодженості кодувальників, що може бути продемонстрована через кореляцію між

¹ Дослідження ініційоване університетом Сент-Галлена (Швейцарія) та охоплювало п'ять тематичних проєктів, присвячених історії, літературі, мові, релігії та економіці сучасної України. Цей проєкт передбачав як кількісне опитування, так і якісне інтерв'ювання та фокус-групові дискусії. Координація якісним соціологічним дослідженням відбувалася у Центрі міської історії Центрально-Східної Європи (Львів, Україна) під керівництвом доц., к.с.н. Вікторії Середи.

оцінками однієї вибірки відповідних змінних групою кодувальників (інтер-кодувальна (inter-coder) надійність) чи одним кодувальником за різних обставин (інтра-кодувальна (intra-coder) надійність) [4, р. 21]. Інтер-кодувальна надійність – це широкоживане поняття для визначення міри того, наскільки незалежні кодувальники однаково оцінюють характеристики повідомлення чи артефакту і при цьому приходять до однакових висновків. Як синонім у цьому контексті застосовують поняття «згода» (agreement), яким вимірюють рівень того, наскільки кодувальники схильні давати ідентичні оцінки об'єктам [9]. Рекомендований рівень згоди між кодувальниками у контент-аналітичних дослідженнях – 0,9 або 90 % співпадінь [4, р. 22]. Коефіцієнти 0,9 і вище майже завжди прийнятні, 0,8 і вище – прийнятні у більшості випадків і 0,7 і вище – можуть прийматися у деяких дослідженнях і для деяких коефіцієнтів. У той час як відсоток згоди виступає «ліберальним» показником надійності, існують також «консервативні» показники (на зразок π Скотта, κ Коена та α Кріппендорфа) [9].

С. Морозов, говорячи про роботу з проєктивними методами у психологічному дослідженні, виділяє такі рівні контролю інтерпретативної надійності [3]: 1) надійність підбору первинної інформації для аналізу (адже, по-перше, протоколюються не всі реакції, і, по-друге, для подальшої інтерпретації використовується далеко не все: на якісь блоки інформації дослідник звертає більше уваги, якісь інші – ігнорує); 2) надійність ідентифікації показників (максимально можлива однозначність інтерпретації показника); 3) надійність висновку, що опирається на сукупність окремих показників; 4) однозначність остаточного діагнозу. Для соціології загалом і авторського дослідження зокрема особливо важливими є перші три рівні контролю, адже вони відповідають традиційним етапам у роботі з даними в соціології (збирання даних, аналіз та інтерпретація, науковий висновок).

Отож, надійність кодувальника або оцінювальника – одна з найважливіших складових перевірки надійності в роботі з проєктивними методами, адже саме кодувальник присвоює числові значення змістовим показникам. Показником надійності кодувальника, на думку А. Анастасі, виступає коефіцієнт кореляції між оцінками двох кодувальників [1, с. 119]. Однак такий підхід до встановлення надійності критикується представниками медійних досліджень, які застосовують коефіцієнти кореляції ϕ Кендалла, r Пірсона та r_s Спірмена лише для встановлення зв'язку між змінними, адже сама природа кодування не передбачає тієї незалежності змінних, якої вимагають статистичні процедури [6, р. 242–257].

У своїй статті М. Ломбард, Дж. Снайдер-Дач та К. Брейкен аналізують 200 досліджень зі сфери масової комунікації, реалізованих у період з 1994 по 1998 роки, і вказують на те, що далеко не у всіх них (у 69 % випадків, тобто в 137 текстах) фігурують показники інтер-кодувальної надійності [8]. Тобто вказування цього показника у публікаціях досліджень на основі контент-аналізу не так давно не було нормою академічної культури навіть на Заході. Аналізуючи підходи до оцінки інтер-кодувальної надійності К. Кріппендорф, М. Ломбард, Дж. Снайдер-Дач радять не використовувати відсоток згоди між оцінювальниками, хоч він і найпоширеніший і найпростіший для обрахунку, адже вказаний індекс ігнорує можливість того, що оцінки кодувальників можуть збігатися випадково. Тому В. Скотт у 1955 році запро-

понував інший показник, який ґрунтується на відношенні очікуваних та реальних співпадінь оцінок [10]. Через декілька років свій показник – k Коена – обрахував Я. Коен. Його обчислення здійснюється за такою формулою [5, р. 214], у якій $\text{Pr}(o)$ – спостережуваний рівень згоди, а $\text{Pr}(e)$ – рівень згоди, очікуваний випадково:

$$k = \frac{\text{Pr}(o) - \text{Pr}(e)}{1 - \text{Pr}(e)}.$$

Важливо, що і π Скотта, і k Коена дозволяють оцінити інтер-кодувальну надійність лише для двох кодувальників та застосовуються до категоріальних даних. Для обчислення надійності більше, ніж двох оцінювальників, використовують коефіцієнти типу k Флейша [6, р. 32] чи коефіцієнт множинної кореляції W Кендалла. Особливості обрахунку k Коена та π Скотта відповідають логіці лог-лінійної моделі mover-stayer (яку не надто доречно перекладають як «мінливі-віддані»), що застосовувалась для перевірки стабільності відповідей респондентів у панельних дослідженнях [2, с. 117–130].

Ще один сучасний дослідник, К. Кріппендорф, наголошує на проблемності використання коефіцієнтів асоціації, кореляції та консистентності для встановлення надійності контент-аналізу [7, р. 412]. Його рекомендації полягають, по-перше, у розробленні детальних та зрозумілих інструкцій для кодувальників та їхній незалежній одне від одного роботі, по-друге, у застосуванні коефіцієнтів згоди (на зразок π Скотта для номінальних шкал) замість коефіцієнтів кореляції та відсотка згоди між кодувальниками, по-третє, у встановленні прийнятного рівня згоди i , нарешті, у виваженій перевірці усіх значимих висновків на предмет надійності [7].

До схожих висновків приходять М. Ломбард, Дж. Снайдер-Дач та К. Брейкен. Серед важливих рекомендацій стосовно вимірювання інтер-кодувальної надійності при контент-аналізі вони також вказують неможливість застосування у цій процедурі лише відсотка згоди між кодувальниками, коефіцієнтів кореляції та тесту k^2 [9]. Дослідники мотивують це іншою природою даних, яка не передбачає коваріації. Як показники інтер-кодувальної надійності вони пропонують спеціально розроблені в комунікаційних дослідженнях π Скотта, k Коена та α Кріппендорфа [9]. Коефіцієнт k Коена можна обрахувати за допомогою програми SPSS та Microsoft Excel, α Кріппендорфа – у програмі Krippendorff's alpha 3.12a (версія станом на лютий 2015 року).

Отож, забезпечення та перевірка надійності кодування передбачає два основні напрями роботи: **змістовий** і **статистичний**. Змістовий блок об'єднує формування інструкцій та навчання кодувальників. Рекомендації стосовно цього знаходимо у Ф. Белла [4, р.17–22]:

- 1) чітко окреслити змінні та їхні характеристики. Вони мають бути взаємовиключними і вичерпними (*mutually exclusive and exhaustive*), адже основний виклик контент-аналізу – визначити змінні та їхні категорії так, щоб кодувальники могли їх стабільно й однозначно класифікувати;
- 2) навчити кодувальників застосовувати окреслені критерії для конкретних випадків, тобто якісно провести інструктаж;
- 3) перевірити інтер-кодувальну надійність *post factum*.

Щоб проілюструвати можливості встановлення надійності кодування проєктивних малюнків, використано дані інтерв'ювання для всеукраїнського дослідження «Регіон, нація та інше: Інтердисциплінарне та міжкультурне переосмислення України». Метод проєктивного малюнку «свого простору» поєднувалися з глибинними інтерв'ю для тематичного підпроєкту, що мав на меті виявити особливості історичної пам'яті та локальної ідентичності громадян України. Польовий етап проходив упродовж 2012–2014 років в 14 областях України. У кожній із них було обрано місто (адміністративний центр), районний центр та село, де було опитано вісім (у випадку міста та містечка) або шість (у випадку села) осіб. Питальник був складений як напівструктурований і об'єднував блоки запитань, що стосувалися біографії респондента, особливостей його чи її зв'язку з місцем проживання, ознайомленості з локальною історією, а також специфіки історичного нарративу у тому чи іншому регіоні. Проєктивний малюнок завершував інтерв'ю. Респонденти могли відмовитися малювати (скажімо, якщо це була старша людина), тоді його чи її просили виконати вербальний опис «свого простору», однак для подальшого опрацювання використано лише «фізичні» малюнки.

У процесі підготовки до кодування проєктивних малюнків було розроблено спеціальні таблиці, які так само конструювалися в процесі наукових дискусій з працівниками відділу методології та методів соціології Інституту соціології НАН України. Згодом проводили інструктаж та навчання групи кодувальників. Необхідність розроблення детальної інструкції та проведення ґрунтовної роботи з оцінювальниками пояснюється серед іншого і тим, що «ми впізнаємо лише те, що уже знаємо» [11, р. 94]. Без проведення такої попередньої роботи кодувальники можуть просто «не помічати» ті категорії, наявність яких вони мали б фіксувати. Після виконання ними кодування здійснювалась статистична перевірка надійності шляхом обрахунку відсотку згоди та к Коена (для двох кодувальників та дихотомічних змінних) та інтра-класового коефіцієнта кореляції ICCS, який використовують в ситуації, коли працювало більше, ніж два кодувальники номінальних змінних. Тобто було дотримано покроковості процедур, на які вказував Ф. Белл.

До проведення методологічного експерименту з кодування проєктивних малюнків були залучені студенти кафедри історії та теорії соціології Львівського національного університету ім. Івана Франка, які проходили виробничу практику у Центрі міської історії Центрально-Східної Європи у 2013–2014 роках. Це були студенти четвертого та п'ятого (спеціалісти та магістри) курсів. У дослідженні надійності кодування використано результати кодування п'яти осіб (четверо дівчат і один хлопець). Перед початком роботи усі вони проходили інструктаж, під час якого на прикладі конкретних малюнків чи інтерв'ю студенти спершу під керівництвом інструктора, а згодом самостійно присвоювали значення тим чи іншим категоріям.

Кодувальний лист містив перелік змінних, кожна з яких мала власне віяло варіантів відповіді, а також коментарі стосовно того, що саме в конкретному випадку підлягає кодуванню. Загалом, більшість змінних можна поділити на дві категорії: категорія наявності (кодувальник вказує на те, чи присутній певний елемент на малюнку, чи ні) та категорія судження (кодувальник приймає рішення стосовно інтенсивності проявлення певної характеристики, або ж кодує характеристику, виходячи із заданого віяла).

Успішність кодування категорії наявності безпосередньо залежить від випадкових, як от уважність чи втома, та систематичних особистих характеристик. До других можна зарахувати загальну культурну компетентність, скажімо, кодувальник має вміти побачити на малюнку простору ті сакральні об'єкти, які належать не лише до його чи її релігійної традиції. Яскравим прикладом такої похибки були результати кодування наявності дому на проективному малюнку – одна із кодувальниць була схильна бачити його вдвічі частіше, ніж це робили інші. Стосовно категорії судження, то тут особливо важливо враховувати систематичні «зсуви», які можуть бути у кодувальників. Наприклад, дехто може регулярно присвоювати вищі або нижчі значення залежно від власних переконань та оцінок. Саме тому для проектів, які ґрунтуються на кодуванні, дуже важливо залучати декількох кодувальників, щоб виявити та мінімізувати вплив як випадкових, так і систематичних чинників.

Ця стаття має на меті проілюструвати особливості кодування категорій наявності. Їм у кодувальних таблицях відповідали змістові запитання, які передбачали бінарну відповідь «так» / «ні», на зразок то, чи зображено на малюнку вулицю, парк або індустріальний об'єкт. У процесі перевірки надійності кодування цих змінних було використано k Коена та для зівставлення – відсоток згоди. На відміну від тесту k_2 , який оперує частотами, k Коена побудована на ймовірностях і вказує на те, чи співпадіння в оцінках кодувальників не могло статися випадково. Цей показник може варіювати від -1 (ідеальне неспівпадіння) до 1 (ідеальне співпадіння), а 0 означає те, що кодувальники могли дійти до згоди випадково. Ілюстрація ситуацій з розподілу оцінок двох кодувальників – k_1 і k_2 – запропонована у таблиці 1.

Таблиця 1

Моделювання ситуацій при різних значеннях каппи

Ідеальне неспівпадіння, k			Ідеальне співпадіння, k			Випадковий збіг, k		
	K2: так	K2: ні		K2: так	K2: ні		K2: так	K2: ні
K1: так	0	50	K1: так	50	0	K1: так	25	25
K1: ні	50	0	K1: ні	0	50	K1: ні	25	25

Обрахунок k Коена здійснюється на основі ймовірного підходу до визначення рівня згоди між кодувальниками, тому в ідеальній ситуації значення присвоєних категорій мали б бути незалежними. Проте оскільки студенти оцінювали одні й ті самі малюнки, це обмеження потрібно враховувати. Як можливе рішення дослідники також пропонують застосовувати коефіцієнти АС1 та α Аїкіна [6, р. 70], які засновані на більш реалістичному припущенні, що тільки частина спостережуваних оцінок може збігтися випадково, однак в основі свого обрахунку вони мають той самий алгоритм, що і k Коена, і різняться від неї хіба обрахунком очікуваних частот. Саме тому це дослідження орієнтуватиметься на k Коена, хоча про її обмеження також потрібно пам'ятати під час інтерпретації результатів.

Інтерпретація значення k Коена зазвичай здійснюється за такою схемою [12, р. 362]:

- 0,01-0,20 – відсутність згоди;
- 0,21-0,40 – слабкий рівень згоди;
- 0,41-0,60 – середній рівень згоди;
- 0,61-0,80 – значний рівень згоди;
- 0,81-0,99 – майже ідеальна узгодженість.

Натомість, ми пропонуємо замість запропонованих п'яти виділити три інтервали, пам'ятаючи, також, про їхню умовність:

- 1) менше 0,33 – недостатня узгодженість;
- 2) 0,34–0,66 – задовільна узгодженість;
- 3) понад 0,67 – достатня узгодженість.

За таких умов значення менше 0,33 однозначно відкидаються, понад 0,67 – приймаються, а стосовно значень в інтервалі між 0,34 та 0,67 рішення залежить від: 1) рівня узгодженості між іншими кодувальниками; 2) наповненості категорій – ідеться про вимогу, яка необхідна для тесту χ^2 , однак в контексті k Коена не фігурує – частоти не менше 5 у комірках, стосовно яких перевіряється рівень згоди; 3) наближеності до одного із крайніх значень; 4) відсотка згоди між кодувальниками. Щодо останнього, то також потрібно враховувати існування «парадоксу каппи», коли низьким значення k Коена відповідає значний відсоток згоди, про який згадують

Е. Віера та Дж. Гаррет [12, р. 362]. Саме тому прийняття рішення щодо задовільності рівня згоди має ґрунтуватися на комплексі умов, а не лише на одній із них.

Щоб проілюструвати особливості роботи з k Коена, наведемо декілька таблиць із результатів методологічного експерименту із залучення декількох кодувальників. Кожна комірка містить показник k Коена для двох кодувальників, наприклад, $K1$ і $K2$ або $K3$ і $K5$, а для порівняння в дужках наведено відсоток згоди між ними, котрий, хоч і не є надто надійним показником, однак часто використовується у дослідженнях, що ґрунтуються на присвоєнні кодувальниками значень певним категоріям. Також зазначимо, що усі k Коена є статистично значимими на рівні $p=0,001$.

Приклад, який ілюструє значення k Коена при кодуванні бінарної змінної (наявність людей на малюнку), наведено у таблиці 2. Комірка на перетині колонки та рядка вказує на рівень згоди між двома кодувальниками. Наприклад, значення, що виділене жирним шрифтом – 0,552 (96,0) – показник згоди (k Коена та відсоток співпадінь у присвоєні значень) між $K2$ та $K4$, другим та четвертим кодувальником.

Таблиця 2

Згода стосовно зображення людей (так або ні)

	K1	K2	K3	K4
K2	0,687 (97,3)			
K3	0,604 (97,7)	0,618 (96,9)		
K4	0,517 (96,9)	0,552 (96,0)	0,853 (99,1)	
K5	0,479 (95,5)	0,510 (94,7)	0,508 (96,0)	0,552 (96,0)
	Карра (%)	Карра (%)	Карра (%)	Карра (%)

Отож, при зівставленні кодування присутності людей на малюнках каппа у 8 випадках із 10 входять до інтервалу задовільної надійності, а у двох випадках – достатньої, у той час як відсоток згоди становить понад 95 % майже в усіх випадках. Цей результат вказує на те, що учасники методологічного експерименту легко ідентифікують антропоморфні фігури на малюнках «свого простору». Щоправда, люди присутні лише на 2,7 % – 5,8 % усіх проєктивних малюнків, а тому замість відсотка згоди варто акцентувати увагу лише на значеннях k Коена.

Щодо оцінок наявності на проєктивних малюнках культових споруд (таблиця 3), то у цьому випадку кодувальники продемонстрували дуже високу однаковість: усі значення k Коена потрапляють до третього інтервалу (достатня узгодженість), а у половині випадків вони становлять понад 0,8. Це можна пояснити тим, що існує поширений спосіб зображення, наприклад, церков, які маркуються хрестом на куполі, а тому кодувальники легко їх відрізняють. Аналогічно, є досить поширені способи маркування мечетей (зображення мінарету з півмісяцем поряд з будівлею) та синагог (будівля із зіркою Давида). До того ж, респонденти інколи підписували зображені об'єкти, що значною мірою полегшувало процес кодування. Нарешті, ця категорія досить чітко окреслена, а тому класифікувати церкву, мечеть або синагогу як культову споруду не вимагає значних зусиль, хоча і залежить від культурної компетентності. Будівлі цього типу, згідно з оцінками учасників методологічного експерименту, присутні на кожному п'ятому малюнку.

Таблиця 3

Згода стосовно зображення культових споруд (так або ні)

	K1	K2	K3	K4
K2	0,823 (93,7)			
K3	0,855 (95,1)	0,832 (94,2)		
K4	0,836 (94,7)	0,732 (91,1)	0,846 (95,1)	
K5	0,682 (89,3)	0,713 (91,5)	0,742 (91,5)	0,746 (92,0)
	Каппа (%)	Каппа (%)	Каппа (%)	Каппа (%)

Досить цікава ситуація виникла із кодуванням дому (таблиця 4). У порівнянні з рештою кодувальників, які вказували на його присутність у більш, ніж третині випадків, одна із учасниць експерименту була схильна бачити дім на кожному третьому із чотирьох малюнків. Врахування цієї системної похибки і вилучення одного рядка з аналізу призводить до зростання значення k Коена, яка тепер варіює в інтервалі від 0,478 до 0,706 і її середнє значення становить 0,612 (без вилучення одного із кодувальників середнє значення k Коена для усіх п'яти кодувальників становить 0,485). Цей випадок показує, як систематична похибка впливає на надійність результатів загалом.

Таблиця 4

Згода стосовно зображення дому (так або ні)

	K1	K2	K3	K4
K2	0,694 (84,9)			
K3	0,506 (75,5)	0,586 (80,8)		
K4	0,478 (74,2)	0,706 (86,6)	0,699 (86,2)	
K5	0,364 (67,5)	0,278 (58,6)	0,293 (60,0)	0,243 (56,0)
	Карра (%)	Карра (%)	Карра (%)	Карра (%)

Стосовно інших елементів зображення, то у випадку спільності оцінок наявності на малюнках громадських споруд каппа варіює в інтервалі 0,375–0,681, індустріальних об'єктів – 0,427–0,799, пам'ятників – 0,489–0,780, парків та скверів – 0,498–0,812. В усіх випадках значення належать до задовільного рівня надійності, хоча у випадку громадських споруд та індустріальних об'єктів один із кодувальників вказував значно меншу кількість одиниць, аніж решта. Це можна пояснити тим, що ці типи споруд інколи важко ідентифікувати, а самі поняття можуть включати в себе різнотипні об'єкти.

Під час роботи з малюнками також була присутня бінарна змінна, яка фігурувала не як категорія наявності, а як категорія судження: згода стосовно розміщення малюнку на аркуші (по центру або збоку). У її випадку каппа становила від 0,444 до 0,620, що відповідає варіативності в межах задовільного.

Таблиця 5

Згода стосовно розміщення на аркуші (центр або збоку)

	K1	K2	K3	K4
K2	0,444 (75,1)			
K3	0,620 (86,6)	0,470 (76,0)		
K4	0,527 (79,5)	0,591 (80,4)	0,575 (81,3)	
K5	0,576 (84,8)	0,550 (79,5)	0,591 (84,8)	0,516 (78,6)
	Карра (%)	Карра (%)	Карра (%)	Карра (%)

Отож, результати методологічного експерименту стосовно кодування проєктивних малюнків ілюструють задовільний і достатній рівень надійності даних, що стосуються категорії наявності. Найбільша однастайність проявляється в тих випадках, коли існує певний «звичний» спосіб малювання того чи іншого об'єкта, як у випадку із зображенням людей чи культових споруд. Значною мірою досягнути вищого рівня надійності допомагають підписи об'єктів та записані коментарі респондентів під час виконання цього завдання.

На цьому етапі можна зробити декілька **загальнометодологічних висновків-рекомендацій**, які безпосередньо стосуються надійності кодування проєктивних малюнків:

1) перед початком кодування:

- 1.1) здійснення інтерпретативного аналізу вибірки зображень для уточнення категорій у кодувальних таблицях;

- 1.2) апробація кодувальних таблиць на іншій вибірці матеріалів із проведеного дослідження;
 - 1.3) обговорення кодувальних таблиць у середовищі дослідників-спеціалістів у заданій темі, яка підлягатиме кодуванню;
 - 1.4) розроблення детальної інструкції для роботи із кодувальними таблицями;
 - 1.5) проведення навчання групи кодувальників з використанням реальних випадків з масиву зображень;
- 2) у процесі кодування:**
- 2.1) контроль за самостійністю реалізації поставленого завдання;
 - 2.2) регулярна перевірка якості роботи (наприклад, кожен учасник методичного експерименту один із 50 малюнків кодує в присутності інструктора);
 - 2.3) проведення обговорення з групою кодувальників після завершення їхньої роботи (виявлення проблемних моментів, які треба враховувати під час здійснення подальшого аналізу результатів);
- 3) під час аналізу закодованих матеріалів:**
- 3.1) розроблення оптимальної схеми прийняття рішення стосовно задовільності рівня згоди, наприклад, три рівні значення k – незадовільна, задовільна, достатня;
 - 3.2) врахування наповненості категорій, особливо у випадку незначного або надто великого наповнення значенням комірок у таблицях, на основі яких обраховується рівень згоди, і вибір на користь k , а не відсотка згоди;
 - 3.3) зіставлення з одновимірними частотними та відсотковими розподілами, які допомагають виявити системні відмінності у інтерпретації;
 - 3.4) залежність якості кодування від природи даних (наприклад, більша згода стосовно чітко окреслених культурних категорій, як-от культурні споруди);
- 4) під час підбирання матеріалів для публікації:**
- 4.1) детальний опис методології проведення дослідження і процедури кодування його результатів;
 - 4.2) наведення різних прикладів, котрі як підтверджують, так і спростовують авторські твердження.

Щодо перспектив для подальшого дослідження, то актуальною залишається можливість зваження значень відповідно до значень k , подальше експериментування з процедурою кодування і статистичними показниками надійності роботи кодувальників, а також уточнення розробленої схеми роботи за допомогою інших емпіричних досліджень.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Анастаси А. Психологическое тестирование. Седьмое издание / А. Анастаси, С. Урбина. – СПб. : Питер, 2009. – 688 с.
2. Аптон Г. Анализ таблиц сопряженности / Г. Аптон. – М. : Финансы и статистика, 1982. – 143 с.
3. Морозов С. Засоби контролю діагностичних якостей психологічних тестів : навч. посібник / С. Морозов. – К. : ІСДО, 1994. – 68 с.

4. *Bell P.* Content analysis of visual images / P. Bell / Handbook of Visual Analysis / Ed. by T. Van Leeuwen and C. Jewitt. – London : Sage Publications, 2003. – P. 10–34.
5. *Cohen J.* Weighted Kappa: Nominal Scale Agreement with Provision for Scaled Disagreement of Partial Credit / Jacob Cohen // Psychological Bulletin. – 1968. – 70 (4). – P. 213–220.
6. *Gwet K. L.* Handbook of Inter-Rater Reliability: The Definitive Guide to Measuring the Extent of Agreement Among Multiple Raters. 3rd Edition / Kilem Li Gwet. – Gaithersburg: Advanced Analytics, LLC, 2012. – 294 p.
7. *Krippendorff K.* Reliability in Content Analysis: Some Common Misconceptions and Recommendations / Klaus Krippendorff // Human Communication Research. – 2004. – 30 (3). – P. 411–433.
8. *Lombard M.* Content analysis in mass communication research: An assessment and reporting of intercoder reliability / M. Lombard, J. Snyder-Duch, C. C. Bracken // Human Communication Research. – 2002. – 28. – P. 587–604.
9. *Lombard M.* Practical Resources for Assessing and Reporting Intercoder Reliability in Content Analysis Research Projects / M. Lombard, J. Snyder-Duch, C. C. Bracken [Electronic resource]. – Acces mode: http://ils.indiana.edu/faculty/hrosenba/www/Research/methods/lombard_reliability.pdf.
10. *Scott W.* Reliability of content analysis: the case of nominal scale coding / William Scott // Public Opinion Quarterly. – 1955. – 19(3). – P. 321–325.
11. *Van Leeuwen T.* Semiotics and Iconography / Theo Van Leeuwen / ed. by T. Van Leeuwen and C. Jewitt. – London : Sage Publications, 2003. – P. 92–119.
12. *Viera, A.* Understanding Interobserver Agreement: The Kappa Statistic / A. J. Viera, J. M. Garrett // Research Series: Family Medicine. – 2005. – 37(5). – P. 360–363.

Стаття надійшла до редколегії 04.08.2015

Прийнята до друку 12.08.2015

RELIABILITY OF CODING PROJECTIVE DRAWINGS: EXAMPLE OF USING COHEN'S K

N. V. Otrishchenko

*Institute of Sociology of National Academy of Sciences of Ukraine,
Shovkovychna Str., 12, Kyiv, Ukraine, 01021,
trishky.natalka@gmail.com*

While ensuring the reliability of coding visual materials (and projective drawing is an example of such type of data), researcher has to take into account two main lines of work: informative and statistical. Informative block integrates preparation of reliable instructions and coding tables, which had to be tested and discussed with experts, providing trainings for group of independent coders, and checking the quality of their work-in-progress and overall results. Statistical block involves decisions about mathematical procedures, which would be used for testing reliability, and relevant calculations.

Author reveals problems with using percentage of agreement between independent coders and gives reasons in favor of Cohen's k. This statistical coefficient was developed by Jacob Cohen in 1960s, and it is commonly used in mass-media studies. Cohen's k is calculated as ratio of the difference between observed and expected by chance percentage

of agreement between coders and the difference between one and expected percentage of agreement. This coefficient could be used as a measure of agreement between only two coders, and it is designed for work with nominal variables.

Author shows how Cohen's k as reliability measure could be used for analyzing coded projective drawings. Those drawings were created during interviews in 14 Ukrainian regions, which were collected in 2012–2014 for «Region, nation and beyond» project, founded by University of Saint Gallen in Switzerland. Projective drawings depicted informant's perception of his or her «own space». Group of five independent coders (BA and MA students from the department of history and theory of sociology from Ivan Franko Lviv National University) worked with two types of variables: categories of presence (binary nominal) and categories of judgment (nominal or ordinal). Cohen's k was calculated for categories of presence, and the percentage of agreement was used as a comparison to it. Also author developed her system of making decisions about the reliability of coding based on the meaning of Cohen's k .

Finally, this article contains general guidelines for coding visual materials according to the stage of the research: from developing coding tables to preparing materials for publication.

Keywords: Cohen's k , coding, reliability, projective drawing.