

RECENZJE

Zbigniew Sawiński
Polska Akademia Nauk

Pasjonaci instrukcji obsługi pralki

The Reviewer's Guide to Quantitative Methods in the Social Sciences. Red. Gregory R. Hancock i Ralph O. Mueller. New York and London: Routledge 2010, 432 s. + xvi.

Książka zaintryguje każdego, kto weźmie ją do ręki. Prawie pół tysiąca stron, 31 rozdziałów, 45 autorów. Rozdziały uporządkowane nie według logiki wyvodu, lecz alfabetycznie – według tytułów. I wreszcie adresat książki. Nie są nim badacze stosujący metody ilościowe w codziennej pracy. Nie są nim również studenci, którzy chcieliby zgłębić tajniki owych metod. Książka adresowana jest do recenzentów publikacji naukowych. Ma pomóc w rzetelnej i wyważonej ocenie, jak dalece proponowane zastosowania metod ilościowych są zgodne z zasadami sztuki.

Każdy z rozdziałów prezentuje jedną metodę ilościową. Ich wewnętrzny układ jest identyczny i rozpoczyna się od krótkiej charakterystyki omawianej metody. Następnie, w tabeli, podane zostają właściwości metody, które należy uwzględnić w wprowadzeniu do artykułu, przy omawianiu stosowanych technik badawczych, przy omawianiu uzyskanych wyników oraz w części podsumowującej – zwanej „dyskusją”. Pozostała część każdego rozdziału poświęcona jest wyjaśnieniu powodów, dla których stosowaną metodę należy w publikacji naukowej opisać tak, a nie inaczej.

Przypuszczam, że w tym miejscu wielu Czytelników straci zainteresowanie książką, którą chciałbym zaprezentować na łamach „Studiów Socjologicznych”. Instrukcja obsługi pralki wydaje się bardziej fascynująca! Powody zniechęcenia dość łatwo wskazać. Wąskie ujęcie, formalizm wykładu, specyficzny adresat. Z książki niewiele można się dowiedzieć, gdy omawianych metod nie zna się wcześniej. Czy nie lepiej skorzystać z monografii poświęconych poszczególnym metodom? Chociażby z popularnych wśród badaczy zielonych książeczek wydawnictwa Sage?

Otóż, i tak, i nie. Bez wątplenia monografie wyjaśniają istotę omawianych metod, określają dziedzinę ich zastosowań oraz sposób interpretacji uzyskanych wyników. Monografie uczą, jak samodzielnie stosować metody zgodnie ze standardami przyjętymi w nauce. Równocześnie jednak monografie muszą dowieść użyteczności opisywanych metod. Przeglądając spektakularne przykłady, w których wszystko układa się w sposób wyśmienity, łatwo do omawianej metody stracić dystans.

O zgrozo, obserwuję to na własnym przykładzie. W codziennej pracy preferuję metody, które poznałem wiele lat temu, na początku kariery badacza. I nie będę ukrywał, że metod tych nadużywam. Co więcej, podświadomie oczekuję, że inni łatwo pojmą to, co uzyskałem dzięki ich zastosowaniu. I, chyba, oczekuję niesłusznie. Inni przywykli korzystać z innych metod. Rozmawiamy przez to jak gęś z prosięciem.

W nauce coraz częściej zwraca się uwagę na fakt, że równie ważne jak samo uzyskanie rezultatu, jest jego zakomunikowanie – czy wręcz „wypromowanie”. Wydawane w ostatnich latach podręczniki metodologii badań często uzupełniane są rozdziałami poświęconymi sposobom prezentacji uzyskanych wyników. Fragmenty takie można znaleźć chociażby w *Metodologii badań psychologicznych* Jerzego Brzezińskiego (2007: 599–603) czy w przetłumaczonych na język polski podręcznikach badań społecznych Earla Babby’go (2008: 477–498) i badań jakościowych Davida Silvermana (2009: 341–424). Uważni Czytelnicy być może uznają, że odwołując się do Silvermana pokręciłem coś z numerami stron. Nie pokręciłem. David Silverman poświęcił około jednej piątej podręcznika prowadzenia badań jakościowych sposobom skutecznego komunikowania ich wyników! Kwestie te uwypukla się również w podręcznikach metod analizy danych. Przykładem jest recenzowana rok temu w „Studiach” (Sawiński 2010) książka Donalda Treimana *Quantitative Data Analysis* (Treiman 2009). Autor nie tylko poświęcił osobny rozdział kwestiom prezentacji wyników badań, lecz również dzieli się swoim doświadczeniem w tym zakresie omawiając każdą z metod.

Znaczenie, jakie obecnie przykładą się do skutecznego komunikowania uzyskanych wyników, wyznacza perspektywę, z której warto spojrzeć na książkę pod redakcją Hancocka i Muellera. Niekiedy, prezentując wyniki, badacze nadmiernie eksponują istotę stosowanej metody, przez co odbiorcy trudno jest wydobyć sens uzyskanych rezultatów. Z drugiej strony, opis zastosowanej metody bywa zbyt powierzchowny, co zdarza się zwłaszcza wtedy, gdy badacz metodę dobrze zna i stosuje kolejny raz. Lakoniczny opis wywołuje wątpliwości, czy metoda została zastosowana w sposób uprawniony i czy formułowane wnioski są należycie uzasadnione. Złoty standard sposobu prezentacji metody mieści się gdzieś pomiędzy tymi dwiema skrajnościami. Hancock i Mueller starają się trafić właśnie w ten punkt.

Książka obiecuje w tytule, że dotyczy zastosowań metod ilościowych w naukach społecznych. To dość enigmatyczne pojęcie, toteż omówienie zawartości książki poprzedzę analizą dyscyplin, które reprezentują autorzy poszczególnych rozdziałów. Pozwoli to lepiej zrozumieć dobór i sposób ujęcia uwzględnionych metod, a także ocenić przydatność recenzowanej książki dla socjologów.

Otóż, na 45 autorów aż 25, czyli ponad połowa, jako afiliację wymienia departament psychologii ogólnej, psychometrii bądź psychologii edukacyjnej. Druga grupa, licząca 10 autorów, zajmuje się badaniami edukacyjnymi i pomiarem w edukacji. Statystycy i metodolodzy badań lokują się na trzecim miejscu (7 autorów). Socjolog zaś jest jeden. To Paul D. Allison – postać znana zarówno z licznych publikacji dotyczących metod analizy danych, jak również z wkładu w badania nad nierównościami osiągnięć w nauce. Jest on jednak samotnikiem w gronie autorów, którzy na zadania nauk społecznych patrzą poprzez pryzmat badań eksperymentalnych, tworzenia wszelakiego rodzaju testów czy oceny skuteczności terapii stosowanych w psychologii i w medycynie. Należy się przez to spodziewać, że w książce odnajdziemy niewiele metod, które pozwalają osadzić badane fakty w kontekście zjawisk i procesów społecznych – co stanowi w gruncie rzeczy o istocie refleksji socjologicznej.

Jeśli w ten sposób pojmujemy potrzeby socjologa, to nie pomylimy się zbyt w przewidywaniach dotyczących zawartości książki. W sumie aż 18 spośród 31 rozdziałów dotyczy budowy i testowania wskaźników, bądź łączenia pojedynczych wskaźników w bardziej złożone konstrukty. Tego rodzaju metody pełnią jedynie funkcję środka wobec faktycznego celu badań socjologicznych, jakim jest wyjaśnianie zjawisk społecznych. Do omawianej grupy zaliczyłbym trzy rozdziały dotyczące różnych aspektów rzetelności i trafności pomiaru, cztery rozdziały na temat różnych metod analizy struktur ukrytych (*latent structure analysis*), a także dość sporą liczbę rozdziałów poświęconych metodom „wydobywania” ukrytych wymiarów z baterii pytań i wskaźników. W tym ostatnim przypadku chodzi o takie metody, jak skalowanie wielowymiarowe, analiza kanoniczna, analiza czynnikowa, analiza grupowania (*cluster analysis*) czy analiza dyskryminacyjna. Zaliczyłbym również, co może spotkać się z protestem niektórych socjologów, różne warianty modelowania równań strukturalnych, wieloczynnikową analizę wariancji, czy technikę znaną jako MTMM (*multitrait-multimethod*). Uważam jednak, że skoro istota tych metod opiera się na pomiarze wielowskaźnikowym, to model pomiaru decyduje o wizji badanego zjawiska.

Kolejna grupa, obejmująca 4 rozdziały, dotyczy badań eksperymentalnych. Zaliczyłbym do niej dwa rozdziały na temat projektowania schematów tego rodzaju badań oraz – co znów może wydać się decyzją kontrowersyjną – dwa rozdziały poświęcone analizie wariancji. Socjologowie korzystają z pewnych elementów tej techniki, na przykład porównując średnie w podgrupach badanej zbiorowości czy obliczając odsetek wariancji międzygrupowej (współczynnik η^2). Jednakże proponowane w książce ujęcie analizy wariancji ogranicza się do zastosowań tej metody w badaniach eksperymentalnych, przez co, dla socjologów, większość omawianych kwestii jest po prostu nieprzydatna.

Pozostałe 9 rozdziałów książki mogę z pełnym przekonaniem polecić socjologom. Podzieliłem je na dwie grupy, z których pierwsza obejmuje 3 rozdziały prezentujące kwestie ogólne, związane z prowadzeniem badań i interpretacją ich wyników, druga zaś 6 rozdziałów opisujących konkretne metody analizy danych.

Zacznijmy od rozdziałów prezentujących kwestie ogólne. W rozdziale 30 „Survey Sampling, Administration, and Analysis” Laura M. Stapleton zawarła kompendium wiedzy na temat zarządzania badaniami z uwzględnieniem standardów jakości. Mówi o właściwym zdefiniowaniu badanej populacji, o wymogach dotyczących budowy kwestionariusza, o wyborze schematu doboru próby i ocenie jakości operatu, z którego próba będzie dobierana, o kontroli prowadzenia wywiadów, o odsetku realizacji próby (*response rate*), o źródłach odmów, o ważeniu uzyskanych wyników oraz o zasadach estymacji wielkości błędów pomiaru. Badacze znajdą tu w pigułce wszystko, co stanowi przedmiot ustaleń komitetów metodologicznych międzykrajowych projektów badań porównawczych, takich jak Europejski Sondaż Społeczny czy PISA (Programme for International Student Assessment). Rzecz nie do pogardzenia w czasach, gdy ograniczenie się w notce pod tabelą do stwierdzenia, że liczebność próby wyniosła n , zaś próba była dobrana „losowo” – nie wystarcza już nawet w studenckich pracach semestralnych.

Z kolei drugi z grupy rozdziałów dotyczących kwestii ogólnych prezentuje dość rudymenarną refleksję dotyczącą stosowania metody naukowej (rozdział 7: „Effect Sizes and Confidence Intervals”). Jego autorzy, Geoff Cummings i Fiona Fidler, pracujący na uniwersytecie w Melbourne (jedyni cudzoziemcy w amerykańskim *teamie* pozostałych autorów), na standardy publikowania wyników badań próbują spojrzeć z perspektywy filozofii nauki. Jak słusznie zauważają, wśród współczesnych badaczy dominuje confirmacyjne podejście do opisu badanych faktów, czyli wyniki badania służą weryfikacji wcześniej sformułowanych hipotez. Problemem jest jednak to, że w wielu wypadkach podejście takie prowadzi do bezrefleksyjnego stosowania testu istotności. Egzemplifikacje zjawiska, o którym piszą Cummings i Fidler, łatwo odnaleźć otwierając na chybił trafił dowolny numer periodyku naukowego, nie wykluczając czasopism z najwyższej półki. Parametry opisujące siłę zależności, jak na przykład współczynniki regresji β , z reguły opatrzone są odpowiednią liczbą gwiazdek informującą o ich statystycznej istotności. Dobrze, ale owa istotność dotyczy hipotezy zerowej, która głosi, że wartość parametru w populacji jest równa 0. Mówiąc po ludzku, autor dowodzi jedynie tego, że w świetle uzyskanych wyników zależność zachodzi. Tymczasem rzeczywistym przedmiotem zainteresowań autora i czytelników jest siła owych zależności. Który z uwzględnionych czynników wywiera największy wpływ na badane zjawisko? Czy w miarę upływu czasu wpływ ten rośnie, czy maleje? Cummings i Fidler podpowiadają, że w statystyce problem ten rozwiązuje estymacja przedziałowa. Dla każdego szacowanego parametru siły zależności należy po prostu podać przedział ufności – na przykład 95-procentowy. Podobnie jak podaje się granice wyników w prognozach wyborczych. Tylko w ten sposób spełnić można jeden z głównych postulatów metody naukowej, jakim jest kumulatywność badań.

Kumulatywności badań dotyczy również trzeci z rozdziałów, które zaliczyłem do wykraczających poza opis konkretnej metody analizy danych. Jest to rozdział 19 na temat metaanalizy. Jego autorka, Natasha Beretvas, ogranicza zakres swoich rozważań do ilościowych aspektów tej metody, czyli, mówiąc w pewnym uproszczeniu, do zestawiania ze sobą ilościowych charakterystyk tego samego zjawiska otrzymanych w różnych badaniach. Autorka twierdzi, że wymogiem przeprowadzenia trafnych porównań jest rzetelny opis sposobu doboru próby i procedur realizacji każdego z zestawianych ze sobą badań. Albowiem istotą metaanalizy jest dążenie wyeliminowanie czynników – zwanych artefaktami – prowadzących do specyficznego „skrzywienia” wyników każdego badania. Skrzywienie to wynikać może z faktu przeprowadzenia badania w szczególnej podzbiorowości, na przykład w grupie studentów, z niewielkiej próby, z wysokiego odsetka odmów czy wreszcie ze sposobu formułowania zadawanych pytań. Szczególny problem w metaanalizie stanowi konieczność przyjęcia założeń dotyczących tych części populacji, o których niewiele wiadomo. Dlatego, że w podzbiorowościach tych badań w ogóle nie zrealizowano, bądź istniejące badania nie są dostatecznie wiarygodne. Dopiero pokonanie tych wszystkich szczebli daje rezultaty homogeniczne, które mogą być porównane ze sobą za pomocą narzędzi statystycznych.

Pozostała do omówienia grupa 6 rozdziałów opisujących metody analizy danych najczęściej stosowane przez socjologów. Tym razem nie będę wnikał w szczegóły

i ograniczę się do wymienienia tych metod. W rozdziale 5 znajdziemy omówienie korelacji i innych miar zależności między badanymi cechami. Rozdział 18 dotyczy modelowania log-liniowego, rozdział 21 metody regresji wielorakiej, zaś 17 regresji logistycznej. W rozdziale 10 omówiono modele wielopoziomowe, zaś w 31 metodę analizy zdarzeń, określaną w literaturze anglosaskiej jako *survival*, *event history*, *hazard*, *transition* bądź *duration analysis*. Ostatni z wymienionych rozdziałów jest autorstwa Paula D. Allisona. Jest to zarazem ostatni rozdział książki, przez co można z pewnego rodzaju satysfakcją odnotować, że ostatnie słowo należy do socjologa. Ciekawe, czy Allison celowo wybrał dla opisywanej metody określenie *Survival Analysis*? Gdyby przykładowo nazwał ją *Event History Analysis* – co na gruncie socjologii jest określeniem bodajże najczęściej stosowanym – to uplasowałby się ze swoją problematyką gdzieś w środku książki, między *Effect Sizes* a *Factor Analysis*. A tak, w jakimś stopniu wyróżnił naszą dyscyplinę.

Żarty, żartami, lecz wróćmy do zasadniczego wątku recenzji, czyli oceny przydatności książki Hancocka i Muellera dla socjologów. Chciałbym przekonać Czytelników, że istnieją przynajmniej dwa powody, dla których warto sięgnąć po tę książkę. Pierwszy jest oczywisty. Jeśli ktoś zainwestuje pieniądze w *reviewing process* w dobrym czasopiśmie, to nie po to, aby pieniądze te wyrzucić w błoto. Książkę uważam za najobszerniejszy i najbardziej aktualny zestaw zasad, jak pisać artykuł naukowy, skodyfikowanych przez autorów na co dzień zajmujących się oceną tego rodzaju publikacji. Wypadałoby jeszcze dodać, że naukowy dorobek wielu z nich wyznacza standardy obowiązujące w tym zakresie.

Drugi powód nie jest już tak oczywisty. Zacznijmy od stwierdzenia, że projekty badawcze realizowane są przede wszystkim tam, gdzie znajdują się na to publiczne środki. Dodajmy do tego fakt, że wśród strategicznych celów wielu współczesnych społeczeństw przewija się poprawa jakości edukacji. W rezultacie, jednym z bardziej wartościowych źródeł wiedzy na temat mechanizmów społecznego uwarstwienia stały się... wyniki badania PISA. Cztery dotychczas zrealizowane edycje objęły dziewięć lat obserwacji (2000–2009). Dane z pojedynczej edycji obejmują ponad pół miliona rekordów, które opisują związki między polityką edukacyjną a kształtowaniem się społecznych różnicowań w 67 krajach z całego świata. Ani się obejrzeliśmy, jak wyniki niszowego badania nabrały takiej rangi. Obecnie żaden szanujący się socjolog nie może ich zignorować.

Do analizy badań prowadzonych w ramach tradycyjnej, socjologicznej metodologii, takich jak World Values Survey czy International Social Survey Programme, w zasadzie wystarcza wiedza z 9, spośród wszystkich 31 rozdziałów książki. Jednakże świat badań ulega zmianom. Aby korzystać z lawinowo rosnących publicznych zasobów danych niezbędne staje się poszerzenie socjologicznego warsztatu o pojęcia i techniki, wcześniej stosowane wyłącznie w badaniach edukacyjnych, psychologicznych czy medycznych. Prędzej czy później staniemy wobec konieczności dowiedzenia się, co to jest alfa Cronbacha, czy jak liczyć tetrachoryczne korelacje. Przypuszczam, że w gabinetach wielu socjologów książka Hancocka i Muellera, rzucona gdzieś w kącie wraz z innymi, w pewnym momencie przewędruje na biurko. Coraz więcej będzie w niej fragmentów zaznaczonych kolorowym markerem.

Zaczniemy zaginać narożniki ważniejszych stron, aby móc ogarnąć, czym są współczesne nauki społeczne w wydaniu ilościowym. Czym jest semantyka owego języka, którym porozumiewamy się na konferencjach, który znajdujemy w opisach badań ściąganych z Internetu, czy który wyznacza kryteria oceny publikacji.

Przed rozpoczęciem prania warto dokładnie przeczytać instrukcję obsługi pralki. Po to, aby na przyjęciu epatować perfekcyjnie wypraną koszulą.

Literatura

- Babbie, Earl. 2008. *Podstawy badań społecznych*. Warszawa: WN PWN.
- Brzeziński, Jerzy. 2007. *Metodologia badań psychologicznych*. Warszawa: WN PWN.
- Sawiński, Zbigniew. 2010. *Kreowanie idei na podstawie wyników badań*. „Studia Socjologiczne” 2(197): 205–213.
- Silverman, David. 2009. *Prowadzenie badań jakościowych*. Warszawa: WN PWN.
- Treiman, Donald J. 2009. *Quantitative Data Analysis. Doing Social Research to Test Ideas*. San Francisco: Jossey-Bass.