


Zbigniew Sawiński 
Instytut Filozofii i Socjologii PAN



MISJA

SURVEY DATA HARMONIZATION IN THE SOCIAL SCIENCES **A POTRZEBA NOWEGO PARADYGMATU W BADANIACH** **PORÓWNAWCZYCH**

Artykuł rozpoczyna się od analizy historycznych przyczyn obecnego kryzysu w międzynarodowych badaniach porównawczych korzystających z harmonizacji *ex-ante* – podejścia, którego istota polega na powtarzaniu tych samych pytań w różnych krajach. Podejście stworzono w latach zimnej wojny w celu oceny efektów amerykańskiej propagandy, skąd na wiele lat trafiło do nauk społecznych, przyczyniając się do rozwoju międzynarodowych programów badawczych, jak WVS, ISSP, ESS czy PISA. Dopiero trudności w prowadzeniu badań w czasie pandemii wywołały utratę zaufania do harmonizacji *ex-ante*. Książka *Survey Data Harmonization in the Social Sciences* ukazała się w szczytowym momencie kryzysu. Choć autorzy nie kwestionują dokonania harmonizacji *ex-ante*, zdecydowanie zachęcają badaczy do szukania nowych podejść. W artykule zaproponowano, aby duże międzynarodowe programy badawcze obejmujące dziesiątki krajów uzupełnić mniejszymi projektami, które pozwalają poświęcić więcej uwagi specyfice zjawisk w badanych krajach. Historia pokazuje, że możliwość osiągnięcia lokalnych korzyści zawsze stanowiła jeden z głównych powodów zaangażowania krajów w międzynarodowe badania.

Słowa kluczowe: historia nauk społecznych; badania sondażowe; międzynarodowe programy badawcze; harmonizacja danych; pandemia

Mission

Survey Data Harmonization in the Social Sciences **and the Need for a New Paradigm in Comparative Research**

The article begins by examining the historical causes of the current crisis in cross-country surveys based on *ex-ante* harmonization, which involves asking the same questions across countries. Originally developed during the Cold War to assess the effects of U.S. propaganda, it was later adopted by the social sciences, contributing to international research programs such as WVS, ISSP, ESS, and PISA. Only recently have difficulties conducting surveys during the pandemic led to a loss of confidence in *ex-ante* harmonization. Although the authors of *Survey Data Harmonization in the Social Sciences*,

IFiS PAN, zbigniew.sawinski@ifispan.edu.pl, ORCID 0000-0003-1756-9853.

Źródło finansowania: NCN 2018/31/B/HS6/02043.

Tekst opublikowany na warunkach licencji Creative Commons Uznanie autorstwa-Użycie niekomercyjne-Bez utworów zależnych 3.0 Polska (CC BY-NC-ND 3.0 PL).

published at the height of the crisis, do not deny the merits of ex-ante harmonization, they strongly encourage researchers to seek new approaches. The article proposes that cross-country survey programs encompassing dozens of countries should give way to smaller projects that can devote more attention to country-specific issues. It is historically known that domestic benefits significantly increase an interest in cooperation between countries.

Key words: history of social sciences; survey research; cross-country survey programs; data harmonization; pandemic

*Im dalej popatrzyysz wstecz,
tym możesz dalej zobaczyć do przodu*
Winston Churchill

Wprowadzenie

Niekiedy droga prowadząca do naukowych celów okazuje się błędna i trzeba szukać nowych możliwości. Tak właśnie stało się z podejściem zwanym harmonizacją ex-ante, uważanym za metodologiczny fundament badań porównawczych prowadzonych z wykorzystaniem technik sondażowych. Istota harmonizacji ex-ante sprowadza się do zadania we wszystkich krajach tych samych pytań, przy jednoczesnym przeprowadzeniu badania w tym samym terminie i na próbach dobranych według identycznego schematu. Podejście pojawiło się pół wieku temu i od tego czasu zdominowało badania porównawcze. Przyczyniło się do powstania sztandarowych programów badawczych, jak Eurobarometr, World Value Survey (WVS), International Social Survey Programme (ISSP), Międzynarodowy Program Oceny Umiejętności Uczniów PISA czy Europejski Sondaż Społeczny (ESS).

Dopiero od niedawna zaczęto mówić o ograniczeniach harmonizacji ex-ante. Zakwestionowano jej teoretyczne podstawy, a w szczególności założenie, że zadawanie tych samych pytań pozwala badać zjawiska w różnych krajach w ekwiwalentny sposób. Ponadto, utrzymanie we wszystkich krajach jednolitych standardów uznano za nierealne. Kraje różnią się pod względem infrastruktury badawczej oraz gotowości respondentów do udziału w badaniach, co wpływa na jakość zbieranych danych.

Książka *Survey data harmonization in the social sciences* ukazała się w momencie, gdy budowany przez lata paradygmat harmonizacji ex-ante rozsypał się jak domek z kart. Jeszcze parę lat temu, zanim zaczęła się pandemia, głosy krytyczne pojawiały się sporadycznie. Trudności z prowadzeniem badań podczas pandemii spowodowały jednak, że rozczarowanie badaczy zaczęło rosnąć w szybkim tempie. W 2024 roku, w Europejskim Sondażu Społecznym – projekcie uznawanym za wzorcowy w dziedzinie porównań między krajami – podjęto decyzję o zastąpieniu wywiadów osobistych ankietami wypełnianymi

samodzielnie przez badanych. W pewnym uproszczeniu można powiedzieć, że obie metody mają się do siebie tak, jak pytanie zamieszczone w formularzu testu do pytania zadanego podczas grupy fokusowej. Tak zasadniczą zmianę w sposobie zbierania danych należy traktować jako początek nowego programu badawczego. A dotychczasowy ESS posłuży historycznej ocenie zmian w poglądach społeczeństw europejskich w pierwszych dwóch dekadach obecnego stulecia.

Książka ukazała się w takim momencie, że chcąc nie chcąc przypadła jej rola punktu zwrotnego w historii badań porównawczych. W 2024 roku kryzys osiągnął apogeum i zmiana obecnego paradygmatu stała się palącą kwestią. Co mają zrobić badacze, którzy jeszcze dwa lata temu wierzyli, że mimo przejściowych trudności standardy trzymają się mocno, natomiast dziś, jak to się kolokwialnie mówi, poczuli się wyrolowani? Pozostaje im szukać najbardziej aktualnego źródła wiedzy na temat harmonizacji *ex-ante*. Dlatego książka stanowi naturalny punkt wyjścia dla nowych propozycji.

Trzeba jednak mieć na uwadze, że książka nie przedstawia sposobu na wyjście z kryzysu. Aby tak się stało, musi upłynąć jeszcze wiele wody. Twórcy książki ograniczyli swoją propozycję do określenia przestrzeni dla takich działań, dokonując przeglądu badań porównawczych o różnorodnym charakterze i tematyce, prowadzonych przez ośrodki naukowe z całego świata, w których metody harmonizacji *ex-ante* nie wystarczały i wymagały uzupełnienia o inne rozwiązania. W ten sposób książka dała badaczom nadzieję, że nie są skazani na harmonizację *ex-ante*, a bogactwo dostępnych podejść stwarza szanse na wyjście z obecnej zapaści.

W artykule proponuje się spojrzeć na kryzys harmonizacji *ex-ante* z perspektywy historycznej, co pozwala docenić znaczenie dwóch kwestii. Po pierwsze, pojawienie się tego podejścia nie wynikało z logiki wcześniejszego rozwoju badań porównawczych. Wprowadzono je po drugiej wojnie światowej w celu monitorowania skuteczności amerykańskiej propagandy. Zastosowanie podejścia w naukach społecznych do porównań między krajami ma charakter wtórny i od początku budziło zastrzeżenia.

Po drugie, harmonizacja *ex-ante* nie sprzyja budowaniu więzi między krajami. Przestrzeganie przez kraje jednolitych standardów wymaga nadzoru. Rolę taką najczęściej pełni ponadnarodowa instytucja, jak Eurostat czy OECD, bądź przejmuje ją grupa badaczy reprezentujących kraje dysponujące większym potencjałem naukowym. Przejęcie zarządzania programem badawczym prowadzi do większego dostosowania jego założeń do realiów jednego czy grupy krajów, co z kolei skutkuje nierówną dystrybucją korzyści.

Autorzy książki nadzieję na przezwycięzenie kryzysu wiążą z rosnącymi możliwościami technicznymi integracji danych z różnych źródeł i coraz doskonalszymi algorytmami analizy dużych zasobów danych. Pozwoliłoby to na osadzenie projektów wykorzystujących harmonizację *ex-ante* w szerszych zasobach

wiedzy badawczej, dzięki czemu luki tego podejścia stałyby się mniej dokuczliwe.

W innym kierunku zmierza propozycja przedstawiona w artykule. Jeśli badania mają mieć sens dla uczestniczących w nich krajów, to tematykę i narzędzia badawcze należy silniej związać z lokalnymi możliwościami i potrzebami. Dlatego przyszłość badań porównawczych leży w projektach obejmujących niewielką liczbę krajów. Z jednej strony zwiększa to szanse na zachowanie krajowej specyfiki, zaś z drugiej nie zamyka drogi do poszukiwań uniwersalnych wzorców badanych zjawisk.

Geneza badań porównawczych

Korzyści z zestawiania ze sobą danych z różnych krajów zauważono w połowie XIX wieku, gdy socjologia stawiała się odrębną dyscypliną naukową. W ramach jednego z nurtów, rozwijanego głównie przez Adolphe'a Quételet, możliwości odkrycia praw społecznych wiązano z dostępnością danych na temat dużych i zróżnicowanych zbiorowości. Z tego punktu widzenia integrację danych z różnych krajów uznano za sposób na poszerzenie wiedzy budowanej na podstawie krajowych badań (Szacki, 1981; Stigler, 1990).

Aby zachęcić kraje do współpracy, Quételet podjął inicjatywę zorganizowania międzynarodowego kongresu w Brukseli w 1853 roku. Na nim za kwestię centralną uznano standaryzację, bez której nie wyobrażano sobie łączenia danych z różnych krajów. Standaryzacja miała objąć synchronizację badań w czasie, co dotyczyło spisów powszechnych, już wtedy prowadzonych przez wiele krajów, a także kwestię ujednolicenia stosowanych klasyfikacji, co między innymi dotyczyło edukacji czy gospodarki (Levi, 1854).

Jednak w obliczu rywalizacji między krajami zbyt daleko idąca współpraca nie była możliwa (Wojciuk, 2016). Badania dostarczały przesłanek dla podejmowanych działań i każdy kraj realizował je z myślą o własnych potrzebach, dostosowując do lokalnych uwarunkowań. Niewielkie były szanse na to, aby dość mglista możliwość poznania uniwersalnych praw skłoniła kraje do dostosowania swoich metod badawczych do tego, co robią inni, ponieważ mogłoby to oznaczać ograniczenie własnych korzyści. Z tych samych powodów z nieufnością podchodzono do kwestii udostępniania zebranych danych.

Dostrzeżono natomiast zasadność dzielenia się wyjaśnieniami dotyczącymi mechanizmów funkcjonowania wybranych zjawisk. Konfrontacja z tym, co dzieje się w innych krajach, ułatwić mogła przeprowadzenie właściwych reform. W latach trzydziestych XX wieku jedną z instytucji wyznaczających trendy w tej dziedzinie stało się Międzynarodowe Biuro Edukacji z siedzibą w Genewie. Biuro organizowało coroczne kongresy, na których krajowi przedstawiciele

dyskutowali o problemach edukacji, odwołując się do doświadczeń z poszczególnych krajów. Na tej podstawie sporządzano raporty zawierające sumaryczne wnioski. Nie poświęcano zbytnej uwagi tworzeniu jednolitych standardów czy wspólnych narzędzi badawczych. Nie uważano tego za potrzebne, skoro kraje poszukiwały przede wszystkim tego, co okazać się mogło przydatne w kontekście własnych celów (Rosselló, 1979).

Natomiast uznano, że warto przygotować opisy systemów szkolnych, aby w trakcie dyskusji nie tracić czasu na szczegóły. Zadanie to powierzono krajowym ekspertom, dzięki czemu systemów szkolnych nie wtłoczono w jednolitą klasyfikację. Przeciwnie – wiele krajów jako osobne kategorie wyodrębniło szkoły, które nie pojawiały się gdzie indziej. Na przykład Belgia wyodrębniła trzy kategorie szkół nauczycielskich: kształcące nauczycieli przedszkoli (normales gardiennes), szkół podstawowych (normales primaires) i szkół średnich (normales moyennes). Egipt jako odrębną uznał kategorię szkół koranicznych. Z kolei Polska wyodrębniła szkoły przysposabiające do zawodu rolnika (ecoles populaires de l'agriculture), uznając je za ważne ogniwo systemu oświaty w obliczu niedoborów sieci szkół powszechnych (Bureau International d'Éducation, 1935).

Standardy przywiezione w teczce

Kiedy w obliczu drugiej wojny światowej część badaczy i instytucji naukowych przeniosła się z Europy za Atlantyk, o kształcie badań zaczęły decydować interesy amerykańskie. W okresie zimnej wojny rząd USA przeznaczał znaczne środki na kształtowanie wizerunku Ameryki w oczach mieszkańców innych krajów. Do realizacji tego celu w 1953 roku powołano Agencję Informacyjną Stanów Zjednoczonych (USIA). Głównym zadaniem agencji, podporządkowanej bezpośrednio służbom rządowym, była koordynacja działań propagandowych w ponad stu krajach świata. Agencja finansowała między innymi wydawanie prasy i działalność stacji radiowych. Zdawano sobie sprawę, że skuteczność działań propagandowych zależy od ich dostosowania do oczekiwań i wyobrażeń mieszkańców poszczególnych krajów. Dlatego też wybór formy i treści przekazu powierzono redakcjom krajowym.

W celu oceny skuteczności działań podjętych w poszczególnych krajach USIA zdecydowała się prowadzić badania sondażowe. We wszystkich krajach postanowiono stosować ten sam zestaw pytań, co – jak należy przypuszczać – upraszczało organizację badania oraz poprawiało przejrzystość raportów przeznaczonych dla polityków. Zasadniczy cel badania stanowiło śledzenie zmian dokonujących się osobno w każdym kraju. Uznano, że aby cel ten zrealizować, w kolejnych edycjach wystarczy powtarzać te same pytania. W tym

kontekście nie miało znaczenia, czy pytania w różnych krajach rozumiano w ten sam sposób.

Stosowana przez USIA zasada zadawania tych samych pytań we wszystkich krajach została na początku lat siedemdziesiątych przyjęta za metodologiczną podstawę Eurobarometru (Reif i Inglehart, 1989). Na rzecz takiej decyzji przemawiał fakt, że w wielu europejskich krajach lokalne agencje badawcze wcześniej realizowały badania dla potrzeb USIA. Schemat ten zaspokajał potrzeby badania zmian osobno w każdym kraju. Natomiast trudno obecnie orzec, czy zebrane dane służyły też do porównywania krajów ze sobą – co bez zbadania sposobu rozumienia pytań nie byłoby uzasadnione.

Pierwszym projektem, w którym przyjęto, że zadanie tych samych pytań wystarcza dla dokonania porównań między krajami, było World Value Survey (WVS). Główny cel tego programu, rozpoczętego w 1981 roku, stanowiło porównanie systemów wartości społeczeństw z całego świata. Choć nie było do tego podstaw, przyjęto założenie, że staranne przetłumaczenie pytań na poszczególne języki zapewni ich jednolite rozumienie. Dziś trudno wskazać przyczyny tej decyzji, szczególnie że w tym czasie dostępne były też inne opcje. Jedna z ówczesnych propozycji głosiła, że zjawiska społeczne w każdym kraju stanowią złożenie komponentów wspólnych i specyficznych. Ograniczenie uwagi do komponentów wspólnych prowadzi do jednostronnego obrazu badanych zjawisk (Przeworski i Teune, 1966).

Pomimo braku podstaw teoretycznych, stosowanie tego samego zestawu pytań we wszystkich krajach szybko zdominowało badania porównawcze. Można sądzić, że atrakcyjność schematu wynikała nie z jego własności poznawczych, lecz z poczucia sprawczości i wspólnego działania, co mogło być atrakcyjne dla krajów rozwijających się. Na ogół na wejściu do nowego projektu nie formułowano wygórowanych wymagań, co pozwalało liczyć na równe traktowanie i uwzględnienie krajowego punktu widzenia podczas przygotowania wspólnych narzędzi.

Jednak nie wszystkie kraje dysponowały zasobami pozwalającymi wnieść równy wkład do wspólnego przedsięwzięcia. Zazwyczaj w ramach każdego projektu wyodrębniała się grupa badaczy reprezentujących kraje chętne do wzięcia na siebie ciężaru prac związanych z przygotowaniem i testowaniem narzędzi badawczych. W rezultacie niektóre kraje od samego początku zyskiwały uprzywilejowaną pozycję, ponieważ sposób podejścia do badanych zjawisk i zadawane pytania lepiej odzwierciedlały ich realia.

Chociaż rola harmonizacji ex-ante stale rosła, do końca lat dziewięćdziesiątych nie wyparła ona innych schematów badawczych. Na przykład w programie ISSP wspólne ustalenia ograniczały się do pytań modułów tematycznych, podczas gdy w przypadku pozostałych pytań pozostawiono swobodę, tak aby umożliwić krajowym zespołom na dołączenie modułu ISSP do innych badań.

W rezultacie część danych harmonizowano dopiero po przeprowadzeniu badania, co zaczęto nazywać harmonizacją *ex-post*. Przy czym podejście to uznano za mniej przydatne w porównaniach między krajami i starano się ograniczyć jego stosowanie do sytuacji koniecznych.

Pierwszym badaniem, które miało podlegać w całości zasadom harmonizacji *ex-ante*, był Europejski Sondaż Społeczny (ESS), uruchomiony w 2000 roku. Jednak zamiaru tego nie udało się zrealizować od początku. W pierwszych latach trwania projektu w niektórych krajach starano się zachować specyfikę wybranych zmiennych, uznając wspólne klasyfikacje za niespełniające krajowych potrzeb. Jeden z przykładów stanowiło wykształcenie. W jego przypadku zastosowanie zalecanej klasyfikacji w obrębie krajów nie pozwalało ujawnić faktycznego sposobu oddziaływania tej zmiennej na badane zjawiska, a w porównaniach między krajami prowadziło do kuriozalnych wniosków (Sawiński, 2013). Jednak pod koniec pierwszej dekady obligatoryjnie wprowadzono we wszystkich krajach jednolite zasady kodowania (Schneider, 2010), po czym do nowego schematu przekodowano dane zebrane w poprzednich rundach.

Niedostosowanie jednolitych narzędzi badawczych do krajowych realiów nie stanowi jedynej bolączki projektów korzystających z harmonizacji *ex-ante*. Trudności wywołuje też wymóg przeprowadzenia badania według tego samego schematu, co obejmuje zebranie danych w tym samym czasie. Nawet jeśli kraje uzgodnią wspólną datę rozpoczęcia badania, to nie oznacza, że wszędzie uda się zebrać na czas potrzebne środki. Może się też zdarzyć, że w niektórych krajach rozpoczęcie badania trzeba przełożyć ze względu na ważne wydarzenia, na przykład wybory. W krajach o słabszej infrastrukturze badawczej realizacja badania zwykle trwa dłużej, zwłaszcza gdy projekt wymaga osiągnięcia uzgodnionego odsetka realizacji próby. Opóźnienia występują też w fazie kodowania i przygotowania danych, ponieważ dane dostarczone przez krajowy zespół wymagać mogą poprawek. Wymienione czynniki powodują, że w niektórych programach badawczych czas oczekiwania na udostępnienie wyników wynosi aż półtora roku. Tak dzieje się między innymi w programach PISA i ESS.

Długi czas oczekiwania na upowszechnienie wyników przynosi straty przede wszystkim krajom dbającym o jakość infrastruktury badawczej. Na wyniki muszą czekać aż do momentu zebrania danych przez wszystkie kraje, również te, które do realizacji badania zbytnio się nie przykładają. Część krajów z przeprowadzeniem badania poradziłaby sobie dużo szybciej, gdyby nie opóźnienia wynikające z wymogów międzynarodowej współpracy. Zniechęca to kraje do udziału w badaniach porównawczych, a czasem prowadzi do przekierowania środków na realizację własnych badań, które można lepiej dostosować do potrzeb krajowych. Przykładem są Włochy, które doszły do takich wniosków po pierwszej rundzie ESS.

Jednak z niektórych programów międzynarodowych trudno się wycofać. Dlatego czasami szuka się sposobu, aby skorzystać z badania, zanim wyniki zostaną oficjalnie ogłoszone. W programie PISA do 2012 roku uczniowie wypełniali test w postaci papierowych zeszytów zadań, co pozwalało krajowym zespołom na analizę wyników bezpośrednio po ich zebraniu. Polski zespół PISA wielokrotnie korzystał z tej możliwości, co pozwalało na bieżąco monitorować efekty zmian w programach szkolnych. Jednak od 2015 roku papierowe zeszyty zastąpiono komputerowym testem, którego wyniki zapisywane są od razu na centralnym serwerze. Od tego czasu każdy kraj musi czekać około półtora roku, aby cokolwiek dowiedzieć się na temat uzyskanych wyników. Przykład ten ilustruje tezę, że na harmonizacji ex-ante tracą przede wszystkim kraje, którym zależy na wykorzystaniu wyników badania do własnych celów.

Początek końca

Choć wobec harmonizacji ex-ante wysunąć można wiele zastrzeżeń, przez lata uznawana była za obiecujący kierunek rozwoju badań porównawczych (Dubrow i Tomescu-Dubrow, 2016). Dopiero ostatnio, w ciągu dosłownie kilku lat, przekonanie o wiodącej roli harmonizacji ex-ante uległo zachwianiu. Najważniejsze okazały się dwa powody. Po pierwsze, w końcu do badaczy zaczęły dochodzić argumenty, że podejście nie spełnia wymogów teoretycznych. Po drugie, w obliczu pandemii spadło zaufanie do badań sondażowych, z którymi harmonizacja ex-ante związana jest od początku.

Prezentację argumentów krytycznych wobec harmonizacji ex-ante warto rozpocząć od kwestii teoretycznych. Podejście dostarcza użytecznych danych do porównań, jeśli spełnione są dwa warunki. Po pierwsze, we wszystkich porównywanych sytuacjach badany konstrukt można zdefiniować w równoważny sposób. Przykładowo, zaufanie do parlamentu można rozumieć w ten sam sposób w krajach o demokracji parlamentarnej i w krajach o ustroju autokratycznym pod warunkiem, że w każdym z tych krajów parlament pozostaje niezależny od rządu. Jeśli jednak w kraju o ustroju autokratycznym rola parlamentu sprowadza się do legitymizacji decyzji rządu, to w takim przypadku trudno mówić o zaufaniu. Przynajmniej nie w znaczeniu, w jakim jest ono rozumiane w systemach politycznych, w których parlament podejmuje decyzje niezależnie.

Drugi postulat teoretyczny głosi, że narzędzie badawcze we wszystkich porównywanych sytuacjach powinno mierzyć to samo, co określa się mianem niezmienniczości pomiaru (measurement invariance). Istotę tego postulatu można wyjaśnić następująco. Załóżmy, że zadano pytanie o ogólną ocenę rządu, a następnie serię pytań dotyczących oceny działań rządu w różnych obszarach, takich jak gospodarka, edukacja czy służba zdrowia. Należałoby oczekiwać, że

ogólna ocena rządu okaże się wypadkową ocen szczegółowych, silniej korelując z ocenami w obszarach uznanych za ważniejsze. Co jednak myśleć o kraju, w którym korelacje z ocenami szczegółowymi okażą się statystycznie nieistotne, natomiast wskaźnik ogólnej oceny rządu będzie silnie skorelowany ze sposobem postrzegania premiera: czy jest przystojny, prawdomówny, czy wierzy w Boga itp. Przy porównywaniu obu krajów postulat niezmienności pomiaru trudno uznać za spełniony.

Podczas gdy ocena spełnienia pierwszego postulatu wymaga odwołania się do teorii, spełnienie postulatu niezmienności pomiaru można ustalić na podstawie wyników badań. Jeśli pomiar dotyczy tego samego konstrukt w dwóch różnych sytuacjach, struktura ładunków czynnikowych łączących konstrukt z wynikami pomiaru powinna być taka sama. Choć sama metoda znana jest od dawna (Van de Vijver i Tanzer, 2004), to dopiero w ostatnich latach zaczęto ją stosować w międzynarodowych programach badawczych do testowania zasadności przyjętych założeń (Koc i Pokropek, 2022; Meuleman i in., 2022).

Przeprowadzone dotychczas testy nie pozostawiają wątpliwości, że w wypadku niektórych konstruktów postulat niezmienności pomiaru nie może zostać uznany za spełniony nie tylko dla porównań między krajami, ale także dla porównań w czasie. O ile różnic w sposobie rozumienia pytań w różnych krajach można było się spodziewać, uwzględniając bogactwo języków i kultur, chyba mniej spodziewano się stwierdzenia zmian w sposobie rozumienia pytań zadawanych w kolejnych edycjach tego samego programu badawczego. Dotychczas badacze uważali, że powtarzanie tych samych pytań wystarczy, aby badać zmiany. Tymczasem zmiany w postrzeganiu niektórych zjawisk okazują się na tyle szybkie, że wymagają odpowiednich modyfikacji treści pytań.

Zagadnienia spełnienia postulatów teoretycznych nie przyciągają jednak na tyle uwagi badaczy, aby w istotny sposób wpłynąć na obraz badań porównawczych. Dlatego za główną przyczynę spadku zaufania do harmonizacji *ex-ante* należy uznać pogarszającą się jakość badań sondażowych. Trudno zapewnić zgodność ze standardami, gdy rośnie niechęć do udziału w badaniach. Ponadto, respondenci stają się coraz bardziej świadomi swojej podmiotowości. Nie czują się mniej kompetentni od badacza i coraz częściej decydują, na które pytania odpowiedzą, a które uznają za niepotrzebne. Czasami nawet godziwe gratyfikacje nie wystarczą, aby przekonać wylosowane osoby do wzięcia udziału w badaniu i zastosowania się do procedury. Dlatego od pewnego czasu wśród badaczy narastało przekonanie, że misterna konstrukcja coraz bardziej wyśrubowanych standardów musi w pewnym momencie się zawalić.

Do ostatecznego upadku harmonizacji *ex-ante* przyczyniła się pandemia. Zamiast ankietów z laptopami, zaczęto wysyłać respondentom papierowe ankiety. Rozwiązanie to przyjęto między innymi w dziesiątej rundzie ESS – przeprowadzonej w 2020 i 2021 roku. Ta z pozoru niewielka zmiana wywołała

katastrofę. Badani poruszali się po papierowych ankietach według własnej logiki, ignorując reguły przejść, zamazując lub zmieniając wcześniejsze odpowiedzi. Wiele pytań zostawiali pustych bez żadnego komentarza. W rezultacie część wypełnionych ankiet nie nadawała się do połączenia z ankietami webowymi czy wywiadami CAPI. Aby ratować projekt, międzynarodowy zespół ESS postanowił przekonać krajowe zespoły, aby „dopasowały” odpowiedzi z papierowych ankiet do wymogów skryptu kierującego przebiegiem wywiadu w wersji komputerowej. Wyjaśniono, że bez spełnienia tego warunku nie będzie można dokonać harmonizacji danych z poprzednimi rundami ESS. W ten sposób krajowe zespoły postawiono pod ścianą. Bo jak rozliczyć środki na badanie, jeśli międzynarodowe konsorcjum nie przyjmie zbioru danych?

Strategia

Książka stanowi czwartą kolejną pozycję wydawnictwa Wiley poświęconą zastosowaniu metod sondażowych w badaniach porównawczych. Pierwsza ukazała się w 2003 roku i uznać ją można za próbę uporządkowania zasad harmonizacji *ex-ante* na podstawie doświadczeń zebranych w projektach WVS i ISSP (Harkness i in., 2003). W drugiej książce, z 2010 roku, wzbogacono podejście o nowe koncepcje, w tym Total Survey Error, zarazem uwzględniając projekty rozpoczęte w międzyczasie, w tym ESS (Harkness i in., 2010). W trzeciej, wydanej w 2019 roku, dokonano wyłomu, decydując się poświęcić część książki metodom harmonizacji *ex-post* (Johnson i in., 2019). I wreszcie najnowsza książka z 2024 roku. Tutaj proporcje odwróciły się i harmonizacja *ex-ante* z głównego bohatera stała się statystą.

Najnowsza książka powstała w szczególnych okolicznościach. W nauce nie kwestionuje się wcześniejszego dorobku, gdyż brak kumulacji wprowadza chaos. Zamiast tego szuka się luk w aktualnej wiedzy, co pozwala wskazać kierunki dalszych badań. Co jednak zrobić, gdy cała gałąź badań naukowych od lat nieubłaganie zbliża się do krawędzi przepaści? Można cofnąć się do momentu zejścia z właściwej drogi, oddzielić późniejsze ustalenia grubą kreską i od nowa rozpocząć przynajmniej niektóre sprawy.

Jednak trudno byłoby liczyć na poparcie środowiska naukowego dla tak radykalnej propozycji. Międzynarodowe programy badawcze wciąż mają liczną grupę zwolenników, przez których harmonizacja *ex-ante* bywa traktowana jako rodzaj świętości. Przedstawienie opinii, że programy te wymagają głębokich reform, byłoby równoznaczne z naukowym seppuku. Co więcej, byłoby wyjątkowo nieprzyjemną rolą podjąć się przekazania środowisku naukowemu faktu, że ponadpięćdziesięcioletni dorobek badaczy z całego świata trafi w najlepszym wypadku do archiwów.

Aby uniknąć dyskusji na temat ograniczeń harmonizacji *ex-ante*, twórcy książki potraktowali swoją misję jako forpocztę reform potrzebnych w badaniach porównawczych. Książka nie oferuje gotowych rozwiązań, lecz stara się zachęcić badaczy do samodzielnych poszukiwań. Prezentuje podejścia wykraczające poza kanony harmonizacji *ex-ante*, pozostawiając ich twórcom miejsce na przedstawienie osiągniętych korzyści i napotkanych ograniczeń. Dzięki temu w jednym miejscu skupiono doświadczenia zebrane po obu stronach Atlantyku, w badaniach przekrojowych i podłużnych, w projektach międzynarodowych i krajowych czy w badaniach nastawionych na różne cele, jak sondaże opinii i spisy powszechnie. Prezentację tak szerokiego obszaru doświadczeń uzupełniają rozdziały poświęcone metodom analitycznym zaprojektowanym pod kątem badań porównawczych.

State-of-the-Art

Ocena stanu wiedzy w danej dziedzinie stanowi integralny składnik każdej propozycji naukowej. Jednak pisanie o ograniczeniach harmonizacji *ex-ante*, czego w książce nie sposób uniknąć, nie przekłada się ani na pozytywny wizerunek piszącego, ani na przyjazny stosunek do poruszanych zagadnień. Być może z tego względu twórcy książki postanowili oddać głos autorom rozdziałów. Trudno o bardziej przekonujące argumenty niż własne, negatywne doświadczenie.

W tym duchu o harmonizacji *ex-ante* piszą autorzy rozdziału poświęconego trudnościom we wdrażaniu międzynarodowych standardów w krajach Azji Wschodniej (Iwai i in., 2024). Badacze z krajów tego regionu zauważają, że z biegiem czasu sposób rozumienia przez respondentów niektórych zjawisk ulega zmianom i zmiany te dokonują się w niejednakowym tempie w różnych krajach. Za główną trudność międzynarodowych badań uznają brak języka, jaki pozwoliłby stworzyć wzorzec wspólnego kwestionariusza. Język angielski jest na tyle obcy, że wiele jego terminów nie ma odpowiedników w językach tego regionu. Do tego dochodzą różnice w unikaniu skrajnych odpowiedzi, czym respondenci z poszczególnych krajów różnią się znacząco. Pojawiają się też trudności innego typu. Niektóre międzynarodowe narzędzia badawcze objęte są prawem autorskim, co wyznacza sposób przetłumaczenia pytań w krajach objętych licencją, podczas gdy pozostałe kraje mają pod tym względem swobodę. Zdarza się też, że tłumaczenie wzorcowego kwestionariusza w pewnym momencie zostaje uznane za niewłaściwe z powodu błędów popełnionych wiele lat wcześniej. Wtedy pojawia się dylemat, czy pozostać przy poprzedniej wersji, czy zastąpić tłumaczenie właściwym. Wniosek z doświadczeń krajów Azji Wschodniej brzmi jednoznacznie. Nie wszystkie trudności związane z harmonizacją *ex-ante* dadzą się przewyciężyć.

Inne spojrzenie na rolę harmonizacji *ex-ante* przedstawia rozdział na temat programu WVS (Kizilova i in., 2024). Autorzy prezentują pogląd, że sukces tego programu osiągnięto właśnie dzięki zastosowaniu harmonizacji *ex-ante*, co nie sprzyja mówieniu o jej ograniczeniach. Tymczasem niektóre z tych ograniczeń badaczom są doskonale znane. Do przykładów należy pytanie o dawanie łapówek, które jako wskaźnik korupcji nie sprawdza się w krajach, gdzie wręczenie prezentu w biurze lub w urzędzie uważa się za oczywiste. Dla badaczy dawno stało się jasne, że zadanie tego pytania ma sens jedynie w niektórych krajach. W tym wypadku autorzy rozdziału znaleźli się w trudnej sytuacji. Przyznanie, że pytanie o łapówki wskazuje na ograniczenia harmonizacji *ex-ante*, wymagałoby napisania w kolejnym zdaniu, że program WVS ma te same ograniczenia. Zaburzyłoby to narrację, toteż autorzy rozdziału postanowili przekonać czytelników, że w ramach harmonizacji *ex-ante* można sobie poradzić z tego rodzaju trudnościami. Wyjaśniają w związku z tym, że „[...] WVS planuje zwiększyć zaangażowanie naukowców z regionów dotychczas niedoreprezentowanych i zintensyfikować współpracę [...] w fazie projektowania kwestionariuszy kolejnej rundy badania w latach 2023–2026” (Kizilova i in., 2024, s. 51).

Najbardziej radykalne podejście do harmonizacji *ex-ante* znaleźć można w programie Afrobarometr (Logan i in., 2024). Wśród uczestników programu występuje przekonanie, że podstawową korzyścią z udziału w międzynarodowych badaniach jest promowanie wizerunku własnego kraju i jego otwartości na współpracę. Na przestrzeganie standardów badawczych patrzy się przez palce. Dla krajów tego obszaru badania mają przede wszystkim wymiar praktyczny i służą rozpoznaniu lokalnych potrzeb. W obliczu różnic kulturowych oraz w warunkach niejednakowego rozwoju infrastruktury badawczej realizacja wspólnych ustaleń wymagałoby dodatkowych nakładów, z czego korzyści byłyby niewielkie.

Briefing

Kluczową część książki przyrównać można do spotkania przy ognisku. Głos oddano badaczom, którzy przedstawili rozwiązania pozwalające wyjść poza rutynowe ramy harmonizacji *ex-ante*. Formuła spotkania przy ognisku pozostawia uczestnikom swobodę co do wyboru prezentowanych zagadnień, a przy tym nie wymaga, aby różne propozycje układały się w spójny obraz.

Zainicjowane przez twórców książki spotkanie stworzyło pewnego rodzaju mapę rozwiązań wypracowanych i sprawdzonych w badaniach porównawczych. Mapę, na której nie ma podziału na rozwiązania bardziej i mniej przydatne, ani wskazówek, jaką dalszą drogę powinny wybrać badania porównawcze.

Nie można wykluczyć, że wśród przyszłych rozwiązań nie będzie miejsca na obecne programy badawcze, co stanie się wówczas, gdy nowe propozycje stanowiąc będą złożenie elementów zaczerpniętych z różnych programów. Dlatego prezentując dorobek badań porównawczych warto to zrobić jak najszerzej. Rozwiązania dziś postrzegane jako niszowe jutro mogą zostać uznane za najbardziej obiecujące.

Do oceny użyteczności przedstawionych propozycji posłużyć się można przykładowymi problemami, które jak się wydaje, w najbliższym czasie wymagać będą rozwiązania. Pierwszy, z dwóch przytoczonych w tym miejscu przykładów, dotyczy sytuacji, gdy informacji na temat badanego zjawiska dostarcza więcej niż jedno badanie, przez co zachodzi konieczność dokonania wyboru: które z badań wykorzystać, a które pominąć. Spośród wystąpień przy ognisku podobny problem wystąpił w projekcie Survey Data Recycling (Tomescu-Dubrow i in., 2024). Projekt ma na celu stworzenie uniwersalnej platformy harmonizacji badań sondażowych, co podzielono na dwa etapy. Pierwszy polega na zgromadzeniu w jednym miejscu danych z badań przeprowadzonych w różnym czasie i w różnych krajach. Cel drugiego zaś stanowi ustalenie zasad przekształcania oryginalnych pytań i odpowiedzi w zmienne docelowe. W projekcie SDR dodatkowo przewidziano, aby różnice w sformułowaniach pytań, bądź inne charakterystyki badań, mogące zaburzyć porównania, oflagować za pomocą zmiennych kontrolnych, co pozwala na ich eliminację podczas analizy.

Do czasu opublikowania książki w projekcie SDR zebrano dane z 3329 badań sondażowych, przeprowadzonych w latach 1966–2017 (Tomescu-Dubrow i in., 2024, s. 147). Obejmuje to najbardziej znaczące międzynarodowe programy badawcze, uzupełnione o wybrane badania krajowe. Przy tak bogatej bazie danych zdarzają się sytuacje, że informacje na określony temat zbierano w więcej niż jednym badaniu. W takich sytuacjach badacz musi zdecydować: czy z dostępnych badań wybrać jedno, czy też posłużyć się więcej niż jednym badaniem, łącząc wyniki według ustalonego klucza.

W projekcie SDR proponuje się, aby dokonując wyboru wziąć pod uwagę jakość badań. Wskaźniki jakości dotyczyć mogą fazy projektowania badania, na przykład tego, czy badanie główne zostało poprzedzone badaniem pilotażowym, oceny sposobu jego przeprowadzenia, chociażby w postaci odsetka realizacji próby, oceny kompletności dokumentacji badania, na przykład tego, czy dostępny jest codebook opisujący zmienne i ich wartości, czy wreszcie oceny zebranych danych, na przykład pod kątem częstości pojawiania się braków odpowiedzi. Propozycja opiera się na założeniu, że badania oceniane wyżej pod względem jakości dostarczają bardziej wiarygodnych danych.

Skupienie uwagi na wskaźnikach jakości badań ma też ograniczenia. Kryteria jakości mogą się wykluczać, co wymaga wyboru między nimi. W programie ESS

jeden z priorytetów stanowi dotarcie do jak największej liczby respondentów, co przekłada się na stosunkowo wysokie wskaźniki realizacji próby, lecz odbywa się kosztem wydłużenia przeciętnego czasu realizacji badania. W Eurobarometrze, gdzie badacze przywiązują większą wagę do uzyskania obrazu społeczeństwa w określonym momencie, wywiady realizuje się przez okres trzech tygodni. W tak krótkim czasie nie udaje się dotrzeć do wszystkich wylosowanych osób, co prowadzi do niższych wskaźników realizacji próby (Jabkowski, 2021).

Na problem wyboru badań można spojrzeć nie tylko od strony ich jakości, lecz również mając na uwadze własności danych pochodzących od respondentów. Podejście takie również zaprezentowano podczas spotkania (Durand, 2024). W tym przypadku pierwszym krokiem jest zawężenie danych do zmiennych bezpośrednio związanych z badanym zagadnieniem, co ułatwia przedstawienie wyników wszystkich uwzględnionych badań w jednolitym formacie i umieszczenie ich w tej samej bazie danych. To z kolei pozwala na wydzielenie trzech głównych źródeł zróżnicowania wyników: ze względu na kraj, czas realizacji badania oraz projekt, do którego należy badanie. Często wystarcza to do podjęcia decyzji, które z rozpatrywanych badań w satysfakcjonujący sposób odzwierciedlają badane zjawisko. Jeśli wątpliwości nadal pozostają, to wtedy skorzystać można z kryteriów zaproponowanych w wystąpieniu na temat ekwiwalentności wyników różnych badań (Pokropek, 2024). Nic też nie stoi na przeszkodzie, aby wnioski z analizy danych pochodzących od respondentów połączyć z oceną jakości badań, tak jak proponuje się w projekcie SDR.

Drugim przykładem zagadnienia ilustrującego potencjał zawarty w propozycjach przedstawionych na spotkaniu jest kwestia harmonizacji krajowych badań panelowych. Badania te, zapoczątkowane w latach sześćdziesiątych, służą ocenie skuteczności reform i działań podejmowanych w ramach polityki publicznej. Na ogół są to duże projekty, obejmujące od kilkunastu do kilkudziesięciu tysięcy osób, do których wraca się wielokrotnie w ustalonych odstępach czasu. Ze względu na związek badanych zagadnień ze specyfiką każdego kraju, harmonizacja ex-ante nie znajduje tu zastosowania. W zamian proponuje się wykorzystanie metod harmonizacji ex-post, uznawanych za bardziej adekwatne w wypadku badań, których nie projektowano z myślą o porównaniach.

Podczas spotkania dokonano prezentacji unikatowej inicjatywy w dziedzinie harmonizacji krajowych badań panelowych, projektu o nazwie Cross-National Equivalent File (CNEF). Projekt obejmuje jedynie badania, w których wizyty u badanych odbywają się corocznie. Ograniczenie projektu do badań o tym samym cyklu realizacji zmniejsza liczbę potencjalnych trudności z porównaniem ich wyników. Obecnie w projekcie bierze udział dwanaście krajów, a chęć przyłączenia się wyraziło kolejnych sześć (Lillard, 2024).

Projekt CNEF otwiera nowe perspektywy dla porównań między krajami, ale zawiera też wyzwania. Badania panelowe o rocznym cyklu realizacji

należą do najbardziej kosztownych przedsięwzięć badawczych. Stwarza to barierę wejścia, gdyż wielu krajów na rozpoczęcie takiego badania po prostu nie stać. Przykład stanowi Polskie Badanie Panelowe POLPAN, realizowane od 1988 roku, w którym ograniczone środki pozwalają na wizyty u badanych nie częściej niż co pięć lat (Tomescu-Dubrow i in., 2021). Z tego powodu POLPAN nie może zostać bezpośrednio włączony do projektu CNEF. Choć z pewnością przyniosłoby to pożytek dla nauki, gdyż POLPAN dostarcza unikatowej wiedzy na temat przemian towarzyszących przejściu od gospodarki centralnie sterowanej do gospodarki rynkowej w Europie Środkowo-Wschodniej.

Jednym ze sposobów przezwyciężenia tych trudności jest przekształcenie wyników krajowych badań panelowych na format roczny. Prezentacje wygłoszone tego wieczoru dostarczyć mogą wskazówek w tym zakresie. W badaniu National Longitudinal Survey of Youth (NLSY) musiano zmierzyć się z podobnym problemem, gdyż w pewnym momencie zamiast co roku, badanych zaczęto odwiedzać co dwa lub co cztery lata (Cooksey i in., 2024). Przydatne okazać się również mogą doświadczenia zebrane w Wielkiej Brytanii podczas harmonizacji badań podłużnych (O'Neill i Hardy, 2024), czy pochodzące z Niemiec, gdzie łączono dane biograficzne partnerów zbierane w różnych badaniach (Weiß i in., 2024). Bardziej ogólnych wskazówek dotyczących harmonizacji badań panelowych dostarczyć może projekt dotyczący harmonizacji badań podłużnych z jedenastu krajów (Wey i Fortier, 2024). W sumie, próbując stworzyć rozwiązanie problemu dotyczącego harmonizacji badań, warto uwzględnić doświadczenia zebrane w różnych projektach. Choć trudno liczyć na znalezienie gotowego rozwiązania, bogactwo doświadczeń i pomysłów sprawi, że łatwiej będzie znaleźć właściwą drogę.

Ognisko już dogasało, gdy twórcy książki dokonali podsumowania. Spotkanie stworzyło okazję do refleksji nad harmonizacją *ex-ante*. Podejście przyczyniło się do zgromadzenia ogromnych zasobów danych, stanowiących unikatowe źródło wiedzy na temat przemian w krajach z całego świata. Obecnie pojawiają się szanse, aby potencjał ten nadal mógł służyć kolejnym pokoleniom badaczy. Postęp technologiczny systematycznie obniża koszty integracji danych z różnych źródeł. Ułatwia łączenie wyników badań sondażowych z danymi z publicznych rejestrów czy z zasobami Internetu. Wiele wskazuje na to, że w najbliższym czasie intensyfikacji ulegnie interdyscyplinarna współpraca. Wszystko to sprzyjać będzie osadzeniu harmonizacji *ex-ante* w szerokich zasobach wiedzy badawczej i pomoże przezwyciężyć niektóre z jej deficytów.

Testimonial

Gdy pierwsze promienie wschodzącego słońca rozświetliły ciemności nocy, przy ognisku pozostał już tylko jeden człowiek. Przychodził w to miejsce od lat i widział już niejedną misję. Zwykle kończyły się niepowodzeniem – poza jedną, która w 2011 roku przywiodła na polanę badaczy z programu PISA.

Losy harmonizacji ex-ante oraz programu PISA mają sporo wspólnego. Geneza obu programów sięga okresu zimnej wojny, gdy stanowiły odpowiedź na wyzwania, z którymi musiały zmierzyć się amerykańskie władze. Za początek ciągu zdarzeń, jaki doprowadził do rozpoczęcia prac nad programem PISA, uważa się szok, jakiego doświadczyło amerykańskie społeczeństwo po wystrzeleniu Sputnika przez Związek Sowiecki w 1957 roku. Za powód porażki uznano niedostateczny poziom amerykańskiego szkolnictwa. Szczególnie niepokojące było to, że badania nad edukacją, prowadzone przez UNESCO w większości krajów świata, w tym w ZSRS, nie dały żadnego wcześniejszego ostrzeżenia (Heyneman, 1999). W tej sytuacji amerykański rząd postanowił stworzyć własny program badawczy. Miał on dostarczyć realnych ocen potencjału systemów edukacji z różnych krajów, a ponadto wiedzy na temat skutecznych rozwiązań, jakie można byłoby wdrożyć w amerykańskich szkołach.

W 1961 roku powołano w tym celu Organizację Współpracy Gospodarczej i Rozwoju OECD (Tröhler, 2013). Początkowo próbowano skorzystać z badań prowadzonych przez International Association for the Evaluation of Educational Achievement (IEA) – organizację, która jako pierwsza do oceny edukacji zaczęła stosować testy osiągnięć. Okazało się jednak, że IEA nie prowadzi porównań w systematyczny sposób. Organizacja zrzeszała badaczy akademickich, którzy uczestnictwo w kolejnych edycjach prowadzonych przez IEA programów badawczych uzależniali od swoich własnych naukowych celów. Pod koniec lat osiemdziesiątych dla OECD stało się jasne, że konieczne będzie stworzenie własnego programu. Prawie dziesięć lat zajęło przygotowanie badania PISA, którego realizację rozpoczęto w 2000 roku. W projektowaniu badania posłużono się zasadami harmonizacji ex-ante. Dotyczy to zarówno testu umiejętności – zawierającego te same zadania we wszystkich krajach, jak również ankiet wypełnianych przez uczniów, rodziców i dyrektorów szkół.

Od początku istnienia programu PISA wiadomo było, że ocena jego efektów zależy od zdolności dostarczenia rozwiązań potrzebnych amerykańskiej edukacji. Z tego powodu prawdziwym testem skuteczności programu okazał się raport: „Lessons from PISA for the United States” (OECD, 2011). Raport w nieoczekiwany sposób wykazał, że gdy potrzebne są naprawdę użyteczne wskazówki, to analiza krajowego kontekstu staje się ważniejsza od wniosków z porównania wyników badań. Dane PISA, obejmujące tysiące zmiennych w szczegółach opisujących stan edukacji w kilkudziesięciu krajach świata, wykorzystano do

wytypowania ośmiu krajów osiągających w programie bardzo dobre wyniki. Zasadniczą część raportu zajęły szczegółowe opisy funkcjonowania szkolnictwa w każdym z wybranych krajów, z zamiarem wydobycia mechanizmów sprzyjających sukcesom edukacji. Jednak najbardziej znamieny wydaje się fakt, że zadania tego nie powierzono badaczom z zespołu PISA. Za bardziej kompetentnych uznano krajowych ekspertów od oświaty.

Niesamowite, pomyślał mężczyzna, grzebiąc kijem w wygasłym ognisku. Lekcje dla USA nie pozostawiły wątpliwości, że wyniki badań niewiele znaczą bez odniesienia do krajowej rzeczywistości. Tymczasem badacze wciąż wierzą, że dla zrozumienia istoty zjawisk wystarczą wyniki badań.

Debriefing

Aż nadszedł dzień, w którym uczestnicy misji, przedzierając się przez gęstwinę badań porównawczych, znaleźli drogę prowadzącą wprost do naukowych celów. Na końcu drogi czekały dwie osoby: kobieta i mężczyzna. Miło was widzieć – kobieta powitała przybyłych. Chciałabym podziękować za książkę. Pomogła rozwiązać wiele mitów na temat badań porównawczych.

Coraz więcej badaczy zyskuje przekonanie – kontynuowała kobieta – że w porównaniach między krajami kończą się czasy dominacji jednego podejścia. Dlatego w waszej książce doceniam otwartość na różne rozwiązania. Szerokie możliwości wyboru ułatwiają osadzenie badań w krajowych realiach, zarówno w wymiarze badanych zjawisk, jak i pod względem doboru metod badawczych. Daje to krajom większą szansę na osiągnięcie własnych korzyści i sprzyja poszukiwaniu uniwersalnych wyjaśnień.

Jesteśmy wdzięczni za słowa potwierdzające sens naszej misji – odpowiedział jeden z badaczy. Rozpoczynając ją, wierzyliśmy, że zintegrowanie międzynarodowych programów badawczych z wiedzą pochodzącą z innych źródeł pomoże uwolnić je od ograniczeń harmonizacji ex-ante i zachować ich dorobek dla nauk społecznych. Chcielibyśmy poznać waszą opinię na ten temat.

Mężczyzna i kobieta wymienili spojrzenia. Harmonizacja ex-ante – zaczął mężczyzna – nie stanowi dogodnego punktu wyjścia dla budowania przyszłości badań porównawczych. Zastosowanie we wszystkich krajach tych samych narzędzi badawczych powoduje, że badania są nastawione na wydobycie tego, co wspólne. Poza zakresem pozostawia się to, co dla krajów specyficzne, a bez czego trudno wyjaśnić, dlaczego kraje różnią się w wymiarze badanych zjawisk.

Dlatego warto przerwać to błędne koło. Międzynarodowe programy badawcze nie potrzebują doskonalszych metod harmonizacji. Potrzebują czegoś, co nazwać można deharmonizacją. Wymagają prześledzenia zależności wyników

od sposobu, w jaki badane zjawiska osadzone są w realiach każdego kraju. Potrzebują też uwzględnienia stanu infrastruktury badawczej oraz podejścia respondentów do badań, gdyż w tym zakresie różnice między krajami bywają znaczne i łatwo mogą wypaczyć sens porównań (Kyyrä i Wilke, 2014).

Wiele wskazuje na to – dodała kobieta – że lukę po dużych, międzynarodowych programach badawczych wypełnią projekty angażujące niewielką liczbę krajów, ściśle współpracujących przy wąsko określonym zagadnieniu (Kopycka i in., 2023). Historia uczy, że główny powód, dla którego kraje angażują się w międzynarodową współpracę, stanowi potrzeba uzyskania wewnętrznych korzyści. Korzyści takie możliwe są tylko wtedy, gdy kraje godzą się współpracować w sposób otwarty, z poszanowaniem swoich racji. Tylko wtedy badania porównawcze są w stanie osiągnąć naukowe cele, odkrywając to, co wspólne w zjawiskach, które w swojej istocie są wyjątkowe.

Podziękowania

Chciałbym podziękować dwóm anonimowym recenzentom za uwagi i pomocne sugestie dotyczące pierwszej wersji artykułu. Jestem niezwykle wdzięczny Kazimierzowi Maciejowi Słomczyńskiemu, którego wiedza i życzliwość pomogła mi uporządkować poglądy w wielu poruszanych kwestiach. Na kształt przedstawionych w artykule propozycji wpływ miały doświadczenia zebrane w projekcie „Dynamika niepewności zatrudnienia młodych: uwarunkowania, trajektorie i skutki w perspektywie porównawczej” (NCN 2018/31/B/HS6/02043). Realizacja tego projektu zbiegła się z okresem pandemii oraz kryzysu w badaniach porównawczych, co zmusiło do szukania nowych podejść. W tym zakresie wiele propozycji przygotowaliśmy wspólnie z Anną Kiersztyn i Katarzyną Kopycką.

Na koniec chciałbym przeprosić czytelników, których razi metaforyczna konwencja użyta w części artykułu. Jedynym celem wprowadzenia fikcyjnych osób i sytuacji było uwypuklenie kwestii uznanych za kluczowe.

Bibliografia

- Bureau International d'Éducation. (1935). *Annuaire international de l'éducation et de l'enseignement 1935*. Bureau International d'Éducation.
- Cooksey, E., Gardecki, R., Lunney, C., Roose A. (2024). Harmonization in the National Longitudinal Surveys of Youth (NLSY). W: I. Tomescu-Dubrow, C. Wolf, K.M. Słomczynski, J. C. Jenkins (red.), *Survey Data Harmonization in the Social Sciences* (s. 73–88). Wiley. <https://doi.org/10.1002/9781119712206.ch5>

- Dubrow, J. K., Tomescu-Dubrow, I. (2016). The rise of cross-national survey data harmonization in the social sciences: emergence of an interdisciplinary methodological field. *Quality & Quantity*, 50(4), 1449–1467. <https://doi.org/10.1007/S11135-015-0215-z>
- Durand, C. (2024). On Using Harmonized Data in Statistical Analysis: Notes of Caution. W: I. Tomescu-Dubrow, C. Wolf, K.M. Slomczynski, J. C. Jenkins (red.), *Survey Data Harmonization in the Social Sciences* (s. 347–365). Wiley. <https://doi.org/10.1002/9781119712206.ch20>
- Harkness, J. A., Van de Vijver, F. J., Mohler, P. P. (red.). (2003). *Cross-cultural survey methods*. Wiley.
- Harkness, J. A., Braun, M., Edwards, B., Johnson, T. P., Lyberg, L., Mohler, P. Ph., Pennell, B.-E., Smith, T. W. (red.). (2010) *Survey Methods in Multinational, Multiregional, and Multicultural contexts*. Wiley. <https://doi.org/10.1002/9780470609927>
- Heyneman, S. P. (1999). The sad story of UNESCO's educational statistics. *International Journal of Educational Development*, 19(1), 65–74.
- Iwai, N., Mo, T., Kim, J., Wu,, C.-I., Wang, W. (2024). Harmonization in the East Asian Social Survey. W: I. Tomescu-Dubrow, C. Wolf, K.M. Slomczynski, J. C. Jenkins (red.), *Survey Data Harmonization in the Social Sciences* (s. 107–124). Wiley. <https://doi.org/10.1002/9781119712206.ch7>
- Jabkowski, P. (2021). *Sondażowe zwierciadło europeizacji*. Wydawnictwo Naukowe UAM.
- Johnson, T. P., Pennell, B.-E., Stoop, I. A. L., Dorer, B. (red.). (2019). *Advances in Comparative Survey Methods: Multinational, Multiregional, and Multicultural Contexts (3MC)*. Wiley.
- Kizilova, K., Diez-Medrano, J., Welzel, C., Haerpffer, C. (2024). Harmonization in the World Values Survey. W: I. Tomescu-Dubrow, C. Wolf, K.M. Slomczynski, J. C. Jenkins (red.), *Survey Data Harmonization in the Social Sciences* (s. 41–56). Wiley. <https://doi.org/10.1002/9781119712206.ch3>
- Koc, P., Pokropek, A. (2022). Accounting for cross-country-cross-time variations in measurement invariance testing. A case of political participation. *Survey Research Methods*, 16(1), 79–96. <https://doi.org/10.18148/srm/2022.v16i1.7909>
- Kopycka, K., Kiersztyn, A., Sawiński, Z., Bieńkowski, S., Sovpenchuk, V. (2023). Use of Panel Surveys to Measure Employment Precarity in a Cross-National Framework: An Integrated Approach to Harmonize Research Concepts and Longitudinal Data. *Survey Research Methods*, 17(3), 353–393. <https://doi.org/10.18148/srm/2023.v17i3.7989>
- Kyyrä, T., Wilke, R. A. (2014). On the reliability of retrospective unemployment information in European household panel data. *Empirical Economics*, 46, 1473–1493. <https://doi.org/10.1007/s00181-013-0718-1>
- Levi, L. (1854). Résumé of the Statistical Congress, held at Brussels, September 11th, 1853, for the purpose of introducing unity in the Statistical Documents of all Countries. *Journal of the Statistical Society of London*, 17(1), 1–14.
- Lillard, D. R. (2024). Harmonization of Panel Surveys: The Cross-National Equivalent File. W: I. Tomescu-Dubrow, C. Wolf, K.M. Slomczynski, J. C. Jenkins (red.),

- Survey Data Harmonization in the Social Sciences* (s. 169–188). Wiley. <https://doi.org/10.1002/9781119712206.ch10>
- Logan, C., Mattes, R., Kibirige, F. (2024). Harmonization in the Afrobarometer. W: I. Tomescu-Dubrow, C. Wolf, K.M. Slomczynski, J. C. Jenkins (red.), *Survey Data Harmonization in the Social Sciences* (s. 57–72). Wiley. <https://doi.org/10.1002/9781119712206.ch4>
- Meuleman, B., Żóltak, T., Pokropek, A., Davidov, E., Muthén, B., Oberski, D. L., Billiet, J., Schmidt, P. (2022). Why measurement invariance is important in comparative research. A response to Welzel i in. (2021). *Sociological Methods & Research*, 52(3), 1401–1419. <https://doi.org/10.1177/00491241221091755>
- O'Neill, D., Hardy, R. (2024). Harmonization of Survey Data from UK Longitudinal Studies: CLOSER. W: I. Tomescu-Dubrow, C. Wolf, K.M. Slomczynski, J. C. Jenkins (red.), *Survey Data Harmonization in the Social Sciences* (s. 189–206). Wiley. <https://doi.org/10.1002/9781119712206.ch11>
- Organisation for Economic Co-operation and Development (OECD). (2011). *Lessons from PISA for the United States*. OECD Publishing. <https://doi.org/10.1787/9789264096660-en>
- Pokropek, A. (2024). Comparability and Measurement Invariance. W: I. Tomescu-Dubrow, C. Wolf, K.M. Slomczynski, J. C. Jenkins (red.), *Survey Data Harmonization in the Social Sciences* (s. 323–332). Wiley. <https://doi.org/10.1002/9781119712206.ch18>
- Przeworski, A., Teune, H. (1966). Equivalence in cross-national research. *Public Opinion Quarterly* 30(4), 551–568. <https://doi.org/10.1086/267455>
- Reif, K., Inglehart, R. (red.). (1989), *Eurobarometer: The Dynamics of European Public Opinion. Essays in Honour of Jacques-René Rabier*. MacMillan Academic. <https://doi.org/10.1007/978-1-349-21476-1>
- Rosselló, P. (1979). Historical note. W: *International Conference on Education. Recommendations 1934-1977* (xi-xxii). UNESCO.
- Sawiński, Z. (2013). The International Standard Classification of Education ISCED: the standard from Sevrès or Satan's trick? *Edukacja*, 125(5), 115–134.
- Schneider, S. L. (2010). Nominal comparability is not enough:(In-) equivalence of construct validity of cross-national measures of educational attainment in the European Social Survey. *Research in Social Stratification and Mobility*, 28(3), 343–357. <https://doi.org/10.1016/j.rssm.2010.03.001>
- Slomczynski, K.M., Wolf, C., Tomescu-Dubrow, I., Jenkins, J.C. (2024). On the Future of Survey Data Harmonization. W: I. Tomescu-Dubrow, C. Wolf, K.M. Slomczynski, J. C. Jenkins (red.), *Survey Data Harmonization in the Social Sciences* (s. 367–376). <https://doi.org/10.1002/9781119712206.ch21>
- Stigler, S. M. (1990). *The History of Statistics. Measurement of Uncertainty before 1900*. Harvard University Press.
- Szacki, J. (1981). *Historia myśli socjologicznej*. Tom 1 i 2. Wydawnictwo Naukowe PWN.
- Tomescu-Dubrow, I., Slomczynski, K.M., Wymulek, I., Powalko, P., Li, O., Tu, Y., Slarzynski, M., W. Zielinski, M., Lavryk, D. (2024). Harmonization for Cross-Na-

- tional Secondary Analysis: Survey Data Recycling. W: I. Tomescu-Dubrow, C. Wolf, K.M. Slomczynski, J. C. Jenkins (red.), *Survey Data Harmonization in the Social Sciences* (s. 145–167). Wiley. <https://doi.org/10.1002/9781119712206.ch9>
- Tomescu-Dubrow, I., Slomczynski, K. M., Sawiński, Z., Kiersztyn, A., Janicka, K., Życzyńska-Ciołek, D., Wysmulek, I., Kotnarowski, M. (2021). The Polish Panel Survey, POLPAN. *European Sociological Review*, 37(5), 849–864. <https://doi.org/10.1093/esr/jcab017>
- Tröhler, D. (2013). The OECD and Cold War Culture: thinking historically about PISA. W: H. D. Meyer, A. Benavot (red.), *PISA, power, and policy: The emergence of global educational governance* (s.141–161). Symposium Books.
- Van de Vijver, F. J., Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology*, 54(2), 119–135. <https://doi.org/10.1016/j.erap.2003.12.004>
- Weiß, B., Schulz, S., Schmid, L., Sterl, S., Haensch, A.-C. (2024). Harmonizing and Synthesizing Partnership Histories from Different German Survey Infrastructures. W: I. Tomescu-Dubrow, C. Wolf, K.M. Slomczynski, J. C. Jenkins (red.), *Survey Data Harmonization in the Social Sciences* (s. 249–268). Wiley. <https://doi.org/10.1002/9781119712206.ch14>
- Wey, T. W., Fortier, I. (2024). Maelstrom Research Approaches to Retrospective Harmonization of Cohort Data for Epidemiological Research. W: I. Tomescu-Dubrow, C. Wolf, K.M. Slomczynski, J. C. Jenkins (red.), *Survey Data Harmonization in the Social Sciences* (s. 227–247). Wiley. <https://doi.org/10.1002/9781119712206.ch13>
- Wojciuk, A. (2016). *Imperia wiedzy*. Wydawnictwo Naukowe Scholar.

