

**Paweł Daniłowicz, Paweł B.
Sztabiński**

**Weryfikacja danych osobowych
otrzymywanych w badaniach z
zastosowaniem wywiadu
kwestionariuszowego**

Przegląd Socjologiczny Sociological Review 30, 307-314

1978

Artykuł został zdigitalizowany i opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej bazhum.muzhp.pl, gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

PAWEŁ DANIŁOWICZ
PAWEŁ B. SZTABIŃSKI
Instytut Filozofii i Socjologii PAN

WERYFIKACJA DANYCH OSOBOWYCH OTRZYMYWANYCH W BADANIACH Z ZASTOSOWANIEM WYWIADU KWESTIONARIUSZOWEGO

Spośród informacji zdobywanych w badaniach kwestionariuszowych na szczególną uwagę ze względu na ich podstawowe znaczenie zwłaszcza w badaniach nad strukturą społeczną zasługują dane osobowe tworzące tzw. metryczkę kwestionariusza. Potrzeba ich ujednoczenia i oceny ich wiarygodności spowodowały podjęcie problematyki standaryzacji i weryfikacji pytań o dane osobowe w kwestionariuszowych badaniach socjologicznych. Od roku 1975 w Zespole Metodologii Badań Socjologicznych IFiS PAN z inicjatywy K. Lutyńskiej rozpoczęto systematyczną refleksję metodologiczną i badania nad problemem standaryzacji i weryfikacji tzw. pytań metryczkowych¹. W rezultacie postanowiono opracować wzorcowy zestaw pytań osobowych. Stosowanie takiego ujednoczonego zestawu pytań zapewni porównywalność otrzymywanych za ich pomocą informacji, a także ułatwi wtórne wykorzystywanie materiałów z badań kwestionariuszowych z zastosowaniem takich pytań. Podjęta przez nas problematyka stanowi kontynuację metodologicznej refleksji i badań w zakresie standaryzacji pytań osobowych zawartych w dwu tomach pt. *Standaryzacja zmiennych socjologicznych*² oraz badań poświęconych weryfikacji danych z wywiadów kwestionariuszowych prowadzonych w ośrodku

¹ K. Lutyńska, *Kilka refleksji w związku z publikacją »Standaryzacja zmiennych socjologicznych« pod redakcją W. Wesotowskiego*, „Przegląd Socjologiczny”, t. 29, 1977. Autorka omawiając problem standaryzacji pytań metryczkowych wskazuje jednocześnie na potrzebę oceny wiarygodności informacji uzyskiwanych za ich pomocą.

² *Standaryzacja zmiennych socjologicznych*, red. W. Wesółowski, t. 1—2, Warszawa 1974.

łódzkim³. Aby dojść do zestawu pytań osobowych, za pomocą których można uzyskać porównywalne informacje, została dokonana analiza „metryczek” 220 kwestionariuszy do badań przeprowadzonych bądź planowanych przez badaczy z IFiS PAN, OBOPiSP, UW i UŁ w latach 1958—1974. Pytania zawarte w tych kwestionariuszach zostały poddane analizie i ocenie z punktu widzenia poszukiwanych informacji i ich użyteczności. Dla całkowitej oceny tych pytań wzięto również pod uwagę sposób ich wykorzystania w opracowaniach będących rezultatem badań.

Analizy te wykazały, że istniejąca w Polsce praktyka formułowania pytań odznacza się dużą różnorodnością zarówno zapotrzebowań badawczych odnoszących się do danej zmiennej, jak i sformułowań pytań. Znaczna ich część miała niepoprawną formalnie konstrukcję, a więc cechowała je niejasność lub nadmierna trudność. Ponadto część tych pytań miała charakter haseł wywoławczych dla ankietera, który w ten sposób każdorazowo był zmuszany do formułowania pytań w sposób dowolny, a więc niejednokrotnie na pewno wadliwy.

Na podstawie wyników przeprowadzonej analizy kwestionariuszy oraz uwzględniając doświadczenia konsultacyjno-pilotażowe Zespołu opracowano zestandaryzowany zestaw pytań o dane osobowe⁴. Zestaw ten obejmuje następujące zmienne: 1) wiek respondenta, 2) miejsce wychowania respondenta, 3) stan cywilny respondenta, 4) wykształcenie respondenta, 5) liczba osób, skład i główne źródła utrzymania członków gospodarstwa domowego, 6) zawód respondenta, 7) wykształcenie współmałżonka, 8) zawód współmałżonka, 9) przychody gospodarstwa domowego, 10) saldo przychodów, 11) przynależność organizacyjna respondenta, 12) wykształcenie ojca respondenta, 13) zawód ojca respondenta, 14) warunki mieszkaniowe gospodarstwa domowego, 15) zasoby gospodarstwa domowego.

Każda z tych zmiennych została zdefiniowana w celu ustalenia jednolitego ich rozumienia. Podstawą naszych propozycji w tym zakresie były tomiki *Standaryzacja zmiennych socjologicznych*, dostępna litera-

³ Problemowi temu poświęcony był zwłaszcza 5 tom pracy *Analizy i próby technik badawczych w socjologii*, red. Z. Gostkowski, J. Lutyński, Wrocław 1975, oraz niektóre analizy z poprzednich tomów tego wydawnictwa.

⁴ Zestaw ten został przedstawiony do oceny Zespołowi Standaryzacji i Weryfikacji Zmiennych Metryczkowych, zorganizowanemu w ramach prac nad problemem węzłowym 11.2: „Ewolucja struktury społeczeństwa socjalistycznego”, którego kierownikiem jest prof. dr W. Wesołowski. Zespół, pracujący pod kierunkiem prof. dra J. Lutyńskiego, a składający się z badaczy z różnych ośrodków socjologicznych, po przedyskutowaniu uznał przedstawiony zestaw jako poprawny formalnie i merytorycznie.

tura socjologiczna i ekonomiczno-statystyczna oraz sens nadawany przez badaczy poszczególnym zmiennym.

Przy formułowaniu pytań kierowano się zasadą wyodrębniania dla poszczególnych zmiennych najbardziej szczegółowych pytań badacza — a więc takich, przy których badacz dąży do uzyskania największej liczby informacji na temat poszczególnych zmiennych. Tak sformułowanym pytaniom badacza zostały przyporządkowane odpowiednie pytania kwestionariuszowe, które mają dostarczyć jak największej liczby informacji dotyczących danych osobowych. Biorąc jednak pod uwagę, że nie wszyscy badacze mają zapotrzebowanie na tak szczegółowe informacje w stosunku do poszczególnych zmiennych, najważniejsze z nich przedstawiono w dwu lub trzech wersjach. Sformułowanie dwu lub trzech wersji pytań w stosunku do niektórych zmiennych było rezultatem analizy opracowań i publikacji relacjonujących wyniki z badań empirycznych, która wykazała, że zmienne te są wykorzystywane z różnym stopniem szczegółowości, co podyktowane bywa zwykle charakterem samych badań. W ostatecznym sformułowaniu pytań uwzględniono również wyniki pogłębionych badań pilotażowych przeprowadzonych w 40 przypadkach⁵. Przedstawione wersje różnią się stopniem szczegółowości informacji przynoszonych przez pytania. Pomiędzy pytaniem w wersji ogólnej a odpowiadającym mu blokiem pytań w wersji szczegółowej istnieje ścisła współzależność, tzn. informacje uzyskane za pomocą pytań szczegółowych można na mocy pewnych reguł wnioskowania sprowadzić do informacji, która zostałaby uzyskana za pomocą pytania ogólnego⁶.

Poprawność formalna i merytoryczna pytań nie przesądza jeszcze o wiarygodności danych uzyskiwanych za ich pomocą. Wiarygodność uzyskiwanych danych można ustalić dopiero po konfrontacji informacji dostarczonych przez pytania z danymi uzyskanymi za pewniejsze. Dopiero taki zabieg jest podstawą pełnej oceny przydatności pytania.

W celu sprawdzenia wiarygodności odpowiedzi uzyskiwanych za pomocą zaproponowanych przez nas pytań metryczkowych od roku 1976 są prowadzone specjalne badania⁷. Podstawowym celem tych badań

⁵ Ogólne zasady przeprowadzania badań pilotażowych, według których przeprowadzono również i ten pilotaż, zawarte są w artykule K. Lutyńskiej, *Pilotaż „pogłębiony”. Koncepcja, realizacja i analiza materiałów pilotażowych w: Analizy i próby ...*, t. 5.

⁶ Pełny tekst pytań we wszystkich wersjach wraz z instrukcją dla ankietorów, poszukiwanymi informacjami oraz uzasadnieniem zawarty jest w pracy P. Daniłowicza i P. B. Sztabińskiego, *Propozycje pytań metryczkowych*, Warszawa 1977.

⁷ Badania te finansowane są z funduszków problemu węzłowego 11.2.

jest więc weryfikacja danych metryczkowych uzyskiwanych w kwestionariuszowych badaniach socjologicznych. Kwestionariusz do badań składa się z około 90 pytań: są to pytania o dane osobowe (w różnych wersjach kwestionariusza uwzględniono wszystkie wersje pytań) oraz specjalnie dobrane pytania dotyczące struktury społecznej (są to pytania najczęściej występujące w badaniach dotyczących tej problematyki). Pytania dotyczące struktury społecznej zastosowano ze względu na konieczność upodobnienia naszego kwestionariusza do innych kwestionariuszy socjologicznych, zarówno pod względem zawartości merytorycznej, jak i czasu trwania wywiadu. Innym równie ważnym celem jest ustalenie i oszacowanie wpływu ewentualnych błędów, popełnianych przez respondentów w odpowiedziach na pytania metryczkowe, na korelacje danych uzyskanych z odpowiedzi na te pytania z danymi uzyskanymi za pomocą pytań dodatkowych.

Badania są oparte na próbie adresowej pracowników zatrudnionych w łódzkim przemyśle kluczowym dobranych według następujących zasad: 1) ustalono udział poszczególnych przemysłów liczony liczbą zatrudnionych w obrębie łódzkiego przemysłu kluczowego, 2) dobrano po jednym zakładzie pracy, typowym dla każdego z tych przemysłów, 3) w każdym zakładzie wylosowano do badań (stosując losowanie odstępowe) taką liczbę respondentów, która odpowiada proporcji ogółu zatrudnionych w danym przemyśle. W założonej próbie liczącej ogółem 450 osób miał być zachowany udział pracowników poszczególnych przemysłów odpowiadający udziałowi tych przemysłów w obrębie łódzkiego przemysłu kluczowego. Na podstawie przebadania opisanej tu próby moglibyśmy wyciągać wnioski dotyczące jedynie zbiorowości o zbliżonej strukturze zatrudnienia. Dla zapewnienia możliwości formułowania wniosków bardziej ogólnych, dotyczących zbiorowości o strukturach bardziej zbliżonych do ogólnopolskiej struktury zatrudnienia, dobrano dodatkowo 150 osób zatrudnionych w jednym z zakładów przemysłu metalowego (struktura zatrudnienia w łódzkim przemyśle kluczowym nie jest typowa w skali kraju). W rezultacie zebrano więc ogółem 600 wywiadów kwestionariuszowych.

Badania były prowadzone od 1 IX do 15 X 1976 r. Wywiady przeprowadzało 35 ankierów ogólnopolskiej sieci IFiS PAN. Przed wystąpieniem do fazy realizacji terenowej badań wszyscy ankierzy uczestniczyli w dwudniowym szkoleniu. Szkolenie było poprzedzone rozdaniem kwestionariuszy wraz z instrukcją. Chodziło o to, aby przed rozpoczęciem szkolenia ankierzy zapoznali się z narzędziem. W pierwszym dniu szkolenia wyjaśniano wątpliwości i pytania ankierów związane z wywiadem, w drugim przeprowadzono kilka wywiadów szkole-

niowych pod kontrolą autorów badań. Przed badaniami wysyłano do wszystkich respondentów listy zapowiadające przybycie ankietera, a zwroty stanowiły wstępną kontrolę aktualności adresów. Każdy ankieter miał obowiązek sukcesywnego oddawania przeprowadzonych wywiadów, ustalono, że po 5 wywiadach miał on obowiązek skontaktować się z organizatorami badań. Forma częściowego zwrotu przeprowadzonych wywiadów miała na celu kontrolę realizowanej próby, sprawdzenie poprawności przeprowadzonych wywiadów i wyjaśnienie ankieterom wątpliwości, jakie nasunęły się im w trakcie badań. Przeciętny czas trwania wywiadów wynosił około 45 min., przy czym rozpiętości między najkrótszym a najdłuższym wywiadem były znaczne, a podyktowane tym, że czas przeprowadzania rozmowy był zależny od liczby osób w gospodarstwie domowym respondenta⁸. W czasie trwania fazy terenowej badań kontrolowano pracę wszystkich ankieterów. Oprócz wspomnianych wyżej nieterenowych metod kontroli polegających na porównywaniu niektórych informacji uzyskiwanych przez ankieterów z danymi wiarygodnymi, będącymi w posiadaniu organizatorów, stosowano także terenowe metody kontroli⁹. Podstawową metodą była ankieta skierowana do respondentów — uczestników badań kwestionariuszowych.

Zasadniczą metodą weryfikacji uzyskanych danych była weryfikacja dokumentalna. Objęto nią nie tylko respondentów, ale również pozostałych członków gospodarstwa domowego. Polegała ona na zebraniu, na podstawie dokumentów istniejących w różnych instytucjach, informacji dotyczących interesujących nas zmiennych, przy czym informacje zawarte w dokumentach uznane były przez nas za bardziej wiarygodne, niż dane uzyskiwane w trakcie wywiadu kwestionariuszowego. Informacje były wypisywane z dokumentów na specjalnie przez nas przygotowane karty weryfikacyjne. Na kartach tych zapotrzebowanie informacyjne sformułowane było na najniższym szczeblu ogólności. Tak np. w przypadku zarobków z głównego miejsca pracy zbierano osobno informacje o zarobkach podstawowych w poszczególnych miesiącach roku, osobno o premiach, nagrodach i pozostałych formach zarobków. Liczono się bowiem z tym, że dokonywanie na tym etapie jakichkolwiek operacji, np. sumowania, może doprowadzić do błędów. Chodziło o to, aby sprowadzić operacje związane z weryfikacją jedynie do wyboru z istniejących dokumentów potrzebnych nam danych.

⁸ W badaniach pytano bowiem także o wykształcenie i zawód współmałżonka oraz o dochody każdego z członków gospodarstwa domowego.

⁹ Informacje te, dotyczące wieku, zawodu i miejsca pracy respondentów, zdobyto w trakcie losowania do badań.

Weryfikacje przeprowadzono w zakładach pracy (dotyczyły one zmiennych: zawodu, wykształcenia, przynależności organizacyjnej, wieku, zarobków, stanu cywilnego), w ADM-ach (co do warunków mieszkaniowych), w ZUS-ie (co do rent, emerytur), w Wydziale Komunikacji (co do posiadania pojazdów) oraz w innych instytucjach. Do dokumentów istniejących we wszystkich tych instytucjach konieczne było podejście krytyczne. Należy zdawać sobie sprawę, że zawarte w nich informacje mogły być nieścisłe. Tak np. mogą występować różnice między zawodem faktycznie wykonywanym a formalnym w umowie zatrudnienia lub też między wykształceniem faktycznym a podanym w ankiecie personalnej (istnieje kategoria wykształcenia udokumentowanego i nie udokumentowanego). Wszystkie osoby przeprowadzające weryfikacje zostały uczulone na możliwość występowania nieścisłości w dokumentach i zobowiązane do odnotowywania takich przypadków. Oczywiście krytyczne podejście do dokumentów znajdzie swoje odbicie w analizach poświęconych wiarygodności informacji uzyskiwanych na drodze wywiadu kwestionariuszowego. Ze względu na to, że losowania dokonywano w zakładach pracy, podstawową część weryfikacji przeprowadzili pracownicy tych zakładów, pracujący w dziale kadr, rachubie i dziale wypłat. Weryfikację informacji dotyczącą członków gospodarstwa domowego przeprowadzili specjalnie dobrani studenci wyższych lat socjologii. Wszyscy weryfikatorzy zostali starannie przeszkoleni, niezależnie od tego otrzymali instrukcję pisemną, a ich praca była wryrykowo kontrolowana przez autorów badań. Faza ta w toku realizacji nastroczała szereg trudności, ponieważ jedynymi informacjami posiadanymi w stosunku do członków gospodarstwa domowego są nazwiska (bez imion), adresy miejsca pracy i główne źródło utrzymania. Na podstawie tych informacji nie we wszystkich przypadkach jest możliwe zdobycie pełnej dokumentacji dotyczącej każdej z tych osób¹⁰.

Jako dodatkowy sposób weryfikacji odpowiedzi respondentów zastosowano (w stosunku do części badanych) technikę autoweryfikacji wywiadu kwestionariuszowego¹¹. Wywiad z autoweryfikacją miał na

¹⁰ W przypadku zarobków zaszła konieczność wyeliminowania również i tych osób, które w ciągu ostatnich 12 miesięcy zmieniły miejsce pracy, gdyż w Karcie Weryfikacyjnej informacja była wówczas niepełna (odnosiła się do krótszego okresu), a nie było możliwości ustalenia poprzedniego miejsca pracy. Jest bowiem prawdopodobne, że wraz ze zmianą miejsca pracy mogły się zmienić również i zarobki. Ponadto niektórzy respondenci podali niezgodną z prawdą informację o miejscu pracy członków gospodarstwa domowego.

¹¹ Metoda ta została zastosowana w naszych badaniach przez J. Krak. Sze-rzej omawia tę technikę i pierwszą próbę jej zastosowania Z. Gostkowski, *O poprawę jakości badań survey'owych*, „Studia Socjologiczne”, 1976, nr 3.

celu uzyskanie od respondentów pełniejszych i bardziej wiarygodnych informacji. Polegał on na pozostawieniu respondentowi, po przeprowadzeniu normalnego wywiadu, kwestionariusza z zapisanymi przez ankietera odpowiedziami. Badany, dzięki powtórnemu przeczytaniu treści pytań i własnych odpowiedzi, miał możliwość skorygowania ewentualnych błędów w odpowiedziach lub w ich zapisie. Pozostawienie kwestionariusza u respondenta pozwala mu przemyśleć problemy, których dotyczą poszczególne pytania, i dokonać uzupełnień lub zmian w udzielonych odpowiedziach. Autoweryfikację zaproponowano 265 osobom. Zgodziło się jej dokonać 210 osób — 79,3%. Wśród osób, które przyjęły kwestionariusz do autoweryfikacji, 79 badanych (37,6%) dokonało zmian polegających głównie na uściśleniu odpowiedzi na pytania dotyczące konsumpcji kulturalnej¹².

Przeprowadzone przez nas badania są pierwszą w Polsce zakrojoną na większą skalę próbą kompleksowej weryfikacji danych osobowych. Wcześniejsze próby ustalenia wiarygodności takich informacji były prowadzone wyrywkowo (tzn. dotyczyły tylko jednej zmiennej) na niewielkich próbach i przy różnych poszukiwanych informacjach dotyczących tej samej zmiennej¹³.

Biorąc pod uwagę weryfikacyjny charakter naszych badań dążyliśmy do eliminowania błędów nie tylko w terenowej fazie badań, lecz także na etapie opracowania materiałów. Dlatego też, między innymi, w fazie opracowania przewidziano dwukrotne kodowanie.

W rezultacie przeprowadzonych przez nas badań weryfikacyjnych możliwe będzie korygowanie informacji uzyskiwanych od respondentów w różnych badaniach. Wiadomo bowiem, że respondenci w zależności od cech społeczno-demograficznych oraz np. od wysokości zarobków mają tendencję do popełniania błędów w różnych kierunkach i o różnej wielkości. Określenie tych tendencji oraz zakresu błędów po-

¹² Wstępne wyniki zastosowania tej techniki w omawianych badaniach omawia J. Krak, *Wstępne wyniki zastosowania technik tzw. autoweryfikacji wywiadu kwestionariuszowego*, komunikat przygotowany na V Ogólnopolski Zjazd Socjologiczny, Kraków 1977 (powiel.).

¹³ Przykładem takiej weryfikacji w odniesieniu do zarobków w głównym miejscu pracy były badania K. Szafnickiego (nie opublikowana praca magisterska), K. Lutyńskiej (*Weryfikacja danych uzyskanych z odpowiedzi respondentów na pytania o wysokość zarobków*, w: *Analizy i próby ...*, t. 5) oraz J. Koniarka (*Weryfikacja zewnętrzna informacji uzyskanych w wywiadzie kwestionariuszowym*, *ibidem*). W badaniach naszych uwzględnione zostały wszelkie doświadczenia i sugestie autorów powyższych publikacji.

zwoli na wprowadzenie korekt do wyników badań. Pozwoli to również zweryfikować hipotezę przyjmowaną przez wielu badaczy mówiącą, że błędy popełniane przez respondentów w odpowiedziach na pytania metryczne zacierają obraz zależności między cechami. Innymi słowy, sprawdzone zostanie twierdzenie, iż brak błędów w odpowiedziach respondentów sprawiłby, że uzyskiwane korelacje osiągałyby inne wartości.