

# Katarzyna Staszyńska

---

## Wiarygodność danych w socjologicznym badaniu ankietowym

---

Przegląd Socjologiczny Sociological Review 37, 163-175

---

1989

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](http://bazhum.muzhp.pl), gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

KATARZYNA STASZYŃSKA – WARSZAWA

## WIARYGODNOŚĆ DANYCH W SOCJOLOGICZNYM BADANIU ANKIETOWYM

Wyniki wielu kontrolowanych badań metodologicznych wskazują, że niski poziom wiarygodności danych uzyskanych w badaniu ankietowym nie jest zdarzeniem wyjątkowym, a typowym. Jeżeli za kryterium wiarygodności danych przyjmuje się ich zgodność z „obiektywnym stanem rzeczy”, to w kontrolowanych badaniach metodologicznych okazuje się, że rozbieżności mogą być bardzo duże. W pewnych badaniach amerykańskich, w których dane o „obiektywnym stanie rzeczy” uzyskano z oficjalnych, urzędowych dokumentów, rozbieżności te sięgały od 2% wszystkich odpowiedzi na pytanie o posiadanie telefonu, przez 10% odpowiedzi na pytanie o posiadanie prawa jazdy i 25% odpowiedzi na pytanie o rejestrację przedwyborczą do 40% wszystkich odpowiedzi na pytanie o świadczenia respondenta na rzecz funduszu zapomogowego (Kahn i Cannell, 1957:176). W polskich badaniach, w których zastosowano podobną procedurę kontroli wiarygodności odpowiedzi uzyskanych na pytania metryczkowe, stwierdzone rozbieżności wynosiły 28% wszystkich odpowiedzi na pytanie o wykształcenie respondenta, a w pytaniach o zarobki respondenta i dochody jego rodziny znaczące rozbieżności uzyskano w około 60% przypadków (Daniłowicz i Sztabiński, 1983:314 – 315).

Ostatnie lata przyniosły wiele interesujących i nowych rezultatów metodologicznych studiów nad wiarygodnością danych uzyskanych w badaniach ankietowych. Stan badań metodologicznych nie jest jednak tak zaawansowany, aby wypracowane techniki kontroli jakości danych mogły być stosowane w typowych substancywnych badaniach ankietowych; w typowych badaniach substancywnych badacz nadal nie dysponuje prostymi i tanimi technikami oceny wiarygodności uzyskanych przez siebie danych. Badania metodologiczne nadal koncentrują się na identyfikacji źródeł obciążenia (bias) danych (por. Sudman i Bradburn, 1974; Bradburn i Sudman, 1979; Schuman i Presser, 1981; Dijkstra i van der Zouwen, 1982).

W artykule tym dokonamy syntetycznego przeglądu współczesnych analiz wiarygodności danych uzyskanych w badaniu ankietowym i zwrócimy uwagę na podstawowe problemy związane z badaniem wiarygodności tych danych. W prezentacji opierać się będziemy przede wszystkim na literaturze obcej; prace polskie, jako lepiej znane polskiemu czytelnikowi, zostały niemal pominięte.

#### I. ZMIENNE WYJAŚNIANE W ANALIZACH WIARYGODNOŚCI DANYCH UZYSKANYCH W BADANIU ANKIETOWYM

W analizach metodologicznych przedmiotem wyjaśniania czyni się cztery charakterystyki jakości danych uzyskanych w badaniu ankietowym:

1) **obciążenie odpowiedzi** (response bias). Obciążenie odpowiedzi definiowane jest jako różnica między odpowiedzią uzyskaną (wynikiem uzyskanym) a odpowiedzią prawdziwą (wynikiem prawdziwym) (Hagenaars i Heinen, 1982:92). Obciążenie odpowiedzi może być analizowane na poziomie jednostkowej odpowiedzi na jednostkowe pytanie kwestionariusza i na poziomie całej próby; w definicji pierwotnej termin ten odnosi się do jednostkowej odpowiedzi;

2) **wariancja odpowiedzi** (response variance). Definiowana jest jako zróżnicowanie odpowiedzi uzyskanych w różnych warunkach (Hagenaars i Heinen, 1982:92), na przykład przez różnych ankierów, przy zastosowaniu różnych technik badawczych, różnych sformułowań pytania. Definicja pierwotna odnosi się do jednostkowej odpowiedzi na pytanie kwestionariusza; termin jest także stosowany do poziomu całej próby (np. do zróżnicowania rozkładów odpowiedzi uzyskiwanych w różnych warunkach);

3) **kompletność odpowiedzi** (response completeness). Definiowana jest jako ilość informacji użytecznych, które zawiera zarejestrowana odpowiedź (Hagenaars i Heinen, 1982:93). Odniesienie tego terminu do poziomu całej próby jest mało zasadne<sup>1</sup>.

4) **stopień realizacji próby** (response rate). Jest to proporcja liczby uzyskanych ankiet lub wywiadów w stosunku do założonej liczebności próby (Alwin, 1977: 13–14). W szerszym znaczeniu termin „response rate” może być zastosowany do oznaczenia proporcji informacji uzyskanych przez badacza w odpowiedzi na jednostkowe pytanie kwestiona-

---

<sup>1</sup> W terminologii J. Lutyńskiego byłaby to operacyjnie zdefiniowana istotność odpowiedzi (por. Lutyński, 1972; 1978; 1979). W starszych pracach Lutyński używał terminu „odpowiedź pożądana” do oznaczenia odpowiedzi przynoszącej informacje poszukiwane przez badacza, w nowszej zastąpił go terminem „odpowiedź istotna”. w propozycji Lutyńskiego nie jest jednak ta cecha jakości odpowiedzi zdefiniowana operacyjnie.

riusza w stosunku do liczby respondentów (inaczej mówiąc: jest to miara wskazująca na częstość pojawiania się odmów odpowiedzi na jednostkowe pytanie kwestionariusza)<sup>2</sup>.

W większości analiz metodologicznych rozważane są źródła i wielkość obciążenia (bias) i wariancji odpowiedzi na poziomie próby. W literaturze termin „bias” używany bywa zamiennie z terminem „error” (błąd), co w ścisłym sensie nie jest uprawnione. Błąd w badaniu ankietowym może, ale nie musi być źródłem obciążenia danych. Błąd systematyczny, to znaczy taki, który oddziałuje w jednym kierunku, rodzi obciążenie danych w każdym wypadku; błąd losowy (w ścisłym sensie taki, który w dużej liczbie pomiarów ma średnią wartość równą lub bliską 0) w szczególnych przypadkach w badaniu ankietowym także może rodzić obciążenie danych<sup>3</sup> (Kahn i Cannell, 1957:168–172).

Operacyjna definicja terminu „obciążenie odpowiedzi” nasuwa szereg poważnych trudności. Zdefiniowaliśmy obciążenie odpowiedzi jako różnicę między wynikiem zapisanym w badaniu ankietowym a wynikiem prawdziwym. Znalezienie kryterium prawdziwości odpowiedzi w socjologicznym badaniu ankietowym jest więc sprawą o kluczowym znaczeniu. Wyodrębnienie kryterium prawdziwości jest trudne już na poziomie koncepcyjnym, jeszcze trudniejsza może okazać się jego operacjonalizacja, a możliwości testowania prawdziwości danych uzyskanych w ankietowych badaniach socjologicznych zwykle nie ma zbyt wiele (Hyman, 1954:226; Kahn i Cannell, 1957:172–174). Jednak w literaturze metodologicznej za wiarygodne uznaje się takie dane (na każdym poziomie analizy ich wiarygodności), które są prawdziwe (por. Dean i Whyte, 1958). Hagens i Heinen (1982:92) podają za innymi takie oto warunki, które spełniać musi kryterium prawdziwości odpowiedzi:

Wartość prawdziwa musi być jednoznacznie (uniquely) zdefiniowana.

Wartość prawdziwa musi być zdefiniowana w taki sposób, który można odnieść do przedmiotu badania.

Jeżeli jest to możliwe w zgodności z dwoma pierwszymi kryteriami, wartość prawdziwa musi być zdefiniowana w terminach operacji, które rzeczywiście mogą być przeprowadzone (nawet wtedy, kiedy ich przeprowadzenie może być kosztowne lub trudne).

---

<sup>2</sup> Stopień realizacji próby (response rate) występuje jako zmienna zależna głównie w badaniach wpływu różnych technik badawczych i technik pomocniczych na realizację celów badawczych (por. badania wpływu informed consent procedure, np.: Singer, Frankel, 1982).

<sup>3</sup> Rozróżnienie pomiędzy błędem losowym a systematycznym w badaniu ankietowym, pod pewnym względem, nie ma wogóle sensu. Za tym twierdzeniem przemawia kilka racji. Po pierwsze, badania ankietowe nie zawsze przeprowadzane są na wystarczająco dużych próbach, aby można było uznać, że średnia błędów losowych zbliża się do 0. Jeżeli liczebność próby jest mała, błąd całkowity (total error) jest więc sumą błędów losowych i błędów systematycznych. W dodatku indentyfikacji rodzaju błędów na poziomie jednostkowym (tzn. na poziomie danych otrzymanych od jednego respondenta w odpowiedzi na jednostkowe pytanie kwestionariusza) nigdy nie jest możliwa.

Postulowane w literaturze kryteria prawdziwości dają się podzielić na trzy rodzaje. Po pierwsze: kryterium prawdziwości jest kryterium psychologicznym (Dean i Whyte, 1958; Warwick i Lininger, 1975; Gorden, 1981, Nowak, 1965). Znaczy to, że informacja uzyskana od badanego może być uznana za prawdziwą, o ile nie została podana z intencją wprowadzenia badacza w błąd i jeżeli badany był zdolny (intelektualnie i emocjonalnie) do podania informacji prawdziwej. Wadą kryterium psychologicznego jest brak jego operacyjnego sensu. W klasycznej pracy Deana i Whyte'a (1958) także nie ma to kryterium operacyjnego sensu. Dean i Whyte rozważając prawdziwość danych subiektywnych wskazują na pewne psychiczne i intelektualne zdolności i procesy badanego, których zajście (lub niezajście) badacz musi założyć, aby ufać w prawdziwość uzyskanych odpowiedzi (Dean i Whyte, 1958:181–182). Podobne założenia postulowane są przy rozważaniu prawdziwości wypowiedzi o faktach obiektywnych (Dean i Whyte, 1958:184). Jedyny operacyjny sens kryterium prawdziwości przybiera w sugestii porównywania danych uzyskanych od respondenta z danymi uzyskanymi od innych informatorów.

W pracy Warwicka i Liningera psychologiczny sens kryterium prawdziwości wyraża się w analizie czynników sprzyjających (*positive forces*) i utrudniających (*negative forces*) właściwe zachowania się respondenta w sytuacji wywiadu, to znaczy w konsekwencji czynników sprzyjających i utrudniających udzielanie odpowiedzi o dobrej jakości (Warwick i Lininger, 1975:183–202). W pracy Gordena psychologiczne znaczenie kryterium prawdziwości nadane jest w obszernej analizie czynników hamujących (*inhibitors*) i ułatwiających (*facilitators*) komunikację w sytuacji wywiadu (Gorden, 1981:104–135). Ścisłe operacyjnego sensu kryterium prawdziwości rozumianego w sensie psychologicznym (definiowanego jako trafność funkcji ekspresyjnej i komunikacyjnej odpowiedzi respondenta) nie podaje także Nowak (1965:322 i dalsze)<sup>4</sup>.

Po drugie: informacja prawdziwa to taka, która może uzyskać potwierdzenie z innego, niezależnego od pierwotnego źródła, które z pewnych powodów można uznać za bardziej wiarygodne (pewne) (Kahn i Cannell, 1957; Hyman, 1954)<sup>5</sup>. Rzadko się zdarza, żeby w studiach socjo-

<sup>4</sup> W modelu Nowaka kryterium psychologiczne rozumiane jest dwojako: odnosi się do trafności funkcji komunikacyjnej i ekspresyjnej odpowiedzi. W modelu tym uwzględnia się także kryterium „obiektywne”, które rozumiane jest jako trafność funkcji poznawczej odpowiedzi.

<sup>5</sup> W pracach Hymana oraz Kahna i Cannella dyskutowane jest także pierwsze i trzecie kryterium prawdziwości, przy czym u Hymana kryterium statystyczne stosowane jest w dwojakim sensie. Przy definicji efektu ankietarskiego netto za wartość prawdziwą przyjmuje się wartość w populacji, natomiast przy definicji zróżnicowania ankietarskiego (*interviewer variance*) wartość prawdziwa to wartość oczekiwana, to znaczy średnia wartość z pomiaru. To drugie znaczenie kryterium statystycznego jest często spotykane w praktyce metodologicznej i w badaniach substancywnych, my uznajemy je za szczególny przypadek znaczenia pierwszego i uważamy, że generalnie posługiwano się nim w sytuacji braku danych o wartości zmiennej w populacji. Hyman czyni to z innego powodu; drugie znaczenie kryterium prawdziwości ma w jego pracy sens teoretyczny.

logicznych dostępne były niezależne źródła wiedzy o danych subiektywnych (o stanach świadomości badanego). Techniki zewnętrznej weryfikacji prawdziwości danych w badaniu ankietowym mogą być dostępne w bardzo szczególnych wypadkach i niemal wyłącznie w odniesieniu do danych obiektywnych (por. Bradburn i Sudman, 1977:1–13; także Hyman, 1954:233–234; Kahn i Cannell, 1957:176–179, a z prac polskich: Słomczyński, 1966; Daniłowicz i Sztabiński, 1983; 1986). Należy poza tym podkreślić, że techniki zewnętrznej weryfikacji odpowiedzi stosowane są niemal wyłącznie przy okazji badań metodologicznych; w badaniach substancyjnych są bowiem bardzo pracochłonne i kosztowne.

Po trzecie: w literaturze spotykane jest jeszcze kryterium prawdziwości odpowiedzi uzyskanych w badaniu ankietowym, które nazwiemy statystycznym (Hyman, 1954; Lueptow, 1977:69). Wartość prawdziwa przy zastosowaniu tego kryterium definiowana jest jako wartość z populacji; jeżeli rozkład cechy w próbie (losowej) jest statystycznie bliski rozkładowi cechy w populacji, to jest prawdziwy. Kryterium statystyczne ma dwie wady: pierwsza – jest sensowne przy ocenie prawdziwości danych na poziomie całej próby (więc rozkładów) i nie może być zastosowane do oceny wiarygodności jednostkowych danych; druga – w większości badań socjologicznych kryterium to nie jest dostępne, bo brak danych o rozkładach badanych cech w populacji. Kryterium to z reguły jest dostępne w odniesieniu do danych metryczkowych, których rozkłady w populacji mogą być uzyskane z badań statystycznych (censusowych).

Z drugiej strony sens pojęciu „wiarygodność” danych uzyskanych w badaniach ankietowych nadają kryteria rzetelności i trafności. Kryteria te zostały przeniesione do metodologii badań ankietowych z teorii testów, która jest metodologicznie najbliższym do badań ankietowych zastosowaniem teorii pomiaru w naukach społecznych i behawioralnych. Kwestia zasadności stosowania ścisłej teorii testów (pomiaru) w odniesieniu do socjologicznych badań ankietowych jest bardzo kontrowersyjna i ze względu na charakter badanych zmiennych, i ze względu na brak możliwości spełnienia podstawowych założeń teorii testów w badaniu ankietowym (por. Sawiński, 1985). Niemniej jednak kryteria wiarygodności danych używane w teorii testów stosowane są, mniej lub bardziej wiernie, w metodologii badań ankietowych.

W myśl tej teorii odpowiedź rzetelna, najprościej mówiąc, to taka odpowiedź, która jest powtarzalna, co znaczy, że niezależna od czynników sytuacyjnych towarzyszących badaniu. Termin „rzetelność” stosowany jako kryterium wiarygodności odpowiedzi uzyskanej w socjologicznym badaniu ankietowym odpowiada terminowi „rzetelność konsensusyjna” stosowanemu w teorii testów (Magnusson, 1981:407). Odpowiedź trafna to natomiast taka, która przynosi informację poszukiwaną przez badacza<sup>6</sup>.

<sup>6</sup> Aby uniknąć nieporozumienia, chcemy podkreślić, że terminologia używana w pracach polskich za Lutyńskim (por. przypis 1) nie odpowiada terminologii zaczerpniętej z teorii testów.

W teorii testów pojęcia te stosowane są do oceny przydatności narzędzia badawczego (testu). Pojęcie „rzetelność” stosowane jest w teorii testów do określenia dokładności, z jaką test mierzy to, co rzeczywiście mierzy (Magnusson, 1981:94)<sup>7</sup>. Pojęcie „trafność” stosowane jest do oceny dokładności, z jaką test mierzy to, co ma mierzyć; inaczej mówiąc, trafność to dokładność, z jaką test mierzy cechy informacji poszukiwane przez badacza (Magnusson, 1981:182)<sup>8</sup>. W metodologii badań ankietowych oba te kryteria stosowane są przede wszystkim do oceny przydatności narzędzia badawczego (w szczególności pytań i skal), ale także do oznaczania pewnych aspektów wiarygodności danych uzyskanych za pomocą technik ankietowych<sup>9</sup>.

Oba te pozornie operacyjne kryteria są bardzo trudne do bezpośredniego zastosowania; uwaga ta dotyczy w szczególności kryterium rzetelności. Zwróćmy uwagę, że odpowiedź rzetelna znaczy dokładnie to samo, co odpowiedź nieobciążona (unbiased). Faktycznie operacyjne zdefiniowanie rzetelności wymaga więc odwołania do kryterium prawdziwości. Trudności z tym związane omówiliśmy powyżej.

Wadą zaprezentowanego wyżej podejścia do zmiennych wyjaśnianych w metodologicznych badaniach wiarygodności danych uzyskanych w badaniach ankietowych jest ich skoncentrowanie na analizie jednostkowych odpowiedzi na jednostkowe pytanie kwestionariusza. Na poziomie próby przyjmuje się za jednostkę analizy rozkłady odpowiedzi na dane pytanie kwestionariusza. W analizach substancywnych w socjologii zaś głównym polem zainteresowań badaczy są związki między zmiennymi, a nie charakterystyki próby (populacji) wyrażone przez czyste rozkłady tych zmiennych. Dystrybucje zmiennych są przedmiotem zainteresowania badaczy w badaniach pollingowych i w badaniach trendów (np. zmian postaw w czasie), a także w badaniach rynkowych. Analizowanie wiarygodności danych na poziomie próby nie wystarcza więc do satysfakcjonującej badacza odpowiedzi na pytanie, w jakim stopniu dane, którymi się posługuje, aby rozwiązać problem badawczy, są wiarygodne; obciążenie danych może powodować nie tylko zmiany siły obser-

---

<sup>7</sup> Rzetelność testu, to proporcja wariancji wyników uzyskanych do wariancji wyników prawdziwych (Magnusson, 1981:94 i dalsze). Pojęcie „rzetelność” odnosi się do zjawisk, które test faktycznie mierzy, niezależnie od tego, co miał mierzyć.

<sup>8</sup> Trafność testu, to statystyczny związek między wynikiem uzyskanym w wyniku zastosowania testu i wynikiem uzyskanym wskutek zastosowania innej, bardziej pewnej metody. Trafność odnosi się do cech, które test miał mierzyć, a nie do cech, które faktycznie mierzy. Rozróżnia się kilka typów trafności: diagnostyczną, prognostyczną, wewnętrzną i teoretyczną (Magnusson, 1981:182 i dalsze).

<sup>9</sup> Wynika to z teoretycznego nastawienia wielu metodologów nauk społecznych, którzy traktują pytanie w kwestionariuszu badania ankietowego jako narzędzie pomiarowe w ścisłym sensie lub sensie bliskim ścisłego (por. Pawłowski, 1977:194-195; Nowak, 1970:102-107).

wowanych związków, ale także ich kierunku. Socjologowie zdają sobie sprawę, że dystrybucje zmiennych są bardzo wrażliwe na obciążenie i rzadko w badaniach substancywnych formułują konkluzje w oparciu o analizy rozkładów; z drugiej strony nie wystarcza aby ustrzec się przed skutkami obciążenia rejestrowanych w badaniach związków między zmiennymi (por. Schuman i Presser, 1981:6–7).

Z tego powodu spotykane w literaturze amerykańskiej probabilistyczne podejście do problematyki obciążenia danych w badaniu ankietowym nie rozwiązuje problemu. Podejście probabilistyczne jest techniką redukcji obciążenia odpowiedzi, stosowaną najczęściej w odniesieniu do pytań drażliwych i osobistych, wykorzystującą motywacje respondentów do szczególnej współpracy z badaczem (Warner, 1965; Horowitz, Shah i Simmons, 1967; Greenberg, 1969; Bradburn i Sudman, 1979). Metoda probabilistycznej redukcji obciążenia odpowiedzi stosowana jest w badaniach ankietowych realizowanych za pomocą wywiadu. Respondent odpowiada na jedno lub dwa pytania wybrane losowo, przy tym ankieter nie wie, na które. W najbardziej ogólnej postaci, respondent odpowiada na jedno z dwóch pytań za pomocą odpowiedzi „tak” lub „nie”:

1. Jestem członkiem grupy A.
2. Nie jestem członkiem grupy A.

Jeżeli znane jest prawdopodobieństwo odpowiedzi na każde z tych pytań, wielkość próby oraz liczba odpowiedzi „tak”, możliwe jest oszacowanie prawdziwej proporcji tych, którzy należą do grupy A.

Metoda probabilistyczna pozwala co najwyżej zredukować obciążenie rozkładów odpowiedzi na dane pytanie, ale brak identyfikacji obciążenia na poziomie jednostkowej odpowiedzi respondenta na jednostkowe pytanie kwestionariusza oraz niemożność zastosowania metody do wszystkich typów pytań stosowanych w socjologicznych badaniach ankietowych wykluczają jej zastosowanie do redukcji obciążenia związków między zmiennymi.

## II. ZMIENNE WYJAŚNIAJĄCE W ANALIZACH WIARYGODNOŚCI DANYCH UZYSKANYCH W BADANIACH ANKIETOWYCH – ŹRÓDŁA OBCIĄŻENIA

W metodologii badań ankietowych źródeł obniżenia wiarygodności (obciążenia) danych poszukuje się w czterech grupach czynników, które w empirycznych analizach metodologicznych pełnią rolę zmiennych niezależnych:

1. Charakterystyki zadania badawczego (task variables). Do tej grupy zalicza się rozmaite charakterystyki narzędzia badawczego i sytuacji badania. W szczególności są to:

- a) formalne charakterystyki narzędzia, to znaczy – kwestionariusza i pytań. W szczególności: kolejność i sekwencja pytań (Schuman i Pres-



ser, 1981: rozdz. II; Schröder i Tennstädt, 1984); długość pytania (Bradburn i Sudman, 1979: rozdz. II; Molenaar, 1982:53–55); struktura pytania – otwarte vs zamknięte (Schuman i Presser, 1981: rozdz. III); struktura możliwych odpowiedzi – kafeteria dwuczłonowa vs wieloczłonowa, obecność pozycji środkowych w kafeteriach (skalach) oraz odpowiedzi „trudno powiedzieć” vs ich brak, pytania logicznie zrównoważone i niezrównoważone (Schuman i Presser, 1981: rozdz. IV, VI i VII; Molenaar, 1982:57–64); występowanie pytań filtrujących lub ich brak (Schuman i Presser, 1981: rozdz. II; McClendon, 1984);

b) treść i sposób sformułowania pytań. W szczególności sformułowanie pytań (Schuman i Presser, 1977); drażliwość pytań (Bradburn i Sudman, 1979: rozdz. I);

c) cechy sytuacji badania, w szczególności: rodzaj zastosowanej techniki badawczej (Groves, 1978; Groves i Kahn, 1979); zastosowanie procedury poinformowanej zgody, ang. informed consent procedure (Singer, 1978; Singer i Frankel, 1982, Lueptow, 1977); wprowadzenie materialnych zachęt (Armstrong, 1975); miejsce przeprowadzenia badania (Lutyńska, 1972); obecność osób trzecich w sytuacji wywiadu (Lutyńska, 1970).

2. Charakterystyki ankierów. Analizowane są dwa rodzaje cech ankierów:

a) cechy nie związane bezpośrednio z rolą ankiera (role-independent albo extra-role characteristics), do których zalicza się:

– cechy społeczno-demograficzne, a w szczególności rasę, płeć, wiek, wykształcenie, pozycję społeczną i religię ankierów;

– podstawowe cechy osobowościowe (basic personality traits), w szczególności extra-vs interowertyzm oraz poczucie emocjonalnego bezpieczeństwa;

– opinie, postawy, przekonania ankiera, które odnoszą się do przedmiotu badania;

– wcześniejsze wobec badania oczekiwania ankiera wobec kształtu odpowiedzi respondentów (pre-survey expectations) (por. Hagenars i Heinen, 1982:94);

b) cechy związane z rolą ankiera (role-restricted characteristics). Zalicza się do nich zachowania ankiera w wywiadzie, a w szczególności: styl przeprowadzania wywiadu, przestrzeganie instrukcji badacza, stosowane techniki sondażu, techniki motywowania i wzmacniania respondentów<sup>10</sup>;

<sup>10</sup> Kahn i Cannell dokonują innego rozróżnienia cech ankiera produkujących obciążenia danych:

a) background factors – kategoria obejmująca cechy społeczno-demograficzne.

b) psychological factors – kategoria obejmująca postawy, oczekiwania, motyw i percepcję.

c) behavioral factors – kategoria obejmująca zachowania ankiera nieadekwatne do roli (Kahn i Cannell, 1957:179-193).

3. Cechy respondentów niezależnie od roli, na przykład: cechy społeczno-demograficzne, cechy osobowościowe, takie jak pamięć i poziom intelektualny oraz cechy respondentów bezpośrednio związane z rolą, czyli zachowania i emocje w trakcie badania, a w szczególności lęk i poczucie zagrożenia w sytuacji badania;

4. Różne interakcje czynników wymienionych w grupach 1, 2 i 3, a więc interakcje między cechami zadania badawczego, cechami ankietera i cechami respondenta (por. De Lamater, 1982: 30–37).

Do określenia związku między cechami ankietera a obciążeniem danych (lub szerzej: zachowaniem respondenta w sytuacji wywiadu) stosowany jest termin „efekt ankieterski”. Tradycja terminologiczna w badaniach nad efektem ankieterskim związana jest niewątpliwie z Hymanem (Hyman, 1954:226–228 i 321–322); w literaturze ten consensus terminologiczny traktuje się jako oczywisty, co rodzi szereg niejasności w odczytywaniu znaczenia nadawanego temu pojęciu przez różnych autorów. Terminologia Hymana używana bywa w sposób daleko odbiegający od precyzyjnych definicji pierwotnych<sup>11</sup>.

Najogólniej mówiąc, efekt ankieterski definiowany jest jako wpływ cech ankietera na odpowiedzi respondentów, przy czym mówi się o efekcie ankieterskim, gdy wpływ ten jest istotny statystycznie. Hyman tak definiuje różne rodzaje efektu ankieterskiego:

Pochodzące od ankieterów zniekształcenie (interviewer distortion) istnieje zawsze wtedy, gdy pojawia się jakiegokolwiek odchylenie od prawdziwej odpowiedzi w odpowiedziach uzyskanych i zapisanych przez ankietera od danego respondenta na dane pytanie. Efekt ankieterski brutto w całym badaniu może być zdefiniowany jako funkcja ogólnej liczby takich indywidualnych odchyień [od wartości „prawdziwej” – K.S.] [...] (Hyman, 1954:226).

Efekty netto mogą być zdefiniowane jako różnice między rozkładem odpowiedzi otrzymanym przez jednego lub więcej ankieterów na jedno lub więcej pytań od danej populacji respondentów a „prawdziwą” dystrybucją odpowiedzi na to pytanie lub pytania w populacji. (Hyman, 1954:228).

Definicja odchylenia między ankieterami (inter-interviewer deviation) oparta jest na pojęciu potencjalnie nieskończonego universum ankieterów. Każdy z tych ankieterów w danych warunkach otrzymałby szczególny rozkład odpowiedzi, gdyby przeprowadził wywiad z każdym respondentem w universum. Zróznicowanie między ankieterami (inter-interviewer variation) pomyślane jest jako wariancja takich pojedynczych rozkładów uzyskanych przez pojedynczych ankieterów w łącznym rozkładzie odpowiedzi uzyskanych przez wszystkich ankieterów. (Hyman, 1954:321).

[...] W sumie wariancja ankieterska (interviewer variance) odpowiada błędowi (error) wokół „wartości oczekiwanej” dla wszystkich ankieterów, podczas gdy obciążenie pochodzące od ankieterów (interviewer bias) netto odpowiada odchyleniu tej wartości oczekiwanej od wartości prawdziwej w populacji”. (Hyman, 1954:322).

<sup>11</sup> W polskiej literaturze pojawia się niezbyt jasne w świetle terminologii Hymana rozróżnienie między efektem ankieterskim a wpływem, które oparte jest na odmiennej teoretycznie perspektywie; rozróżnienie to jest konsekwencją pewnej specyficznej teorii interakcji między ankieterem i respondentem (Lutyńska, 1978:147-148).

Jak łatwo zauważyć, w definicjach Hymana pojawia się jeszcze jedno, w porównaniu z zaprezentowanymi przez nas powyżej, kryterium prawdziwości odpowiedzi. Oprócz kryterium, które my nazwaliśmy statystycznym, przy zastosowaniu którego wartość prawdziwa to wartość z populacji (efekt ankietarski netto), Hyman stosuje inne jeszcze kryterium statystyczne (wariancja ankietarska). Przy zastosowaniu tego kryterium wartość prawdziwa to wartość oczekiwana, czyli średnia wartość uzyskana w danym badaniu.

Badania efektu ankietarskiego są w metodologii badań ankietowych nadal bardzo popularne; nowe rezultaty wskazują jednak, że mimo bogatej tradycji tych badań mechanizm efektu ankietarskiego nadal nie jest dobrze zrozumiany. Nowsze analizy dowodzą, że udział analizowanych, niezależnych od roli cech ankietarów, w wariancji odpowiedzi nie przekracza 5–10% (por. Sudman i Bradburn, 1974; Sudman, Bradburn, Blair i Stocking, 1977; Bradburn i Sudman, 1979). Za istotne ze względu na wpływ, jaki wywierają na odpowiedzi respondentów uważa się tylko dwie niezależne od roli cechy ankietarów: płeć i rasę<sup>12</sup>. Poza tymi dwoma wyjątkami efekt niezależnych od roli cech ankietarów w nowszej literaturze uważa się za znikomy (Alwin, 1977:16). Wskazuje się jednak na konieczność poszukiwania efektu cech ankietera bezpośrednio związanych z rolą (Bradburn i Sudman, 1979:50; Brenner, 1982:134–135). Niektórzy autorzy zwracają uwagę na metodologiczne niejasności w dotychczasowych studiach kwestionujących wagę efektu ankietarskiego dla obniżania wiarygodności danych (Hagenaars i Heinen, 1982; Gove, 1982).

W nowszych analizach zwraca się uwagę, że spośród cech respondentów szczególną rolę w kształtowaniu obciążenia odpowiedzi odgrywają cechy respondentów nie związane bezpośrednio z rolą: płeć, rasa i poziom wykształcenia (DeLamater, 1982:35–37; Sudman i Bradburn, 1974). Gdyby ten rezultat przyjąć za ostateczny, należałoby zweryfikować obszerny wątek analiz teoretycznych, w którym szczególną wagę przywiązuje się do wpływu zachowań i emocji respondenta w sytuacji badania na obciążenie odpowiedzi.

Współcześnie autorzy przypisują znacznie poważniejszą rolę efektowi cech zadania badawczego (task variables effect), niż efektowi ankietarskiemu i efektowi cech respondenta. Najważniejsze rezultaty w tym nurcie badań uzyskali Schuman i Presser (1981). Podkreślić jednak należy, że najpłodniejszy wydaje się w metodologii ten nurt badań, który bada efekt charakterystyk zadania badawczego, cech ankietera i cech

---

<sup>12</sup> Badania Hatchetta i Schumana wskazują, że nie mniej, niż 26% wariancji w Indeksie Liberalizmu Rasowego wyjaśnione jest przez rasę ankietera; efekt płci ankietera zaobserwowała Jastrząb (Hatchett i Schuman, 1975; Jastrząb, 1966). Jest jednak oczywiste, że w badaniach Hatchetta i Schumana efekt rasy ankietera jest w rzeczywistości efektem interakcji cechy ankietera i cechy respondenta.

respondenta w interakcji ze sobą (por. DeLamater, 1982). Ten nurt badań ma bardzo silne uzasadnienie w teorii badań ankietowych i należy oczekiwać, że stanie się podstawowym polem analiz wiarygodności danych uzyskanych w badaniu ankietowym (Bradburn i Sudman, 1979; Dijkstra i van der Zouwen, 1982). Efekt cech zadania badawczego, cech ankietera i cech respondenta rozpatrywanych w interakcji może wyjaśnić w przyszłości pewne zaskakujące rezultaty uzyskane, na przykład w badaniach efektu ankieterskiego i badaniach efektu cech respondenta; rezultaty, które zdają się kwestionować metodologiczne rozprawy teoretyczne, pełniące rolę prac klasycznych. Badanie wpływu cech zadania badawczego, cech respondenta i cech ankietera w interakcji jest jednak metodologicznie znacznie bardziej skomplikowane, niż badanie każdego z tych efektów osobno mimo zaawansowanych statystycznie technik, stosowanych dziś w metodologii badań ankietowych.

Obserwując stan badań nad wiarygodnością danych uzyskanych w badaniu ankietowym nie sposób nie zauważyć, że mimo wieloletniej tradycji tych badań, mechanizm powstawania obciążenia danych daleki jest od empirycznego zrozumienia. Techniki badania wiarygodności danych ankietowych stosowane w kontrolowanych badaniach metodologicznych nie pozwoliły dotąd na wykrycie wszelkich źródeł obciążenia danych i nie mogą być na szeroką skalę stosowane w typowych badaniach substancywnych. Sądzymy więc, że przez długi jeszcze czas socjolog pozostanie bezradny wobec swojej niepewności na temat wiarygodności uzyskanych przez siebie danych i pewności formułowanych na podstawie badań konkluzji.

## Literatura cytowana:

- Alwin, D. F. (1977), *Making Errors in Survey: An Overview*, [w:] Alwin D. F. (red.) *Survey Design and Analysis, Current Issues*. Sage, Beverly Hills 1978:7–26.
- Armstrong J. S. (1975), *Monetary Incentives in Mail Survey*, „Public Opinion Quarterly”, 39:111–116.
- Bradburn N. M. Sudman S., (1979), *Improving Interview Method and Questionnaire Design*, San Francisco.
- Branner M. (1982), *Response Effects of „Role-restricted” Characteristics of Interviewer* [w:] Dijkstra W., van der Zouwen I. (red.), *Response Behaviour in the Survey-interview*, New York:131–165.
- Daniłowicz P., Sztabiński P. (1983), *Wiarygodność informacji uzyskiwanych w odpowiedzi na pytania metryczkowe*, [w:] Lutyńska K., Wejland A. P. (red.), *Wywiad kwestionariuszowy*, Wrocław: 309–319.
- Daniłowicz P., Sztabiński P., (1986), *Zmienne osobowe w badaniach socjologicznych. Analiza weryfikacyjne. Analizy i próby technik badawczych w socjologii*, t. VI., Wrocław.
- Dean J. P. Whyte W. F., (1958), *How do you know if the informant is telling the truth?* [w:] Bynner J., Stribley K. M. (red.), *Social Research: Principles and Procedures*, London 1979:179–188.
- De Lamater J., (1982), *Response-effects of Question Content*, [w:] Dijkstra W., van der Zouwen J. (red.), *Response Behaviour in the Survey-interview*, New York: 13–48.
- Dijkstra W., van der Zouwen J., (1982), *Response Behaviour in the Survey-interview*, New York.
- Gorden R., (1975), *Interviewing: strategy, techniques and tactics*, wydanie zmienione, Homewood, Illinois.
- Gove W. R. (1982), *Systematic Response Bias and Characteristics of the Respondent*, [w:] Dijkstra W., van der Zouwen J. (red.), *Response Behaviour in the Survey-interview*, New York:167–187.
- Greenberg B. in., (1969), *The Unrelated Question Randomized Response Model: Theoretical Framework*, „Journal of the American Statistical Association”, 64:520–539.
- Groves R. M., (1979), *Actors and Questions in Telephone and Personal Interview*, „Public Opinion Quarterly”, 43:190–205.
- Groves R. M., Kahn R.L., (1979), *Surveys by Telephone: A National Comparison with Personal Interviews*, New York.
- Hagenaars J. A., Heinen T. G., (1982), *Effects of Role-independent Interviewer Characteristics on Responses*, [w:] Dijkstra W., van der Zouwen J. (red.), *Response Behaviour in the Survey-interview*, New York: 91–130.
- Hatchett Sh., Schuman H., (1975), *White Respondents and Race-of-Interviewer Effects*, „Public Opinion Quarterly”, 39:523–528.
- Horowitz D. G., Shah B. V., Simmons W. R., (1967), *The Unrelated Question Randomized Response Model*, „Proceedings of the American Statistical Association”, Washington D. C.
- Hyman H. i in., (1954), *Interviewing in Social Research*, Chicago.
- Jastrząb M., (1966), *Wpływ płci ankietera na wyniki uzyskane w badaniach*, [w:] *Analizy i próby technik badawczych w socjologii*, t. I: 147–185.
- Kahn R. L., Cannell Ch. F., (1957), *The Dynamics of Interviewing*. John Wiley and Sons, New York.
- Lueptow L. i in., (1977), *The Impact of Informed Consent Regulations on Response Rate and Response Bias*, [w:] Alwin D.F. (red.), *Survey Design and Analysis, Current Issues*. Sage, Beverly Hills 1978:59–80.
- Lutyński J., (1972), *Analiza procesu otrzymywania informacji w badaniach z zastosowaniem wywiadu kwestionariuszowego*, [w:] *Analizy i próby technik badawczych w socjologii*, t. 4.

- Lutyński J., (1978), *Koncepcje pytania kwestionariuszowego, ich zastosowanie w badaniach, konsekwencje i możliwości*, „Przegląd Socjologiczny”, t. XXX:9–51.
- Lutyński J., (1979), *Pytanie jako narzędzie w surveyowych badaniach socjologicznych*, „Studia Socjologiczne”, 2(73).
- Lutyńska K., (1970), *Osoby trzecie w wywiadzie socjologicznym i ich wpływ na odpowiedzi respondentów*, [w:] *Analizy i próby technik badawczych w socjologii*, t. 3:133–222.
- Lutyńska K., (1972), *Miejsce przeprowadzania wywiadu i jego oddziaływanie*, [w:] *Analizy i próby technik badawczych w socjologii*, t. IV.
- Lutyńska K., (1978), *Ankieterzy i badacze. Z badań nad wpływem ankieterskim*, „Przegląd Socjologiczny”, t. XXX:143–173.
- Mangusson D., (1969), *Wprowadzanie do teorii testów*, Warszawa 1981.
- McClandon, McKee J., (1983), *The Effects of Opinion Filters on Don't Know Responses, Substantive Responses and Attitude Strength*. Referat prezentowany na „Annual Meeting of the American Sociological Association”, San Antonio, Texas, 1984.
- Molenaar N. J., (1982), *Response-effects of „Formal” Characteristics of Questions* [w:] Dijkstra W., van der Zouwen J. (red.) *Response Behaviour in Survey-interview*, New York: 49–89.
- Nowak S., (1965), *Funkcje wskaźnikowe odpowiedzi w badaniach ankietowych*, [w:] *Studia z metodologii nauk społecznych*, Warszawa: 287–366.
- Nowak S., (1970), *Metodologia badań socjologicznych*, Warszawa.
- Pawłowski T., (1977), *Pojęcia i metody współczesnej humanistyki*, Wrocław.
- Sawiński Z., (1985), *Błędy pomiaru w badaniach procesu osiągnięć*. (maszynopis).
- Schröder S., Tennstädt F., (1984), *Towards a Theory of „How People Answer Questions”*, Referat prezentowany na „Annual Meeting of the American Sociological Association”, San Antonio, Texas, 1984.
- Schuman H., Presser S., (1977), *Question Wording as an Independent Variable in Survey Analysis*, [w:] Alwin D.F. (red.) *Survey Design and Analysis, Current Issues*. Sage, Beverly Hills 1978:27–46.
- Schuman H., Presser S., (1981), *Questions and Answers in Attitude Survey*, New York.
- Singer E., (1978), *Informed Consent: Consequences for Response Rate and Response Quality in Social Survey*, „American Sociological Review”, 43:144–162.
- Singer E., Frankel M.R., (1982), *Informed Consent Procedures in Telephone Interviews*, „American Sociological Review”, 47:416–427.
- Słomczyński K.M., (1966), *Wpływ oficjalnej i prywatnej sytuacji wywiadu na odpowiedzi respondentów w środowisku inteligencji*, [w:] *Analizy i próby technik badawczych w socjologii*, t. 1:186–224.
- Sudman S., Bradburn N.M., (1974), *Response Effects in Surveys: A Review and Synthesis*, Aldine, Chicago.
- Sudman S., Bradburn N.M., Blair E., Stocking C., (1977), *Modest Expectations: The Effects of the Interviewers' Prior Expectations on Responses*, [w:] Alwin D. F. (red.) *Survey Design and Analysis, Current Issues*. Sage, Beverly Hills 1978:47–58.
- Warner S. L., (1965), *Randomized Response: A Survey Technique for Eliminating Error Answer Bias*, „Journal of the American Statistical Association”, 60:63–69.
- Warwick D. F., Lininger Ch. A., (1975), *The Sample Survey: Theory and Practice*, New York.