

Franciszek Sztabiński  
Polska Akademia Nauk

## CO WIEMY O RESPONDENCIE? ROZMIAR I ŹRÓDŁA BŁĘDÓW W POMIARZE CECH METRYCZKOWYCH

*Należy zauważyć, że wykazanie błędu wskazuje na  
zaawansowany stan nauki. Każde badanie jest  
obarczone błędem i daleko lepiej jest mieć jego świadomość (...),  
niż być ignorantem w zakresie błędów obecnych w danych.  
Ignorowanie błędu nie oznacza jego braku.*

Herbert H. Hyman<sup>1</sup>

Pytania o cechy demograficzne i społeczne respondenta (tzw. „metryczka”) stanowią stałą i integralną część każdego kwestionariusza we wszystkich badaniach ilościowych. Służą one nie tylko do scharakteryzowania badanej zbiorowości, lecz także pozwalają opisać jej zróżnicowanie. Jednocześnie tak ważne informacje obarczone są zazwyczaj błędem. Celem artykułu jest określenie rozmiaru tego błędu i ich źródeł. Błąd traktowany jest jako brak replikowalności pomiaru, a więc niezgodność odpowiedzi respondenta na to samo pytanie zadane w trakcie powtórnego pomiaru (w badaniu kontrolnym). Analizą objęte zostały materiały z kontroli terenowej projektów badawczych realizowanych w okresie od 1997 do 2000 roku (w sumie 7029 losowo dobranych wywiadów kontrolnych z 83 badań). Na podstawie analizy można sformułować wniosek o powszechności błędu pomiaru cech społeczno-demograficznych. Jego wielkość zależy jednak od rodzaju cechy. Podstawowe źródła błędów w pomiarze zmiennych metryczkowych to: pamięć respondenta, waloryzacja społeczna oraz obiektywny brak wiedzy. Uzyskane wyniki są prezentowane na tle wyników innych badań (w tym tzw. badania weryfikacyjne), w których ocena danych formułowana jest na podstawie ich porównania z danymi dokumentalnymi.

### Znaczenie pytań metryczkowych

Pytania o cechy demograficzne i społeczne respondenta (tzw. „metryczka”) stanowią stałą i integralną część każdego kwestionariusza do wszystkich badań ilościowych. Służą one nie tylko do scharakteryzowania badanej zbiorowości, lecz tak-

---

Uwagi do Autora lub prośbę o nadbitki prosimy kierować: Franciszek Sztabiński, Instytut Filozofii i Socjologii PAN, ul. Nowy Świat 72, 00-330 Warszawa, e-mail: fsztabin@ifispan.waw.pl

<sup>1</sup> Hyman 1954: 4.

że pozwalają opisać jej zróżnicowanie (a więc określić, czym różnią się postawy, opinie i zachowania kobiet i mężczyzn, osób starszych i młodych, o różnym wykształceniu, różnych dochodach itd.). Co więcej, przyjmuje się, że pewne cechy demograficzne (np. płeć, wiek) są „pierwotnymi ogniwami” warunkującymi inne cechy społeczne, ale także poglądy i przekonania. Znajduje to zresztą odzwierciedlenie w stosowanej terminologii: mówi się bowiem często o „zmiennych zależnych” i „zmiennych niezależnych” (właśnie ze względu na nie analizuje się te pierwsze).

Zestaw pytań zawartych w metryczce, jest zawsze pochodną trzech czynników: typu i problematyki badania, badanej populacji oraz przewidywanego planu analizy (postawionych hipotez badawczych). Przykładowo, jeśli przedmiotem naszego badania jest użytkowanie i kupowanie określonego produktu, to oczywiste jest, iż chcemy dowiedzieć się jacy ludzie, o jakich cechach używają tego produktu, a jacy go nie używają, ale także, jacy ludzie, o jakich cechach kupują ten produkt, a jacy nie (nie każdy bowiem, kto np. jada jogurty, musi je kupować i odwrotnie – nie każdy kto je kupuje, jada je). Jeśli tak określimy nasze zapotrzebowanie informacyjne, to rzeczą oczywistą jest, iż musimy pytać o cechy użytkowników (nieużytkowników) produktu, ale również o cechy osób, które produkt ten kupują (nie kupują go). Z kolei w badaniach opinii istotną zmienną metryczkową są poglądy polityczne respondenta oraz jego przynależność organizacyjna. W badaniach społecznych o charakterze akademickim istotnymi cechami mogą być przykładowo: środowisko, w którym respondent wychowywał się, a także zawód i wykształcenie ojca respondenta.

Podobnie charakter badanej zbiorowości również w dużej mierze określa zawartość „metryczki”. Jeśli badanie prowadzimy na próbie rolników indywidualnych, ważnymi cechami są: wielkość gospodarstwa, rodzaj produkcji rolnej, wyposażenie w maszyny itp. W badaniu instytucji – na przykład sklepów – cechami takimi są: powierzchnia sprzedaży, liczba pracowników, obrót itp., ale także stanowisko i staż pracy (w zawodzie i w danej placówce) naszego rozmówcy.

Jeśli natomiast jedną z naszych hipotez jest to, iż wśród podstawowych czynników określających orientacje życiowe ludzi uwzględnić należy znak zodiaku, to w metryczce musimy zapytać respondentów o dokładną datę urodzenia (dzień, miesiąc i rok), nie zaś tylko o ich wiek<sup>2</sup>.

Pytania metryczkowe pełnią zatem rolę kluczową: braki w tej części kwestionariusza powodują, że nie jest możliwa analiza (np. ze względu na płeć, wiek, wykształcenie itd.) odpowiedzi respondentów na pytania zamieszczone w zasa-

---

<sup>2</sup> O znaczeniu i funkcjach zmiennych społeczno-demograficznych piszą m.in. Daniłowicz i Sztabiński 1986: 15–20.

dnicznej jego części (tzw. merytorycznej). Jednocześnie musimy mieć świadomość, iż tak ważne informacje obarczone są zazwyczaj błędem i to często większym niż informacje merytoryczne (Sztabiński F. 1997: 104–118). Cechą pytań metryczkowych jest to, iż ich przedmiotem są zazwyczaj wydarzenia przeszłe (np. data urodzenia lub wiek poszczególnych członków gospodarstwa domowego, historia zawodowa respondenta, przebieg edukacji, dochody w określonym przedziale czasu) oraz fakty będące ich konsekwencją lub pochodną (np. stan cywilny respondenta, wykształcenie jego lub współmałżonka). Innymi słowy, w badaniach surveyowych respondenci bardzo często są pytani nie tylko o to, co się w ich życiu wydarzyło, ale także, kiedy to miało miejsce. Są to zatem pytania retrospektywne, wymagające od respondenta odwołania się do pamięci. Jednocześnie od dawna wiadomo, że odpowiadając na pytania retrospektywne respondenci popełniają błędy: zapominają, że pewne wydarzenia w ogóle miały miejsce lub też niektórzy z nich przypominają sobie wydarzenia, które w ogóle nie miały miejsca (ta ostatnia sytuacja jest znacznie rzadsza). Są to tzw. błędy „pominięcia” i „popelnienia”. Relatywnie najczęstsze są jednak sytuacje, w których respondenci pamiętają, iż dane wydarzenie miało miejsce, ale w sposób błędny umiejscawiają je w czasie (tzw. efekt teleskopowy). Błędne umiejscowienie w czasie danego wydarzenia może polegać na wydłużaniu dystansu czasowego, a więc raportowaniu, iż miało ono miejsce wcześniej niż faktycznie (wsteczny efekt teleskopowy) – tzw. *backward telescoping* lub skracaniu dystansu czasowego (kompresja) – raportowaniu, iż miało ono miejsce później niż faktycznie (uprzedni efekt teleskopowy) – tzw. *forward telescoping* (zob. Auriat 1996: 24–33, Gaskel, Wright i O’Muircheartaigh 2000: 77).

Skoro pytania metryczkowe obarczone są błędem, którego istnienie sobie uświadamiamy, naturalnym jest pytanie o rozmiar tego błędu w przypadku pomiaru poszczególnych cech. Jest to zatem pytanie o to, co naprawdę wiemy o respondencie.

## **Błąd pomiaru zmiennych metryczkowych**

Charakterystyczną cechą pytań metryczkowych jest fakt, jak wspomniano, iż w większości przypadków odwołują się one do wydarzeń przeszłych (np. daty narodzin, ślubu, posiadanie dzieci), których zapamiętywanie różni się od zapamiętywania innych wydarzeń – codziennych (np. zakupów, wizyty u lekarza). Są to wydarzenia, których się nie zapomina, poza przypadkami klinicznymi: amnezją, chorobą Alzheimera. Dlatego też podstawowym problemem w przypadku

tego typu danych jest raportowanie dat (Poulain, Riandey i Firdion 1991: 65–88). W jaki sposób można jednak określić, czy informacje metryczkowe uzyskiwane od respondentów są obciążone błędem, czy nie? Są dwa podstawowe sposoby: jeden to powtórne zadanie tych samych pytań, drugi natomiast to odwołanie się do danych dokumentalnych.

Idea „powtórnego pomiaru” wywodzi się w prostej linii z nauk przyrodniczych, gdzie jedną z cech charakterystycznych jest replikowalność wyników: wynik uznaje się za pewny, jeśli pomiary przeprowadzone w tych samych warunkach dają te same rezultaty (Heisenberg 1975: 55). Z uwagi na pozytywistyczne korzenie badań ilościowych często przyjmuje się, że replikowalność jest koniecznym ich elementem. Służyć ma ona określeniu błędu badania: pomiar jest rzetelny, jeśli jest powtarzalny<sup>3</sup>. Aby zatem ocenić rzetelność pomiaru, pomiar ten musimy powtórzyć (Bryman 1995: 37–38, Jones 2000: 28–29). W badaniach ilościowych w większości przypadków powtórzony pomiar sprowadza się do powtórnego wywiadu z tą samą osobą (*reinterview*), tzn. powtórnego zadania respondentom, w trakcie wywiadu kontrolnego, wybranych pytań z wywiadu zasadniczego<sup>4</sup>. Umożliwia to określenie szacunkowej rzetelności pomiaru, a więc jego ewentualny błąd. Mówiąc inaczej, jeśli pytamy dwa razy o to samo (jeden raz w badaniu zasadniczym, drugi raz w wywiadzie kontrolnym) i uzyskujemy rozbieżne odpowiedzi, to wartość tych informacji staje się problematyczna.

Drugim sposobem oceny jest odwołanie się do danych istniejących. W przypadku cech metryczkowych są to zazwyczaj istniejące dane dokumentalne. Przyjmuje się, iż dane takie są w pełni wiarygodne, ponieważ pochodzą z oficjalnych rejestrów takich instytucji, jak urzędy stanu cywilnego, kancelarie kościelne, działy spraw osobowych w zakładach pracy, ośrodki pomocy społecznej, wydziały komunikacji itd. Jedną z wcześniejszych prób określenia dokładności pomiaru pewnych zmiennych faktualnych było tzw. „Denver Validity Survey” – badanie przeprowadzone w 1949 roku przez Opinion Research Center Uniwersytetu Denver. Dotyczyło ono m.in. udziału respondentów w wyborach municy-

---

<sup>3</sup> Przez rzetelność uzyskanych informacji rozumie się bowiem ich zgodność w pomiarach, to znaczy w sytuacji, gdy o te same fakty lub opinie pytamy niezależnie po raz drugi (Goode i Hatt 1952).

<sup>4</sup> Procedura powtórnego wywiadu ma na celu nie tylko ocenę rzetelności pomiaru, ale generalnie ocenę jakości rezultatów badania. Jej celem jest bowiem określenie, między innymi, rozmiaru błędów zawartości (*content errors*). W Stanach Zjednoczonych już w latach 50. podjęto zakrojony na szeroką skalę program oceny jakości danych z badań realizowanych przez U.S. Bureau of the Census (Bailar 1968: 41–63). O różnych możliwościach wykorzystania powtórnego wywiadu piszą także: Biemer i Forsman 1992: 915–923, Forsman i Schreiner 1991: 279 i nn.

palnych, posiadania przez nich karty bibliotecznej, prawa jazdy oraz telefonu w miejscu zamieszkania (Parry i Crossley 1950: 61-80). Uzyskane deklaracje respondentów były następnie porównywane z danymi dokumentalnymi na poziomie jednostkowym: stąd też mówiono o trafności jednostkowej odpowiedzi (*individual validity*)<sup>5</sup>. Należy podkreślić, iż wskazany powyżej schemat oceny rezultatów był i jest nadal wielokrotnie wykorzystywany w badaniach socjologicznych i demograficznych (zob. m.in.: Auriat 1993, Clausen 1968-69, Miller i Groves 1985, Poulain, Riandey i Firdion 1991, Weaver 2000, Weiss 1968-69).

Mówiąc o ocenie odpowiedzi respondentów na pytania faktualne, dokonywanej na podstawie porównania z danymi dokumentalnymi należy wspomnieć także o tego typu badaniach realizowanych w Polsce. Dotyczyły one nie tylko zmiennych osobowych (Daniłowicz i Sztabiński 1986, Koniarek 1975: 525-548), danych społeczno-demograficznych (Sztaudynger-Kaliszewicz 1966: 257-281, Lutyńska 1975: 117-188), ale także zachowań zdrowotnych (Lutyńska i Szeszenia 1972: 329-354, Sawicki, Wysocki, Twardowska i Obrzut 1969: 25-31) oraz losów osób zwolnionych z zakładów karnych (Pelka-Sługocka i Sługocki 1975: 221-258). Większość z powyżej wskazanych badań była realizowana w ramach projektu analiz weryfikacyjnych danych uzyskiwanych przy zastosowaniu wywiadu kwestionariuszowego, którego autorem był J. Lutyński (1975: 319-392)<sup>6</sup>.

Poniżej przedstawimy wyniki analizy błędu pomiaru zmiennych metryczkowych. W analizach tych błąd określany jest na podstawie oceny zgodności odpowiedzi na wybrane pytania metryczkowe, które uzyskano w trakcie powtórnego pomiaru, realizowanego w badaniach kontrolnych (w pewnym sensie jest to zatem ocena rzetelności odpowiedzi)<sup>7</sup>. Błąd traktowany jest tutaj zatem jako

---

<sup>5</sup> Badanie „Denver Validity Survey” stało się klasycznym: do niego i uzyskanych rezultatów odwoływano się wielokrotnie. Swoistym ewenementem jest fakt, iż jeszcze 20 lat po przeprowadzeniu badania ukazał się artykuł bazujący na jego wynikach (Cahalan 1968-69: 607-621).

<sup>6</sup> Autor ten weryfikacją nazywa analizy zmierzające do sformułowania jakiegokolwiek oceny wiarygodności wyników liczbowych (Lutyński 1975: 325) i wyróżnia dwa podstawowe jej typy. Pierwszy to weryfikacja wewnętrzna, w której jednostkowe informacje i wynik w całości ocenia się na podstawie danych o procesach, których rezultaty stanowią otrzymane informacje (np. kontrola pracy ankietatorów). Drugi typ, to weryfikacja zewnętrzna, w której ocenę wyniku ustala się na podstawie zestawienia rezultatów osiągniętych różnymi drogami (w tym także dane dokumentalne). W tym ujęciu, zarówno powtórzony pomiar (ocena rzetelności), jak i porównanie z istniejącymi danymi (ocena trafności) zaliczyć należy do weryfikacji wewnętrznej, aczkolwiek różnych jej typów.

<sup>7</sup> Mówiąc o badaniach kontrolnych mamy na myśli badania umożliwiające pośrednią i bezpośrednią ocenę materiałów zebranych w badaniach zasadniczych, tzn. ocenę pracy ankietera i wartości informacji określonej poprzez powtórny pomiar (Sztabiński F. 1999: 71-78). W rozumieniu J. Lutyńskiego jest to zatem weryfikacja wewnętrzna i zewnętrzna (por. przyp. 6).

brak replikowalności pomiaru, a więc niezgodność odpowiedzi respondenta na to samo pytanie zadane w trakcie wywiadu zasadniczego i kontrolnego.

W przypadku niektórych zmiennych, takich jak: wiek, wykształcenie respondenta, liczba osób i skład gospodarstwa domowego, analizy te prowadzone będą nie tylko na podstawie materiałów z badań „jednorazowych”, ale także dwóch badań o charakterze ciągłym (umożliwi to zaobserwowanie pewnych tendencji, jeśli takie występują). Uzyskane wyniki zostaną przedstawione na tle wyników analiz weryfikacyjnych, prowadzonych przy wykorzystaniu danych dokumentalnych.

Celem analiz jest wyłącznie określenie rozmiaru błędu w pomiarze cech metryczkowych i ich potencjalnych źródeł. Oznacza to, iż nie będziemy analizować konsekwencji posługiwania się danymi nie w pełni wiarygodnymi dla wyników badań. Inaczej mówiąc, przedmiotem naszych zainteresowań nie jest analiza wpływu błędów na korelacje, a jedynie wskazanie istnienia błędu w każdym badaniu (jeśli tak jest) oraz jego rozmiaru.

To, iż pytania metryczkowe znajdują się w każdym kwestionariuszu nie oznacza, że ich zestaw jest za każdym razem identyczny. Jak wspomniano, zależy to od kilku czynników, niemniej wskazać można pewne stałe pytania, które zadawane są zawsze, w każdym badaniu. W przypadku cech indywidualnych respondenta są to pytania o: płeć, wiek, wykształcenie i sytuację zawodową. Z kolei w przypadku cech kontekstowych, pytania o: liczbę osób w gospodarstwie domowym (ewentualnie skład gospodarstwa domowego), dochody i wyposażenie materialne gospodarstwa (por. Daniłowicz i Sztabiński 1986: 15–20). Spośród wskazanych powyżej, w naszych analizach uwzględnimy praktycznie wszystkie (z wyjątkiem „dochodów gospodarstwa domowego”), aczkolwiek nie w jednakowym zakresie. Wynika to z przedmiotu poszczególnych badań kontrolnych – nie w każdym z nich zadawano powtórnie wszystkie pytania metryczkowe. Przykładowo, dotyczy to informacji o składzie gospodarstwa domowego (chodzi o płeć i wiek poszczególnych jego członków) oraz wyposażenia materialnego gospodarstwa (dodatkowo, pytania te, jeśli były powtarzane, nie zawsze dotyczyły identycznego zestawu dóbr).

Analizą obejmujemy materiały z kontroli terenowej pięciu projektów badawczych, w tym dwóch „ciągłych”, realizowanych w okresie od 1997 do 2000 roku. Próby do badań kontrolnych liczyły od 150 do 240 losowo dobranych przypadków, a w badaniach „ciągłych” od 125 do 517 kwartalnie. W sumie analizie poddano 7029 wywiadów kontrolnych z 83 badań, zrealizowanych w ramach 5 projek-

---

<sup>8</sup> Analizowane materiały zgromadzone zostały w trakcie badań kontrolnych, których przedmiotem była ocena realizacji terenowej projektów prowadzonych przez różne instytucje badawcze, zarówno marketingowe, jak i akademickie. Autor pragnie podziękować Kierow-

tów<sup>8</sup>. Wszystkie badania kontrolne realizowane były przy zastosowaniu techniki wywiadu kwestionariuszowego. Dopuszczono dwie formy realizacji wywiadu kontrolnego: bezpośrednią (jako wywiad osobisty) oraz pośrednią (jako wywiad telefoniczny). Uprzednie doświadczenia z badań kontrolnych wskazują bowiem, że wartość uzyskiwanych za ich pomocą informacji jest porównywalna (Sztabiński F. 1995: 49–60). W poszczególnych badaniach udział kontrolnych wywiadów telefonicznych nie przekraczał jednak 10% ogółu kontrolowanych.

## **Rozmiar błędu pomiaru zmiennych metryczkowych - prezentacja wyników**

Poziom zgodności odpowiedzi na poszczególne pytania metryczkowe w pomiarze powtórzonym przedstawia tabela 1 (zob. s. 98).

Obecnie omówimy wyniki porównania dla każdej ze zmiennych.

A. Płeć respondenta. Płeć respondenta jest zmienną tzw. indywidualną, występującą w każdym badaniu, niezależnie od jego celu i przedmiotu. Wydawać by się mogło, że ze względu na obserwowalny charakter tej cechy nie jest możliwy błąd pomiaru. Niestety nie jest tak. W jednym spośród analizowanych badań błąd taki wystąpił i stwierdzono go w 1,0% wywiadów (2 przypadki). W pozostałych badaniach nie zaobserwowano przypadków niezgodności tej cechy. Tak więc maksymalny odsetek odpowiedzi niezgodnych w powtórzonym pomiarze wynosi 1,0%, natomiast średni odsetek zgodności w analizowanych badaniach – 99,9%.

B. Wiek respondenta. Drugą zmienną indywidualną, występującą również w każdym badaniu, jest wiek respondenta. Błąd pomiaru tej cechy wystąpił we wszystkich analizowanych badaniach. Maksymalny odsetek niezgodności w powtórzonym pomiarze wynosi 39,5%, minimalny – 1,1%, natomiast średni odsetek zgodności – 93,9%<sup>9</sup>. Warto zwrócić uwagę na niewielkie różnice odsetka od-

---

nictwu firmy EPI M.A. za udostępnienie materiałów z realizowanych przez nią kontroli terenowych oraz Kierownictwu Instytutu Badania Rynku i Opinii Publicznej SMG/KRC Poland Media, na zlecenie którego realizowana była znaczna część omawianych badań.

<sup>9</sup> Tak wysoki maksymalny odsetek odpowiedzi niezgodnych (prawie 40%) zaobserwowano tylko w jednym badaniu (badanie „A”). W pozostałych nie przekracza on 8%. Jest to, między innymi, efekt świadomych błędów ankieterskich. Badanie to realizowane było bowiem na próbie udziałowej, gdzie jednym z kryteriów doboru był wiek respondenta (ankieterzy „naciągali” to kryterium). Jeśli pominiemy w analizie to badanie, maksymalny odsetek niezgodności w powtórzonym pomiarze wyniesie 7,1%, natomiast średni odsetek zgodności wzrośnie do 96,3%.

Tabela 1. Zgodność odpowiedzi na pytania metryczkowe w badaniach zasadniczych i kontrolnych

Projekt / liczba badań	Rok realizacji /kwartał	Odpowiedzi zgodne (w %)							Samochód	Telefon
		Płeć resp.	Wiek resp.	Wykształce- nie resp	Sytuacja zaw. resp.	Liczba osób w rodzinie	Płeć członka rodziny	Wiek członka rodziny		
<b>A</b>	1997	99,0	60,5		87,7					
<b>B</b>	1997	100,0	92,9	90,6	94,9*	86,8		91,8		
<b>C1</b>	1997/1					87,9	96,3	83,2	96,0	
<b>C2</b>	1997/2					89,2	98,1	90,2		98,4
<b>C3</b>	1998/1					89,3	97,2	85,8		99,2
<b>C4</b>	1998/2					94,3	99,1	88,4		95,2
<b>C5</b>	1999/1					85,7	100	82,3		99,0
<b>C6</b>	1999/2					89,5	100	80,4		100,0
<b>C7</b>	2000/1					83,6	97,4	77,8		98,1
<b>C8</b>	2000/2					85,4	100	83,3		97,4
<b>D</b>	2000	100,0	95,8	74,5	76,9	82,9			88,9	
<b>E1/6</b> badań	1998/1	100,0	95,9	77,0		81,8	98,2	84,9		
<b>E2/6</b> badań	1998/2	100,0	97,3	78,6		85,4	95,9	82,8		
<b>E3/6</b> badań	1998/3	100,0	97,3	72,6		82,7	97,5	82,7		
<b>E4/6</b> badań	1998/4	100,0	98,2	75,8		83,4	94,4	85,1		
<b>E5/6</b> badań	1999/1	100,0	95,5	76,6		83,5	94,3	82,8		
<b>E6/6</b> badań	1999/2	100,0	98,9	80,3		91,5	94,7	76,9		
<b>E7/6</b> badań	1999/3	100,0	97,1	82,1		81,4	96,9	87,5		
<b>E8/6</b> badań	1999/4	100,0	95,8	77,9		86,1	98,7	78,2		
<b>E9/6</b> badań	2000/1	100,0	95,6	83,2		86,6	97,5	84,3		
<b>E10/6</b> badań	2000/2	100,0	96,1	88,7		85,2	96,2	86,2		
<b>E11/6</b> badań	2000/3	100,0	96,6	79,9		84,3	95,9	77,1		
<b>E12/6</b> badań	2000/4	100,0	95,1	79,0		87,1	93,9	79,2		
Średni % odp. zgodnych		99,9	93,9	79,8	82,3	86,1	97,1	82,9	90,4	97,9

\* wg kategoryzacji: pracujący - nie pracujący



powiedzi zgodnych w projekcie badawczym „E”: dla 12 fal tego projektu (72 badania) nie przekraczają one 3,8 punktów procentowych, zaś średni odsetek zgodności wynosi 96,6%.

Powyższe odsetki dotyczą pomiaru z dokładnością do 1 roku. Jeśli przyjmiemy wyższy poziom tolerancji – praktycznie w żadnych bowiem badaniach nie analizuje się danych ze względu na roku urodzenia respondenta (analizy takie prowadzi się według kohort wiekowych), odsetek odpowiedzi zgodnych zwiększa się. I tak, przykładowo w badaniu „A” będą one wynosić odpowiednio:

- 60,5%            z dokładnością do 1 roku
- 87,7%            z tolerancją +/- 1 rok
- 91,3%            z tolerancją +/- 2 lata
- 94,4%            z tolerancją +/- 3 lata
- 95,9%            z tolerancją +/- 4 lata
- 96,9%            z tolerancją +/- 5 lat

W związku z powyższym można oczekiwać, że błąd pomiaru zmiennej „wiek respondenta”, w pewnym zakresie może być zniwelowany ze względu na „szerokość” kohort (osoby wskazujące swój rok urodzenia z tolerancją np. +/- 2 lata „mieszczą się” w kohorcie). Okazuje się to jednak złudne: w analizowanym badaniu (badanie „A”), przyjmując pięcioletnie kohorty wiekowe, odsetek zgodności odpowiedzi wynosi jedynie 85,1%, a więc jest niższy niż wtedy, gdy analizujemy wiek respondenta z dokładnością +/- 1 rok.

C. Wykształcenie respondenta. Kolejną zmienną, w przypadku której spróbujemy określić błąd pomiaru, jest wykształcenie respondenta. Błąd pomiaru tej cechy także wystąpił we wszystkich analizowanych badaniach. Maksymalny odsetek odpowiedzi niezgodnych w powtórzonym pomiarze wynosi 27,4%, minimalny – 9,4%, natomiast średni odsetek zgodności – 79,8%. Warto zwrócić uwagę na dość znaczne różnice odsetka odpowiedzi zgodnych w projekcie badawczym „E” – dla 12 fal tego projektu (72 badania) sięgają one 16 punktów procentowych.

Wskazane powyżej odsetki dotyczą pomiaru według następujących 9 kategorii wykształcenia: niepełne podstawowe, podstawowe, zasadnicze zawodowe, niepełne średnie, ukończone średnie zawodowe, ukończone średnie ogólnokształcące, pomaturalne, niepełne wyższe oraz wyższe. Jeśli przyjmiemy wyższy poziom tolerancji pomiaru, tzn. wyłącznie według poziomów wykształcenia (podstawowe, zasadnicze zawodowe, średnie, pomaturalne, wyższe), odsetek odpowiedzi zgodnych wzrasta nawet do 90%.

D. Sytuacja zawodowa respondenta. W przypadku tej zmiennej baza materiałowa, na której opieramy wnioski, jest znacznie uboższa: dotyczy jedynie 3 projek-

tów. Niezależnie od tego błąd pomiaru tej cechy wystąpił we wszystkich tych badaniach. Maksymalny odsetek odpowiedzi niezgodnych w powtórzonym pomiarze wynosi 23,1%, minimalny – 12,3%, natomiast średni odsetek zgodności – 82,3%.

Powyższe odsetki dotyczą szczegółowego pomiaru sytuacji zawodowej (z pominięciem danych uzyskanych w badaniu „B”). Były one analizowane według następujących kategorii szczegółowych: stała praca zarobkowa, praca dorywcza, bezrobocie, emerytura/renta, nauka/studia, nie pracuje/zajmuje się domem, urlop wychowawczy/urlop macierzyński, nie pracuje z innych powodów. Jeśli przyjmiemy wyższy poziom tolerancji pomiaru, tzn. wyłącznie według kategorii: pracujący – nie pracujący, odsetek odpowiedzi zgodnych wzrasta nawet do 95%.

E. Liczba osób w gospodarstwie domowym respondenta. Błąd pomiaru tej cechy także można uznać za powszechny, ponieważ wystąpił we wszystkich analizowanych badaniach. Maksymalny odsetek odpowiedzi niezgodnych w powtórzonym pomiarze wynosi 18,6%, minimalny – 5,7%, natomiast średni odsetek zgodności – 86,1%. Warto wrócić uwagę na różnice odsetka odpowiedzi zgodnych w projekcie badawczym „C” oraz projekcie „E”. W pierwszym przypadku dla 8 fal nie przekraczają one 11 punktów procentowych, natomiast w drugim – dla 12 fal (72 badania) sięgają 10 punktów. Jednocześnie, rząd wielkości średniego odsetka odpowiedzi zgodnych jest zbliżony i odpowiednio wynosi 88,1% (projekt „C”) i 84,9 (projekt „E”).

F. Płeć członków gospodarstwa domowego respondenta. Błąd pomiaru tej cechy nie jest tak powszechny jak poprzednich – wystąpił tylko w niektórych badaniach<sup>10</sup>. Przy czym, co należy podkreślić, pominięto tutaj wszystkie przypadki niezgodności liczby osób w gospodarstwie domowym, a dodatkowo brano pod uwagę niezgodność płci któregokolwiek z członków gospodarstwa. Maksymalny odsetek odpowiedzi rozbieżnych w powtórzonym pomiarze wynosi 6,1%, natomiast średni odsetek zgodności – 97,1%. Warto zwrócić uwagę na niewielkie różnice odsetka odpowiedzi zgodnych w projekcie badawczym „C” – dla 8 fal nie przekraczają one 3,7 punktów procentowych oraz w projekcie „E” – 12 fal tego projektu (72 badania) nie przekraczają 4,8 punktów procentowych. Jednocześnie, średni odsetek odpowiedzi zgodnych jest zbliżony i odpowiednio wynosi 98,5% (projekt „C”) i 96,2 (projekt „E”).

G. Wiek członków gospodarstwa domowego respondenta (poza respondentem). Błąd pomiaru tej cechy występuje z kolei we wszystkich analizowanych badaniach (należy podkreślić, iż w analizie pominięto wszystkie przypadki nie-

---

<sup>10</sup> Wynika to być może z dychotomicznego charakteru tej cechy. W przypadku płci respondenta zaobserwowano równie niewielki odsetek niezgodności.

zgodności liczby i płci osób w gospodarstwie domowym, brano natomiast pod uwagę niezgodność wieku powyżej 1 roku któregośkolwiek z członków gospodarstwa). Maksymalny odsetek odpowiedzi rozbieżnych w powtórzonym pomiarze wynosi 23,1%, minimalny - 9,8%, natomiast średni odsetek zgodności - 82,9% (jeśli przyjmiemy wyższy poziom tolerancji: +/- 1 rok, odsetek odpowiedzi zgodnych wzrasta do 90%).

Warto także zwrócić uwagę na różnice odsetka odpowiedzi zgodnych w projekcie badawczym „C” oraz projekcie „E”. W pierwszym przypadku dla 8 fal przekraczają one 12 punktów procentowych, natomiast w drugim - dla 12 fal (72 badania) przekraczają 9 punktów. Jednocześnie, rząd wielkości średniego odsetka odpowiedzi zgodnych jest zbliżony i odpowiednio wynosi 83,9% (projekt „C”) i 82,3% (projekt „E”).

Analizując zmienne dotyczące gospodarstwa domowego, należy podkreślić, iż „łączny” błąd ich pomiaru przekracza 30%. Oznacza to, iż jedynie w niespełna 70 przypadkach na 100 uzyskujemy rzetelne informacje dotyczące liczby osób w gospodarstwie domowym oraz płci i wieku poszczególnych jego członków. Oczywiście ekstremalne sytuacje to te, w których „nic się nie zgadza”. Poniżej przedstawiono w formie zestawienia odpowiedzi jednego z respondentów na pytanie o wiek i płeć wszystkich członków jego gospodarstwa domowego, uzyskane w wywiadzie zasadniczym i wywiadzie kontrolnym.

Tabela 2. Płeć i wiek członków gospodarstwa domowego (poza respondentem)

Pomiar zasadniczy	Pomiar powtórzony
1. Kobieta ur. 1973 r. (żona respondenta)	1. Kobieta ur. 1976 r. (żona respondenta)
2. Kobieta ur. 1998 r. (córka respondenta)	2. Mężczyzna ur. 1997 r. (syn respondenta)
	3. Kobieta ur. 1940 r. (teściowa respondenta)
	4. Mężczyzna ur. 1940 r. (teść respondenta)

H. Wyposażenie materialne gospodarstwa domowego respondenta. Ze względu na znaczne zróżnicowanie zestawu dóbr materialnych będących przedmiotem pytania, naszą analizę ograniczymy wyłącznie do dwóch spośród możliwych: samochodu oraz linii telefonicznej.

W przypadku zmiennej: posiadanie samochodu, baza materiałowa, na której opieramy wnioski jest uboga (jedynie 2 projekty). Niezależnie od tego błąd pomiaru tej cechy wystąpił w obu badaniach. Maksymalny odsetek odpowiedzi niezgodnych w powtórzonym pomiarze wynosi 11,1%, minimalny – 8,2%, natomiast średni odsetek zgodności – 90,4%. Powyższe odsetki dotyczą odpowiedzi na pytanie o posiadanie samochodu, niezależnie od ich liczby. W przypadku pytania o liczbę posiadanych samochodów odsetek odpowiedzi zgodnych spada do poziomu 85% (według kategorii: brak samochodu w gospodarstwie, 1 samochód, 2 samochody, więcej samochodów).

Z kolei w przypadku zmiennej: posiadanie linii telefonicznej, baza materiałowa, na której opieramy wnioski obejmuje 1 projekt (8 badań). Błąd pomiaru tej cechy nie występuje w każdym badaniu. Maksymalny odsetek odpowiedzi niezgodnych w powtórzonym pomiarze wynosi 4,8%, natomiast średni odsetek zgodności – 97,9%.

Podsumowując wyniki powyższych analiz należy stwierdzić, iż błąd rozumiany jako niezgodność odpowiedzi w pomiarze powtórzonym dotyczy wszystkich cech metryczkowych. Przy czym „prawdopodobieństwo” jego wystąpienia jest różne i zależy od charakteru cechy. Wbrew pozorom nie jest tak, iż częściej występuje on przy pomiarze cech kontekstowych (dotyczących gospodarstwa domowego), w przypadku których respondent jest informatorem o innych. Przykładowo, relatywnie najniższy odsetek odpowiedzi zgodnych zaobserwować można w pytaniach o sytuację zawodową respondenta i jego wykształcenie. Poniższa tabela ilustruje opisaną sytuację.

Tabela 3. Zmienne metryczkowe uszeregowane według średniego odsetka odpowiedzi zgodnych uzyskanych w powtórnym pomiarze

Zmienna	%	Ranga
Płeć respondenta	99,9	1
Posiadanie linii telefonicznej	97,9	2
Płeć członków gospodarstwa domowego	97,1	3
Wiek respondenta	93,9	4
Posiadanie samochodu	90,4	5
Liczba osób w gospodarstwie domowym	86,1	6
Wiek członków gospodarstwa domowego	82,9	7
Sytuacja zawodowa respondenta	82,3	8
Wykształcenie respondenta	79,8	9

Z kolei maksymalna różnica odsetka odpowiedzi zgodnych uzyskanych w powtórny pomiarze (tzn. różnica między maksymalnym i minimalnym odsetkiem odpowiedzi zgodnych dla danej cechy) wskazuje na dość duże wahania rozmiaru tego błędu. I znowu, nie jest tak, iż są one większe w przypadku cech kontekstowych. Przykładowo, relatywnie największą różnicę można zaobserwować w pytaniu o wykształcenie respondenta. Opisaną sytuację ilustruje poniższa tabela.

Tabela 4. Zmienne metryczkowe uszeregowane według maksymalnej różnicy odsetka odpowiedzi zgodnych uzyskanych w powtórny pomiarze<sup>11</sup>

Zmienna	Punkty %	Ranga
Płeć respondenta	1,0	1
Posiadanie samochodu	2,9	2
Posiadanie linii telefonicznej	4,8	3
Płeć członków gospodarstwa domowego	6,1	4
Sytuacja zawodowa respondenta	10,8	5
Liczba osób w gospodarstwie domowym	12,9	6
Wiek członków gospodarstwa domowego	13,3	7
Wykształcenie respondenta	18,0	8
Wiek respondenta	38,4	9

### **Charakter błędów: błąd jednostkowego pomiaru czy błąd właściwy każdemu badaniu?**

Odpowiedź na to pytanie w pewnym zakresie udzielona została już wcześniej: jeśli jakiś błąd występuje bowiem w kilkunastu, a w niektórych przypadkach kilkudziesięciu różnych badaniach, to nie można chyba mówić o błędach jednostkowych, którymi obarczone są tylko niektóre pomiary. Co więcej, bliższa analiza dwóch projektów realizowanych przez dłuższy czas (punktowo - projekt „C” i w sposób ciągły - projekt „E”), wskazuje na istnienie dość interesującej tendencji. Otóż średnie odsetki odpowiedzi zgodnych w pomiarze powtórzo-

<sup>11</sup> W przypadku zmiennej: wiek respondenta uwzględniono także wyniki powtórzonego pomiaru w badaniu „A” (realizowanego na próbie udziałowej - zob. przyp. 9.). Jeśli pominiemy to badanie, maksymalna różnica odsetka odpowiedzi zgodnych uzyskanych w powtórny pomiarze dla tej cechy wyniesie tylko 6 punktów procentowych (nie zaś 38,4) i tym samym zmienna ta zajmie czwartą pozycję w rankingu.

nym dla wszystkich zmiennych metryczkowych łącznie, w poszczególnych falach realizacji danego projektu są zbliżone. Zbieżność tę dla tych dwóch projektów ilustruje poniższa tabela.

Tabela 5. Średni odsetek odpowiedzi zgodnych na wszystkie pytania metryczkowe w projekcie „C” i „E” dla poszczególnych fal realizacji

Rok/kwartał realizacji fali projektu	Średni odsetek odpowiedzi zgodnych	
	Projekt „C”	Projekt „E”
1997/1	90,9	
1997/2	94,0	
1998/1		89,6
1998/2	92,9	90,0
1998/3		88,8
1998/4	94,3	89,5
1999/1		88,8
1999/2	91,8	90,4
1999/3		90,8
1999/4	92,5	89,5
2000/1		91,2
2000/2	89,2	92,1
2000/3		88,9
2000/4	91,5	89,1
Maksymalna różnica (pkt. %)	5,1	3,3

W przypadku projektu „C”, w poszczególnych falach badania (analiza 8 fal realizowanych na przestrzeni 4 lat) maksymalna różnica odsetka odpowiedzi zgodnych dla wszystkich pytań metryczkowych wynosi 5,1 punktu procentowego. Dla projektu „E” (projekt ciągle realizowany na przestrzeni 3 lat, analizowany w 12 falach kwartalnych) ten sam wskaźnik wynosi 3,3 punktu. Wynik ten wskazuje więc, iż błąd pomiaru zmiennych metryczkowych jest w miarę „stabilny”, a co więcej – występuje w każdym badaniu, jest ich immanentną cechą.

Niemniej jednak, interesujące wydaje się porównanie zaprezentowanych wyników z wynikami innego badania, aczkolwiek ogniskuje się ono na nieco innych cechach metryczkowych. Chodzi o badanie P. Daniłowicza i P. B. Sztabińskiego (1986), realizowane w ramach projektu weryfikacyjnego<sup>12</sup>. Porównanie wyników tego badania z przez nas analizowanymi zawiera tabela 6.

<sup>12</sup> W badaniu tym błąd pomiaru był określany na podstawie porównania deklaracji respondenta z danymi dokumentalnymi, a więc w inny sposób niż prezentowany w naszych analizach.

Tabela 6. Błąd pomiaru zmiennych metryczkowych: porównanie wyników badań kontrolnych (1997-2000) i weryfikacyjnych (1976-1977)

Zmienna	Odsetek zgodności w badaniu:	
	1997-2000	1976-1977
Płeć respondenta	99,9	100,0
Wiek respondenta	93,9	96,4
Wykształcenie respondenta	79,8	72,4
Posiadanie samochodu	90,4	95,2

Generalnie, odsetki dla poszczególnych zmiennych są zbliżone: różnice wahają się od 0,1 do 7,4 punktu procentowego.

W przypadku zmiennej: wiek respondenta, różnica zmniejszy się, jeśli wskaźnik dla badań 1997-2000 zostanie obliczony z pominięciem projektu „A” (por. przyp. 9 i 11). Wynosić będzie on wtedy 96,3%, a więc różnica zmniejszy się do 0,1 punktu. Interpretując wielkość rozbieżności w pomiarze tej cechy Daniłowicz i Sztabiński stwierdzają, że „rozkłady wieku (...) trafnie odzwierciedlają rozkłady rzeczywiste” (Daniłowicz i Sztabiński 1986: 98).

W przypadku zmiennej: wykształcenie, różnica jest znacznie większa i wynosi 7,4 punktu. Niższy poziom zgodności zaobserwowali wspomniani autorzy, niemniej uznają, iż: „istnieje bardzo silna zależność pomiędzy informacjami z wywiadu a danymi dokumentalnymi. (...) Można zatem orzekać o podobieństwie struktur badanych prób pod względem wykształcenia” (tamże, s. 104).

Z kolei wyższy odsetek zgodności stwierdzili oni w przypadku posiadania samochodu, aczkolwiek różnica nie jest wysoka - wynosi 4,8 punktu procentowego (tamże, s. 164).

Warto tutaj wspomnieć, iż autorzy ci, poza wspomnianymi, analizowali także inne cechy metryczkowe: wykształcenie współmałżonka respondenta, zawód respondenta i współmałżonka, zarobki respondenta i współmałżonka oraz dochody gospodarstwa domowego, warunki mieszkaniowe i zasoby (tamże, s. 97-166). W przypadku wszystkich tych zmiennych stwierdzili rozbieżności między deklaracjami respondentów a danymi dokumentalnymi, co potwierdza tezę o „powszechności” błędów.

Podobny wniosek wysnuć można na podstawie badań M. Poulain, B. Rian-dey i J.-M. Firdion (1991: 65-88), realizowanych na próbie 500 gospodarstw domowych. Autorzy ci stwierdzają, iż mężczyźni nieco częściej niż kobiety popełniają błędy w raportowaniu pewnych faktów, a co więcej, błędy te są popełnia-

ne nawet wtedy, gdy istnieje możliwość uzgodnienia swoich odpowiedzi z drugą osobą<sup>13</sup>. Poniżej przedstawione są wyniki badania w odniesieniu do 3 cech społeczno-demograficznych.

Tabela 7. Odsetek odpowiedzi zgodnych z danymi dokumentalnymi na pytania o cechy społeczno-demograficzne (w rozbiciu na deklaracje mężczyzn i kobiet oraz „uzgodnione”)

Cecha	Odsetek odpowiedzi zgodnych z danymi dokumentalnymi		
	Mężczyzna (mąż)	Kobieta (żona)	„Uzgodnione”
Data zawarcia związku	94,1	98,9	99,6
Narodziny dzieci	93,8	97,9	98,3
Usamodzielnienie się dzieci	44,2	52,2	53,6

## Źródła błędu pomiaru zmiennych metryczkowych

Jednym z podstawowych dogmatów socjologii empirycznej jest „obiektywizm badania”, który rozumiany może być, między innymi, jako „obiektywizm” rezultatów. Chodzi o to, aby rezultaty badania były „obiektywne”, tzn. wolne od zanieczyszczeń, czy inaczej mówiąc – błędów. Tak rozumiany „obiektywizm” jest oczywiście, co staraliśmy się wykazać, wyłącznie postulatem. W naukach społecznych pomiar w pełni obiektywny nie jest możliwy i to z dwóch powodów. Po pierwsze, ze względu na błąd związany z próbą, tzn. błąd losowania, błąd „pokrycia” oraz błąd „braku odpowiedzi”. Po drugie, ze względu na błąd będący pochodną przygotowania narzędzia, a więc związany ze sposobem uzyskiwania danych (w tym związany z osobą ankietera i respondenta) oraz sposobem pytania (Sztabiński P. 1997: 14–35). Błędy, które stanowią przedmiot naszych zainteresowań z pewnością nie są związane z próbą. Ich źródła należy zatem upatrywać w sposobie uzyskiwania danych oraz specyfice pytania.

<sup>13</sup> W badaniu tym wywiady prowadzone były w gospodarstwach domowych równolegle z mężem i żoną (w osobnych pomieszczeniach, bez możliwości wzajemnej konsultacji), a następnie – wywiad w „diadzie” (mąż i żona), z możliwością wzajemnej konsultacji. W kolejnym kroku odpowiedzi uzyskane od poszczególnych osób oraz „wspólne” były porównywane z danymi dokumentalnymi.



Realizacja każdego badania jest praktycznie zawsze związana z tworzeniem pewnych sztucznych warunków, umożliwiających jego prowadzenie (niezależnie od tego, każda z technik wprowadza specyficzne zniekształcenia). W przypadku wywiadu kwestionariuszowego konsekwencją owej sztuczności są przede wszystkim błędy związane z osobą ankietera oraz respondenta i interakcją między uczestnikami procesu badawczego. Ich źródłem mogą być obserwowalne cechy ankietera i respondenta (w tym charakterystyki społeczno-demograficzne, takie jak: płeć, wiek, wykształcenie itp. oraz „pozaspołeczne”: ogólne wrażenie, sposób zachowania), a także ich cechy nieobserwowalne – „psychologiczne” (system wartości, opinie, wyobrażenia, umiejętności, wiedza, motywacje itp.). W przypadku respondenta może być to również skłonność do udzielania społecznie akceptowanych odpowiedzi, „zgadzania się”, ale także wyobrażenia o badaniu, jego celach, ośrodku je realizującym itp. (tamże, s. 46-92)<sup>14</sup>.

Z kolei Sudman i Bradburn (1973: 805-815) w jednym ze swoich artykułów wskazują, iż odpowiedź respondenta jest wypadkową kilku czynników, w tym przede wszystkim: błędów pamięci, społecznej waloryzacji oraz obiektywnego braku wiedzy. Szczególnie interesujące, nie tylko w przypadku zmiennych metryczkowych, są błędy pamięci (wspomniane wcześniej błędy pominięcia oraz błędy będące wynikiem efektu teleskopowego) oraz błędy związane z waloryzacją społeczną, a więc błędy pochodzenia psychologicznego (świadome i nieświadome) oraz pochodzenia społecznego (por. Auriat 1996: 18-19).

Znając zatem możliwe źródła błędów w odpowiedziach respondentów na pytania, spróbujmy zastanowić się, jakie mogą być przyczyny rozbieżności w pomiarze cech metryczkowych w analizowanych badaniach. Wskazując je nie możemy jednak zapominać o specyfice powtórzonego pomiaru i fakcie, iż nie zawsze jest on niezależny (Blair i Sudman 1996: 421-426).

A. Płeć respondenta. Błąd pomiaru tej cechy zasadniczo nie powinien się pojawiać – cecha ta jest bowiem cechą jawną. Niestety, wbrew pozorom jest on dość częsty. Jego podstawowym źródłem jest ankieter, a konkretnie nieświadome błędy (pomyłki) oraz świadome (intencjonalne) zafalszowania. Ta ostatnia sytuacja ma miejsce przede wszystkim w przypadku badań realizowanych na próbach udziałowych, w którym jednym z kryteriów doboru jest płeć (por. przyp. 9.).

B. Wiek respondenta. Z całą pewnością można stwierdzić, iż poza przypadkami klinicznymi, błąd pomiaru tej cechy nie jest spowodowany błędami pamięci,

---

<sup>14</sup> Autor ten dokonuje przeglądu literatury poświęconej problemowi skłonności do „zgadzania się” oraz udzielania społecznie akceptowanych odpowiedzi (s. 78-89).

ani obiektywnym brakiem wiedzy. Wbrew jednak powszechnemu przekonaniu, iż zmienna ta jest bardzo podatna na tzw. „wpływ ankieterski” (obserwowalne i nieobserwowalne cechy ankietera), pomiar tej cechy uznać należy za rzetelny. Odsetek odpowiedzi niezgodnych rzędu 4% jest bowiem, w porównaniu do analogicznego wskaźnika dla innych zmiennych, niski. Pamiętać jednak należy, iż potencjalny błąd, którego źródłem jest ankieter, może pojawić się tylko w sytuacji, gdy współwystępuje z nim świadoma bądź nie, tendencja respondenta do waloryzowania sytuacji i udzielania społecznie akceptowanych odpowiedzi. Powyższe zjawiska nie przesądzają kierunku niezgodności pomiaru: w praktyce zawyżenia wieku respondenta równoważą liczbę przypadków zaniżeń.

C. Wykształcenie respondenta. W przypadku pomiaru tej cechy źródła błędów mogą być nieco inne, aczkolwiek także z dużym prawdopodobieństwem można stwierdzić, iż nie są one spowodowane błędami pamięci. W pewnym stopniu są one efektem obiektywnego braku wiedzy (przykładowo, respondenci nie zawsze wiedzą, do której kategorii zaliczyć tzw. „czteroklasówkę”, przedwojenne gimnazjum czy „małą maturę”). Z pewnością cecha ta jest bardzo podatna na „wpływ ankieterski”. Wykształcenie jest bowiem w naszym społeczeństwie wartością bardzo wysoko waloryzowaną. Stąd też zaobserwować można znaczną przewagę przypadków zawyżeń poziomu wykształcenia nad przypadkami zaniżeń (zarówno w powtórzonym pomiarze, jak i weryfikacji dokumentalnej). Tendencję do zawyżania poziomu wykształcenia w przypadku powtózonego pomiaru tłumaczyć można dodatkowo dwoma czynnikami: charakterem kontaktu ankietera *versus* kontrolera z respondentem oraz błędami ankietera. W pierwszym przypadku (charakter kontaktu) można przyjąć, iż w sytuacji powtózonego pomiaru relacje te mają charakter krótkotrwały, bardziej oficjalny (w porównaniu z kontaktem z ankieterem) i w związku z tym w większym stopniu generują odpowiedzi społecznie pożądane. Z kolei błędy ankieterskie sprowadzają się przede wszystkim do zmiany procedury pytania: nie odczytania kafeterii lub nie wręczenia Karty Respondenta<sup>15</sup>.

D. Sytuacja zawodowa respondenta. W przypadku tej zmiennej źródeł błędów pomiaru upatrywać należy przede wszystkim w: obiektywnym braku wiedzy respondentów oraz faktycznych zmianach sytuacji respondenta (w okresie między pomiarem zasadniczym a powtórzonym/weryfikacją dokumentalną).

---

<sup>15</sup> W obu przypadkach efekt jest ten sam: respondent nie wie, w jakich kategoriach powinien udzielać odpowiedzi, jak szczegółowa winna być jego odpowiedź (chodzi przede wszystkim o poziom wykształcenia: ukończone - nieukończone oraz kierunek: średnie ogólnokształcące - średnie zawodowe).

W pierwszym przypadku - zwłaszcza wśród ludzi starszych oraz rolników - możemy mieć do czynienia z mylną kwalifikacją własnej sytuacji zawodowej (np. „emeryt - rencista”, „emeryt/rencista - pracujący emeryt/rencista”, „rolnik/współmałżonek rolnika jako osoba pracująca - nie pracująca”). W przypadku tej zmiennej możemy jednak mieć też do czynienia z faktyczną zmianą sytuacji zawodowej respondenta: przykładowo, osoba bezrobotna lub zajmująca się domem podjęła pracę zarobkową lub odwrotnie osoba wcześniej pracująca zarobkowo straciła pracę lub przeszła na emeryturę itp. Marginalnym, aczkolwiek nie we wszystkich badaniach, czynnikiem powodującym błąd pomiaru może być charakter kontaktu ankietera z respondentem oraz „wpływ ankierski” (chodzi tutaj przede wszystkim o wszelkiego rodzaju obawy/brak obaw związanych z nieoficjalnymi źródłami zarobkowania - tzw. „praca na czarno”).

E. Liczba osób w gospodarstwie domowym respondenta. Błąd pomiaru tej cechy jest niemal powszechny. Jego źródłem wydaje się niestabilność sposobu definiowania „gospodarstwa domowego” przez respondentów. Przykładowo, w niektórych przypadkach respondenci włączali, bądź wyłączali z gospodarstwa domowego osoby będące ich rodzicami lub dorosłymi dziećmi i ich współmałżonków. W tym przypadku wspomniany błąd pomiaru nie jest więc błędem ankietera, lecz błędem respondenta spowodowanym obiektywnym brakiem wiedzy<sup>16</sup>. Co więcej, wydaje się, iż błąd ten ma charakter losowy, ponieważ liczba przypadków zaniżeń i zawyżeń liczby osób w gospodarstwie w wywiadzie zasadniczym i kontrolnym wzajemnie się równoważy. Nie można także zapominać o innym jeszcze czynniku „błędogennym”, a mianowicie faktycznej zmianie składu gospodarstwa domowego respondenta w okresie między pomiarem zasadniczym a powtórzonym. Przykładowo, dorosły syn mógł się wyprowadzić od rodziców, mógł „wprowadzić” swoją żonę, mogło urodzić się dziecko, ktoś mógł umrzeć. Wszystko to są naturalne zmiany w ramach gospodarstwa domowego i nikt za błąd nimi spowodowany nie jest odpowiedzialny.

F. Płeć członków gospodarstwa domowego respondenta (poza respondentem). Błąd pomiaru tej cechy wynika - podobnie jak w przypadku zmiennej: płeć respondenta - raczej z nieświadomych błędów ankietera (pomyłek), nie zaś zniekształceń ze strony respondenta<sup>17</sup>.

---

<sup>16</sup> Ów obiektywny brak wiedzy w wielu przypadkach może być efektem błędnego sformułowania pytania. Czasami bowiem zadaje się pytanie o liczbę osób w gospodarstwie nie przytaczając definicji gospodarstwa domowego w treści pytania. Oczywiście nie dotyczy to sytuacji, w której liczebność gospodarstwa jest kryterium ich doboru do badania.

<sup>17</sup> Błąd pomiaru tej cechy rozpatrujemy wyłącznie dla przypadków zgodności liczby osób w gospodarstwie domowym respondenta.

G. Wiek członków gospodarstwa domowego respondenta (poza respondentem). Błędy pomiaru tej cechy upatrywać należy przede wszystkim w błędach pamięci respondenta (lub obiektywnym braku wiedzy) oraz, w niewielkim stopniu, we „wplywie ankieterskim”. Błędy pamięci lub braku wiedzy dotyczą przede wszystkim osób z pokolenia starszego niż respondent (rodzice, dziadkowie), aczkolwiek wystąpić mogą również w informacjach na temat wieku młodszego/starszego rodzeństwa/dzieci. Błędy te mogą mieć także związek z wielkością gospodarstwa domowego respondenta: im jest ono mniejsze, tym „łatwiej wszystko zapamiętać”. Tego typu błąd w odniesieniu do wieku współmałżonka/dzieci może być efektem społecznej waloryzacji. Przykładowo, respondent zawyża/zaniża wiek współmałżonka (ponieważ „ma zbyt młodego/starego współmałżonka”) lub też, wiek dzieci („w moim wieku mam zbyt małe/duże dzieci”). Generalnie jednak, niezależnie od źródeł błędów, nie można uznać, że respondent jest kompetentnym informatorem o innych członkach jego gospodarstwa domowego.

H. Wyposażenie materialne gospodarstwa domowego respondenta. W przypadku tych zmiennych (posiadanie linii telefonicznej/samochodu) praktycznie nie sposób uznać błędów pamięci lub braku wiedzy respondenta. Jedynym zatem ich źródłem może być „wplyw ankieterski” i związana z tym waloryzacja: w niektórych środowiskach „wstyd” jest się przyznać, że nie posiada się telefonu lub samochodu (dobra te mogą stanowić według niektórych respondentów wyznacznik statusu społecznego lub materialnego – są to tzw. dobra „prestżowe”). Wplyw ankieterski na uzyskiwane odpowiedzi może jednak mieć „odwrotny” efekt. Respondent może nie przyznać się do posiadania pewnych dóbr (samochód) z obawy przed zagrożeniem bezpieczeństwa osobistego (Sztabiński F. i Sztabiński P. 1994). Błąd pomiaru tej cechy może być także skutkiem obiektywnych zmian w wyposażeniu gospodarstwa (zakup samochodu – pierwszego lub drugiego, bądź jego/ich sprzedaż).

Analizując źródła błędu w pomiarze cech metryczkowych wskazywaliśmy przede wszystkim na respondenta, jego „nieobserwowalne” cechy, które są niejako warunkiem koniecznym aktywizacji „wplywu” ankieterskiego. Musimy jednak zdawać sobie sprawę z istnienia błędów związanych bezpośrednio z ankieterem i jakością jego pracy. Innymi słowy, osobnym źródłem zniekształceń pomiaru cech metryczkowych jest niska jakość pracy ankietera. Podstawowym czynnikiem jest tutaj błędne przekonanie o „łatwości” pytań metryczkowych („takie same pytania lub podobne zamieszczane są przecież w kwestionariuszu do każdego badania”) oraz mniejszym ich znaczeniu („są na końcu, a więc mniej ważne”). Dobrym przykładem „stosunku” ankieterów

do pytań metryczkowych jest fragment transkrypcji wywiadu utrwalonego na taśmie VHS<sup>18</sup>.

Ankieter: O dobrze, prawie wywiad się skończył! Tylko jeszcze Pana zapytam o kilka danych... Zapytam Pana o kilka danych o... takich pomocnych już jakby na marginesie wywiadu.

Taki sposób postrzegania pytań metryczkowych prowadzi do pomijania niektórych z nich lub ich „opowiadania” – zamiast odczytywania. Nie bez znaczenia niewątpliwie jest tutaj również zmęczenie i znużenie wywiadem ankietera. W efekcie, jak wykazują wyniki badań metodologicznych, liczba błędów ankierskich w tej części kwestionariusza jest zazwyczaj najwyższa (Sztabiński F. 1997: 104-118).

Tego typu podejście niektórych ankieterów do „metryczki” bezpośrednio może rzutować na stosunek respondentów do tychże pytań. Respondenci nie uświadamiając sobie znaczenia i nie widząc związku między pytaniami merytorycznymi i metryczkowymi nie zadają sobie trudu udzielania na nie „prawdziwych” odpowiedzi (Sztabiński F. i Sztabiński P. 1994). I znowu, dobrym przykładem jest fragment transkrypcji wywiadu utrwalonego na taśmie VHS (zob. przyp. 18).

Ankieter: To może przejdziemy teraz do metryczki. Teraz Pani zadam kilka pytań dotyczących gospodarstwa domowego

Respondent: Mogę wiedzieć czemu one służą?

Niestety, tego typu zachowania respondenta, zmuszające ankietera do odczytania wprowadzenia do bloku pytań metryczkowych nie zdarzają się często.

## **Uwagi końcowe**

Powyższe rozważania miały na celu zwrócenie uwagi na istnienie błędu w pomiarze cech metryczkowych praktycznie w każdym badaniu. Wyszliśmy z założenia, że im większy błąd w pomiarze tych cech, tym mniej wiemy o respondencie. I chyba rzeczywiście niewiele wiemy na ten temat. Z pewnością jest tak, iż pomiar niektórych zmiennych jest bardziej rzetelny (replikowalny), innych mniej. Musimy się z tym pogodzić. Ale, jeśli rozszerzyć zakres naszego

---

<sup>18</sup> Wywiady te realizowane były w ramach metodologicznych badań poświęconych analizie procesu zbierania informacji w jednym z projektów badawczych. Badania metodologiczne przygotowane były i bezpośrednio nadzorowane przez autora niniejszego artykułu.

wnioskowania na pytania merytoryczne o charakterze faktualnym (np. usamodzielnienie się dzieci, zakup pewnych produktów lub sprzętów), to pomiar tego typu zmiennych – na co wskazywaliśmy – również jest obarczony błędem (por. Silberstein 1989: 125–142, Neter i Wakesberg 1964: 18–55, Sudman i Bradburn 1973: 805–815).

Analizując błędy w pytaniach metryczkowych ograniczyliśmy się wyłącznie do określenia częstotliwości i zakresu ich występowania. Świadomie nie zajmowaliśmy się w ogóle konsekwencjami tego typu błędów (ich wpływem na wartość współczynników korelacji), a więc z pewnego punktu widzenia sprawą najważniejszą.

Naturalną kolejną rzeczą jest chyba bowiem najpierw uświadomienie sobie pewnych „niedoskonałości” (pomiaru), a dopiero później ich konsekwencji i podejmowanie prób przeciwdziałania im.

## Literatura

- Auriat, Nadia. 1993. „My wife knows best”. *A comparison of event dating accuracy between the wife, the husband, the couple, and the Belgium Population Register*. „Public Opinion Quarterly”, vol. 57, nr 2.
- Auriat, Nadia. 1996. *Les defaillances de la memoire humaine. Aspects cognitifs des enquetes retrospectives*. Paris: INED, PUF.
- Bailar, Barbara A. 1968. *Recent research in reinterview procedure*. „Journal of American Statistical Association”, vol. 63, nr 321.
- Biemer, Paul P., Robert M. Groves, Lars E. Lyberg, Nancy A. Mathiowetz i Seymour Sudman (red.) 1991. *Measurement Errors in Surveys*. New York: John Wiley & Sons.
- Biemer, Paul P. i Gösta Forsman. 1992. *On the quality of reinterview data with application to the current population survey*. „Journal of the American Statistical Association”, vol. 87, nr 420.
- Blair, Johnny i Seymour Sudman. 1996. *The survey reinterview: respondent perceptions and response strategies*. „Journal of Official Statistics”, vol. 12, nr 4.
- Bryman, Alan. 1995. *Quantity and Quality in Social Research*. London: Routledge.
- Cahalan, Don. 1968-69. *Correlates of respondent accuracy in the denver validity survey*. „Public Opinion Quarterly” (winter).
- Clausen, Aage R. 1968-69. *Response validity: vote report*. „Public Opinion Quarterly” (winter).
- Daniłowicz, Paweł i Paweł B. Sztabiński. 1986. *Zmienne osobowe w badaniach socjologicznych. Analizy weryfikacyjne*. Wrocław: Ossolineum.
- Forsman, Gösta i Schreiner Irwin. 1991. *The design and analysis of reinterview: An overview*. W: Biemer P., R. Groves, L. Lyberg, N. Mathiowetz, S. Sudman (red.), *Measurement Error in Surveys*. New York: John Wiley & Sons.

- Gaskel, George D., Daniel B. Wright i Colm A. O'Muircheartaigh. 2000. *Telescoping of landmark events. Implications for survey research*. „Public Opinion Quarterly”, vol. 64.
- Goode, William J. i Paul K. Hatt. 1952. *Methods in Social Research*. New York: Mc Graw-Hill.
- Heisenberg, Werner. 1975. *The great tradition: end of an epoch?* „Encounter”, vol. 44, nr 3.
- Hyman, Herbert H. 1954. *Interviewing in Social Research*. Chicago: University of Chicago Press.
- Jones, Russel A. 2000. *Méthodes de recherche en sciences sociales*. Paris: PUF.
- Koniarek, Jerzy. 1975. *Weryfikacja zewnętrzna informacji uzyskanych w wywiadzie kwestionariuszowym*. W: Z. Gostkowski, J. Lutyński (red.), *Analizy i próby technik badawczych w socjologii*, t. V. Wrocław: Ossolineum.
- Lutyńska, Krystyna. 1975. *Weryfikacja danych uzyskiwanych z odpowiedzi respondentów na pytania o wysokość zarobków*. W: Z. Gostkowski, J. Lutyński (red.), *Analizy i próby technik badawczych w socjologii*, t. V. Wrocław: Ossolineum.
- Lutyńska, Krystyna i Neonila Szeszenia. 1972. *Wstępne wyniki zewnętrznej weryfikacji odpowiedzi respondentów w wywiadzie na temat zachowań zdrowotnych*. W: Zygmunt Gostkowski i Jan Lutyński (red.), *Analizy i próby technik badawczych w socjologii*, t. IV. Wrocław: Ossolineum.
- Lutyński, Jan. 1975. *Analizy weryfikacyjne w badaniach z zastosowaniem wywiadu kwestionariuszowego, ich rodzaje i możliwości*. W: Zygmunt Gostkowski i Jan Lutyński (red.), *Analizy i próby technik badawczych w socjologii*, t. V. Wrocław: Ossolineum.
- Miller, Peter V. i Robert M. Groves. 1985. *Matching survey response to official records: an exploration of validity in victimization reporting*. „Public Opinion Quarterly”, vol. 49, nr 3.
- Neter, John i John Wakesberg. 1964. *A study of response errors in expenditures data from household interviews*. „Journal of American Statistical Association”, vol. 59.
- Parry, Hugh J. i Helen M. Crossley. 1950. *Validity of responses to survey questions*. „Public Opinion Quarterly” (spring).
- Pelka-Sługocka, Maria D. i Leszek Sługocki. 1975. *Analiza zgodności między wypowiedziami o faktach, uzyskanymi w wywiadzie kwestionariuszowym a dokumentami*. W: Zygmunt Gostkowski i Jan Lutyński (red.), *Analizy i próby technik badawczych w socjologii*, t. V. Wrocław: Ossolineum.
- Poulain, Michel, Benoit Riandey i Jean-Marie Firdion. 1991. *Enquête biographique et registre belge de population: une confrontation des données*. „Population”, nr 1.
- Sawicki, F., M. Wysocki, L. Twardowska i Z. Obrzut. 1969. *Ocena zgodności pomiędzy informacjami o niektórych chorobach układu oddechowego i serca uzyskanymi w wywiadach i danymi o tych chorobach w dokumentacji w przychodniach*. „Zdrowie Publiczne”, nr 1.
- Silberstein, A. 1989. *Recall effects in the U. S. Consumer Expenditure Survey*. „Journal of Official Statistics”, vol. 5, nr 2.
- Sudman, Seymour i Norman M. Bradburn. 1973. *Effects of time and memory factors on response in surveys*. „Journal of American Statistical Association”, vol. 64, nr 344.
- Sztabiński, Franciszek. 1995. *Kontrola w badaniach surveyowych: pytania i odpowiedzi*.

- „Ask” nr 1.
- Sztabiński, Franciszek. 1997. *Ankieta pocztowa i wywiad kwestionariuszowy*. Warszawa: Wydawnictwo IFiS PAN.
- Sztabiński, Franciszek. 1999. *Kontrola w badaniach ilościowych: ocena pracy ankietowanych czy ocena wartości uzyskanych materiałów?* W: Henryk Domański, Krystyna Lutyńska i Włodzimierz A. Rostocki (red.), *Spojrzenie na metodę. Studia z metodologii badań socjologicznych*. Warszawa: Wydawnictwo IFiS PAN.
- Sztabiński, Franciszek i Paweł B. Sztabiński. 1994. *Polscy respondenci początku lat 90.* „Studia Socjologiczne”, nr 2.
- Sztabiński, Paweł B. 1997. *Ankieterzy i ich respondenci*. Warszawa: Wydawnictwo IFiS PAN.
- Sztaudynger-Kaliszewicz, Anna. 1966. *Badania wiarygodności zeznań w spisie powszechnym*. W: Zygmunt Gostkowski i Jan Lutyński (red.), *Analizy i próby technik badawczych w socjologii*, t. I. Wrocław: Ossolineum.
- Weaver, David A. 2000. *The accuracy of survey reported marital status: evidence from survey records matched to social security records*. „Demography”, vol. 37, nr 3.
- Weiss, Carol H. 1968-69. *Validity of welfare mothers' interview responses*. „Public Opinion Quarterly” (winter).

## WHAT DO WE KNOW ABOUT RESPONDENT? SIZE AND SOURCES OF ERRORS IN BACKGROUND VARIABLES

Questions concerning socio-demographic features of respondent (so-called background variables) are built into each questionnaire schedule. One should assume that asking them we receive information with error. This analysis aims at showing size of this error and its sources. I define error in terms of difference between answers given by the same respondents to the same questions in two interviews. Data comes from control research in national surveys, based on random samples, carried out in 1997–2000. All in all, it includes 7029 randomly selected cases from 83 surveys. The basic conclusion says about widespread scope of error. However, it varies in case of different variables. Errors result from: weaknesses of memory of respondents, tendency to give “socially desirable” answers, and lack of knowledge (e.g. about incomes of the family members). Results of this analysis are discussed against the background of comparable data coming from other research.