

CZĘŚĆ II. ANALIZY WERYFIKACYJNE W BADANIACH Z ZASTOSOWANIEM WYWIADU KWESTIONARIUSZOWEGO

JAN LUTYŃSKI

ANALIZY WERYFIKACYJNE W BADANIACH Z ZASTOSOWANIEM WYWIADU KWESTIONARIUSZOWEGO, ICH RODZAJE I MOŻLIWOŚCI*

„Prowadzący badania może przedstawić swojemu odbiorcy liczbowe oszacowanie tylko tego błędu, którego źródłem jest dobór próbek. Nie jest w stanie tego zrobić w stosunku do wszystkich rodzajów błędów. Jest to słabość i jednocześnie pole do owocnych poszukiwań w zakresie metodologii badań kwestionariuszowych. Działalność na tym polu być może nie wprawia w podniecenie ani nie obiecuje dramatycznych rezultatów. Posiada jednak podstawowe znaczenie dla wszystkich badaczy”.
C. A. Moser, *Survey Methods in Social Investigations*.

Uwagi wstępne

Studia nad technikami otrzymywania danych, takimi jak wywiad kwestionariuszowy, techniki ankietowe czy zdobywanie dokumentów osobistych w wyniku konkursów, mogą mieć na wi-

* Do artykułu w jego pierwotnej wersji wprowadzono poprawki pod wpływem uwag kolegów i koleżanek z Zakładu Metod i Technik Badań Społecznych UŁ, a zwłaszcza mgr B. Tuchańskiej, za które składam tu podziękowanie.

doku różne cele. Jeden z tych celów stanowi po prostu lepsze zrozumienie procesu, w rezultacie którego uzyskuje się materiały i zawarte w nich informacje. Zrozumienie, a więc pogłębienie wiedzy o tym, jak funkcjonuje dana technika, rzadko jednak bywa celem jedynym wspomnianych studiów. Obok niego badacze-metodologowie zmierzają zwykle również do tego, aby ustalić, kiedy, w jakich sytuacjach dana technika jest skuteczna, a kiedy zawodzi, jakie są źródła niepowodzeń i jakie mogą być sposoby ich usunięcia. Chodzi więc również o wykrycie słabości i udoskonalenie danej techniki, które pozwoli zapobiec błędom lub je zminimalizować.

Wymienione cele niewątpliwie dominują obecnie w studiach, poświęconych różnym technikom. Nie są one jednak jedyne, uzupełniają je niekiedy cele do pewnego stopnia odmienne. Występowanie ich wiąże się z faktem, że współcześni socjologowie-empirycy w sformułowaniach wniosków z badań lub w ich uzasadnieniach odwołują się do liczb. Charakteryzując społeczne zjawiska przy użyciu liczb dokonują pomiarów, których przedmiot stanowią społeczne zbiorowości, czasem także składające się na nie jednostki, niekiedy zaś systemy społeczne i różne aspekty ich funkcjonowania.

Posługiwanie się liczbą i pomiarem upodabnia socjologię do bardziej zaawansowanych nauk przyrodniczych i niektórych społecznych. To upodobnienie nie jest jednak całkowite. Świadczy o tym między innymi następujący fakt, który można uznać za symptomatyczny dla pomiaru w społecznych badaniach prowadzonych przez socjologów. Zastosowaniu przez nich pomiaru nie towarzyszy sformalizowana, liczbowa ocena jego wartości, co stanowi regułę w wielu innych naukowych dyscyplinach. W związku z tym można sformułować wniosek, że jednym z głównych celów studiów nad technikami otrzymywania danych powinno być opracowanie metod otrzymywania takich ocen wartości danych.

Stwierdzenia powyższe wymagają krótkiego komentarza. W bardziej zaawansowanych naukach dąży się do tego, aby wartość pomiaru była w jakiś sposób oszacowana. Sposoby są zresztą bardzo różne. Niekiedy ocena dotyczy samego wyniku. Wówczas zazwyczaj na podstawie empirycznych danych drogą obliczeń

ustala się, jaki jest — z określonym, wysokim prawdopodobieństwem — rozmiar ewentualnego błędu pomiaru, w jakich mieści się on granicach¹. Niekiedy ocena odnosi się do wartości metody pomiaru w danych warunkach, zwłaszcza jej trafności (czy i jak dokładnie przy jej użyciu mierzy się to, co chce się mierzyć) i rzetelności (czy i jak dokładnie przy jej użyciu mierzy się to, co się mierzy). Takich ocen, znajdujących również wyraz w liczbach, dokonuje się w odniesieniu do dobrze opracowanych testów psychologicznych².

Wydawać się może, że socjologowie formułują także podobne oceny. Korzystając z metod statystyki określają oni przecież często, jaki jest poziom istotności danego wyniku. Wbrew pozorom nie stanowi to ekwiwalentu liczbowej oceny wartości pomiaru. Testy istotności nie charakteryzują bowiem, czy pomiar danych zjawisk został dokonany dobrze, lecz jedynie to, z jakim prawdopodobieństwem wynik pomiaru odnoszący się do próbki można rozszerzyć na zjawiska w całej populacji, tj. rozszerzyć z jednej zbiorowości na drugą. W związku z tym pomiar zjawisk w zbiorowości stanowiącej próbkę może być obciążony znacznym błędem, niezależnie od tego, że ze względu na wielkość i sposób doboru tej próbki dany wynik może być istotny na bardzo wysokim poziomie dla całej populacji.

Wartość pomiarów dokonywanych w badaniach socjologicznych oceniana jest więc, co najwyżej, tylko częściowo — w zakresie błędów, których źródłem jest dobór próbki (*sampling errors*). Nie ma natomiast efektywnych sposobów pozwalających oszacować rozmiar błędów, których źródłem jest niedoskonałość zastosowanej techniki otrzymywania danych (*non sampling errors, response errors, measurement errors*). W rezultacie nie można więc oszacować wielkości błędu w całości³.

¹ Por. m. in. E. Wilson, *Wstęp do badań naukowych*, Warszawa 1964, zwłaszcza rozdz. 9: „Uchyby pomiarowe”.

² M. Choynowski, *Elementy teorii testów psychologicznych*, „Przegląd Psychologiczny”, 1959, nr 3, s. 161-167.

³ M. H. Hansen, W. N. Hurwitz, W. G. Madow, *Sample Survey Methods and Theory*, Nowy Jork, Londyn, Sidney 1953, t. I: *Methods and Applications*, s. 82 i następne, t. II: *Theory*, s. 280-281. Por.

Następstwa tego stanu rzeczy są bardzo poważne. Socjologowie na ogół nie wiedzą, czy podstawa empiryczna ich generalizacji jest pewna, przy czym stopień tej niepewności nie da się z reguły określić. Świadomość tego stanu rzeczy wpływa na nich samych. Wątpiąc często w wartość wyników swoich badań przywiązują do nich niejednokrotnie mniejszą wagę niż przedstawiciele innych nauk, a w związku z tym są może bardziej skłonni, aby akceptować sądy o rzeczywistości nie oparte na badaniach. Ponieważ mają wątpliwości, czy ich wyniki są pewne, trudniej jest im formułować dalej idące zalecenia praktyczne, z czym zawsze wiąże się duża społeczna odpowiedzialność. W przypadku, gdy badacze posługujący się różnymi technikami dochodzą do odmiennych rezultatów — co zdarza się bynajmniej nierzadko — nie umiemy na ogół ocenić, jaki wynik jest lepszy, na którym można bardziej polegać. Utrudnia to zarówno formułowanie wniosków praktycznych, jak i weryfikację rozbudowanych systemów twierdzeń i teorii, nie mówiąc już o przedstawieniu opisu zjawisk. Oczywiście w tej sytuacji nie można także ocenić dokładnie wartości samych technik w ich poszczególnych zastosowaniach ani skorygować wniosków.

Z zastosowaniem pomiaru i w ogóle ujęć liczbowych w badaniach socjologicznych wiążą się więc istotne braki, nieobce zresztą badaniom innych dyscyplin we wcześniejszych stadiach ich rozwoju. Braki te mogą być zapewne usunięte w rezultacie wielu metodologicznych studiów⁴. Byłyby to — i są zresztą — studia

także C. A. Moser, G. Kalton, *Survey Methods in Social Investigation*, Second Edition, Londyn 1971, s. 378, oraz G. Lissowski, *Z zagadnień doboru próby*, [w:] *Metody statystyczne w socjologii*, red. K. Szaniawski, Warszawa 1968, s. 68; F. Mosteller, *Nonsampling Errors*, [w:] *Encyclopedia of the Social Sciences*, ed. D. L. Sills, vol. 5, The Macmillan Co and the Free Press, 1968, s. 113 - 114.

⁴ Wydaje się jednak, że usunięcie tych braków jest w socjologii trudniejsze niż w innych naukach, z różnych zresztą powodów. Jeden z takich powodów stanowią obawy, że określenie wielkości błędów podważy zaufanie do wyników badań społecznych, których pozycja nie wszędzie jest ustabilizowana, że jeśli wykaże się, iż wyniki te są obciążone dużymi

nad technikami otrzymywania informacji, zmierzające do liczbowego oszacowania wartości tych ostatnich i wyników badań w całości. Taki cel refleksji i badań nad technikami uzyskiwania danych jest odmienny od poprzednio wyróżnionych, pozostaje jednak w ścisłym z nimi związku. W istocie nie można bowiem ustalić, że dana technika posiada jakieś wady czy słabości, gdy nie umiemy ocenić rezultatu jej zastosowania. Oczywiście ocena ta nie musi być wyrażona liczbowo, jednak takie bardziej ścisłe oszacowania byłyby bardziej użyteczne niż oceny w kategoriach: lepszy — gorszy. Realizacja sformułowanego przed chwilą celu metodologicznych studiów jest więc także konieczna dla pełnej realizacji celu, jakim jest udoskonalenie sposobów otrzymywania informacji.

Opracowanie efektywnych metod ścisłej oceny wartości pomiarów w badaniach społecznych prowadzonych przez socjologów stanowi niewątpliwie jeden z ważnych celów badań metodologicznych nad technikami otrzymywania danych. Nie jest to jednak cel równie ważny w przypadku różnych technik i sposobów badania. Nie wydaje się na przykład słuszne, aby go stawiać przed refleksją metodologiczną, dotyczącą studiów monograficznych społeczności lokalnych, w których ujęcia liczbowe nie odgrywają na ogół podstawowej roli, w których zbiera się wiele różnorodnych danych za pomocą różnych, często niestan-

błędami, będzie trzeba zrezygnować z udzielania na podstawie badań odpowiedzi na ważne społeczne problemy, których rozwiązanie uważa się za zadanie socjologii itp. Zapewne główny powód tkwi jednak w trudnościach znalezienia sposobów oceny błędów.

Na marginesie tych uwag warto przypomnieć stwierdzenie podane przez Hymana: „Nie należy utożsamiać braku wiedzy o błędzie z brakiem błędu” oraz anegdotę przytoczoną przez Wilsona: „Raczej nie należy przyswajając sobie poglądów pewnego pracownika, który miał pomierzyć kabel używany przy doświadczeniach podwodnych i na pytanie, czy dla sprawdzenia powtórzył pomiary, odpowiedział: «Nie, bo jeśli zmierzymy coś raz, to wiemy, co jest, a jeśli zmierzymy ponownie i nie dostaniemy tej samej wartości, to nie wiemy, czego się trzymać»”. Powtórzenie pomiaru, jak wiadomo, jest podstawą do sformułowania wyniku i określenia wielkości błędów w wielu naukach (H. Hyman i inni, *Interviewing in Social Research*, Chicago 1954, s. 4; E. Wilson, *op. cit.*, s. 389).

daryzowanych sposobów. Na plan pierwszy wysuwają się natomiast tutaj techniki, w których za pomocą w zasadzie jednakowych zabiegów uzyskuje się jednorodne, bardzo liczne dane, opracowane za pomocą statystycznych metod. Taką technikę stanowi przede wszystkim wywiad kwestionariuszowy, dominujący obecnie w badaniach socjologicznych, tj. wywiad, w którym w celu uzyskania wielu informacji licznym respondentom zadaje się jednakowe pytania w zasadzie w jednakowy sposób.

Postawienie omawianego celu przed metodologicznymi studiami nad tą techniką wydaje się nie tylko najbardziej potrzebne. Można także przypuszczać, że studia tego rodzaju w tym przypadku przyniosą najlepsze rezultaty, o czym będzie jeszcze mowa. Nie oznacza to oczywiście, iż rezultatów tych należy oczekiwać szybko i że można osiągnąć bez wielkiego wysiłku ze strony licznych badaczy-metodologów.

W niniejszym artykule nie stawia się, rzecz jasna, zadania, aby zrealizować wskazany przed chwilą dalekosiężny cel. Temat artykułu nawiązuje jednak do perspektywy, jaką stwarza ów cel. Artykuł dotyczy bowiem analiz mających na widoku jakąkolwiek ocenę wyników otrzymywanych w wywiadzie kwestionariuszowym. Niezależnie bowiem od tego, że w socjologii — i nie tylko zresztą w socjologii — nie opracowano dotychczas efektywnych metod ścisłej oceny wartości tych wyników, wielu badaczy i metodologów stara się je poddać analizie i na tej podstawie chce dojść do jakichś oceniających wniosków. Sposoby stosowane w tych analizach są przy tym bardzo różne. Artykuł zmierza do tego, aby je scharakteryzować, a zwłaszcza wskazać na ich założenia i możliwości.

Zadaniem artykułu nie jest jednak wyczerpująca prezentacja różnych analiz tego typu, lecz ich schematyczna charakterystyka dostarczająca podstaw do określania ich wartości dla uzyskania bardziej ścisłej oceny wyników badań. Materiał tej charakterystyki stanowią przykłady analiz, w których badacze-metodologowie oceniają wartość wyników badań. Między innymi są to analizy zamieszczone w poprzednich tomach niniejszej serii, a także analizy omawiane w publikacjach, w których poddano je

pewnej systematyzacji. Artykuł stanowi jednocześnie wprowadzenie do drugiej części niniejszego tomu *Analiz i prób*, poświęconej omówieniu rozbudowanej próby dokonania oceny wartości rezultatów wywiadu kwestionariuszowego. Były to badania z zastosowaniem tego wywiadu, zorganizowane specjalnie w tym celu. W badaniu tym wykorzystano niektóre myśli zawarte w artykule, choć przeprowadzono je przed jego napisaniem.

Założenia i główne rodzaje analiz weryfikacyjnych

Analizy zmierzające do sformułowania jakiegokolwiek oceny wiarygodności wyników liczbowych nazywać będziemy analizami weryfikacyjnymi lub weryfikacją. Cechy tych analiz, także te, ze względu na które można wyróżnić wśród nich określone rodzaje i typy, nawiązują do cech badań z użyciem wywiadu kwestionariuszowego. Są zwłaszcza cechy następujące:

1. Wyniki badań z użyciem wywiadu kwestionariuszowego, które dostarczają rozwiązania problemu, podobnie jak ten problem odnoszą się zawsze do zbiorowości. Niezależnie od charakteru tych wyników, a więc od tego, czy dotyczą one rozkładu cech w zbiorowości, związku tych cech, układu tych związków, przekształceń zbiorowości w czasie itp. — oparte są one na sądach o jednostkach tworzących daną zbiorowość. Najczęściej, choć nie wyłącznie, są to pojedynczy ludzie. Sądy te przenoszą informacje o jednostkach badania, o tym, do jakiej kategorii badanych zjawisk należy je zaliczyć ze względu na ich właściwości. Dochodzi się do tych sądów w rezultacie zastosowania danej techniki otrzymywania informacji. W konsekwencji bezpośrednich źródeł błędu, którym — ewentualnie — obciążony jest wynik dotyczący zbiorowości, szukać należy w błędach jednostkowych, w tym, iż fałszywe są zdania przenoszące informację o jednostkach badania, o ich przynależności do danych klas badanych zjawisk.

2. W badaniach społecznych, przynajmniej w obecnej chwili, wiele często nieprzewidywalnych trudności następcza wielokrotne powtarzanie tych samych poszukiwań, tj. poszukiwań pro-

wadzonych w tym samym celu, przy użyciu tych samych technik i w stosunku do tego samego przedmiotu, które powinny przynieść te same wyniki. Zastosowanie takiej procedury utrudnia zmienność społecznych zjawisk, występującą także pod wpływem samych poznawczych zabiegów, koszty powtórnych badań itp. W tej sytuacji w praktyce nie można zastosować schematu, wypracowanego w naukach przyrodniczych i polegającego na ustalaniu wyniku dokładnego (prawdziwego) z określeniem prawdopodobieństwa popełnienia błędu w danych granicach. Nie jest to także celowe ze względu na to, że nie ma podstaw, aby przyjąć, jak to się robi w tych naukach, że błędy, którymi obciążone są poszczególne wyniki, mają charakter przypadkowy, losowy, i stąd można je wyeliminować czysto statystyczną drogą (obliczenie średniej wartości wyniku, granic i prawdopodobieństwa błędu). Przekonanie dotyczące charakteru błędów w badaniach społecznych, o tym, że są to także błędy systematyczne, jest słuszne także w odniesieniu do wyników na poziomie zbiorowości, co można uzasadnić empirycznie.

3. W badaniach społecznych z użyciem wywiadu kwestionariuszowego najczęstsza jest sytuacja, kiedy jedną informację o danej jednostce badania, posiadającą znaczenie dla ustalenia wyniku, a więc i rozwiązania merytorycznego problemu, uzyskuje się na podstawie odpowiedzi na jedno pytanie z kwestionariusza. Zazwyczaj jest to także jedyna informacja tego rodzaju, którą przynosi odpowiedź na to pytanie, choć niekiedy z jednym pytaniem związane są dwie lub więcej informacji⁵. Zdarza się

⁵ Należą do nich pytania wprost o opinie, na które pewne odpowiedzi przynoszą dwie informacje (tj. rozstrzygnięcia jednostkowych pytań badacza): że respondent jest zwolennikiem jednego z poglądów przewidywanych przez pytanie lub że nim nie jest (brak opinii) i że jest to dany pogląd. Szereg informacji interesujących badacza przynoszą zwykle pytania o dużej ilości możliwych odpowiedzi, spośród których respondent powinien wybrać wszystkie odpowiednie, np. pytanie typu: „Które z następujących przedmiotów i urządzeń (tu wyliczanie) znajdują się w Pana(i) mieszkaniu?” — gdy badacza interesują kolejno poszczególne przedmioty

także i tak, że badacz łączy informacje uzyskane na podstawie odpowiedzi na kilka pytań z kwestionariusza, uzyskując w ten sposób jednostkową informację wyższego rzędu. Na przykład, przy ustalaniu stopnia religijności jakiejś osoby wykorzystuje się informacje dotyczące wykonywania przez nią poszczególnych praktyk religijnych, jej różnych przekonań odnoszących się do spraw wiary i moralności, stopnia zaangażowania w sprawy religii itp. Każda z tych szczegółowych informacji posiada z reguły znaczenie także dla ustalenia jakiegoś wyniku w zakresie danego szczegółowego zagadnienia (np. dla ustalenia, jaki jest stopień rozpowszechnienia praktyk religijnych) w odniesieniu do całej zbiorowości. Niekiedy te szczegółowe informacje czerpie się nawet z różnych kwestionariuszy⁶. Rzadko natomiast występuje sytuacja, kiedy owe cząstkowe informacje dotyczą tylko tego, jak respondenci odpowiadali na dane pytanie, gdy nie interesują one badacza z punktu widzenia żadnych szczegółowych merytorycznych zagadnień, lecz tylko ze względu na udział w informacji jednostkowej wyższego rzędu, która dopiero przyczynia się do rozwiązania merytorycznego problemu. Ta sytuacja charakterystyczna jest, jak wiadomo, dla badań testowych w psychologii, gdzie ze względu na problem ważny jest dopiero końcowy rezultat uzyskany przez daną jednostkę, nie zaś odpowiedź

lub ich układy. Często więcej niż jedną informację badacz otrzymuje na podstawie odpowiedzi na pytanie otwarte, zwłaszcza żądające przedstawienia poglądów lub narracji. Pojęcie poszukiwanej jednostkowej informacji (jednostkowego pytania badacza) wykorzystałem już w moim artykule pt. *Analiza procesu otrzymywania informacji w badaniach z zastosowaniem wywiadu kwestionariuszowego*, [w:] *Analizy i próby technik badawczych w socjologii*, red. Z. Gostkowski, J. Lutyński, t. IV, Wrocław 1972, s. 73-86. Artykuł niniejszy stanowi kontynuację podjętych tam rozważań.

⁶ Np. gdy informacje uzyskane od męża i żony, a dotyczące ich zadowolenia z małżeństwa, badacz łączy w jedną informację o tym, czy małżeństwo jest szczęśliwe, czy nie. Dwie zgodne opinie o zadowoleniu (lub niezadowoleniu) mogą być traktowane jako wskaźnik tego, iż małżeństwo jest szczęśliwe (lub nieszczęśliwe), opinie niezgodne, że nie jest ani szczęśliwe, ani nieszczęśliwe.

respondenta na jedno pytanie, sama dla siebie, w oderwaniu od pozostałych jego odpowiedzi lub rozwiązań innych zadań⁷.

Z omawianą obecnie łączy się jeszcze inna cecha badań społecznych z użyciem wywiadu kwestionariuszowego, przeciwstawiająca je badaniom testowym w psychologii. Podczas gdy rezultatem zastosowania testu, składającego się z reguły z wielu pytań czy zadań (*items*), jest w zasadzie jeden i polega on na zaliczeniu badanego do odpowiedniej pozycji na skali, wywiad kwestionariuszowy przynosi bardzo dużo rezultatów, mniej więcej tyle, ile jest pytań w kwestionariuszu. Obydwie te cechy pociągają za sobą to, że w interesujących nas studiach weryfikacyjnych trudno jest stosować metody używane przy ocenie wartości psychologicznych testów. Ostatnia cecha w powiązaniu z faktem, iż liczne informacje, które otrzymuje się za pomocą wywiadu kwestionariuszowego, mogą być narażone na zniekształcenie (błąd) w różnym stopniu i z różnych powodów, powoduje, że trudno jest ustalić jakąś jednolitą ścisłą ocenę kwestionariusza w całości.

4. Następne ważne dla nas cechy badań z użyciem wywiadu kwestionariuszowego dotyczą jednolitego w znacznej mierze charakteru procesów poznawczych i innych prowadzących do pojawienia się jednostkowych informacji. Na każdy taki proces składa się, z jednej strony, postępowanie członków ekipy badawczej, jakimi są sam badacz, ankieter i koder, z drugiej — zachowania i przeżycia respondentów. Tym, co zapewnia jednolitość postępowania badawczego oraz — w pewnych granicach — reakcji respondentów, jest jednolitość różnego rodzaju narzędzi badania (kwestionariusz, instrukcje), które same zresztą tworzą część omawianego procesu. Tę cechę interesujących nas tutaj badań można wykorzystać w analizach weryfikacyjnych, zresztą w łączności z cechą inną. W tych badaniach badacz ma — lub może mieć — wiedzę o przebiegu całości procesu, prowadzącego do

⁷ Por. M. Choynowski, *op. cit.* Analogiczna sytuacja występuje bardzo rzadko w badaniach socjologicznych. Przykład jej znajdujemy m. in. w pracy: P. F. Lazarsfeld, W. Thielens, *The Academic Mind. Social Scientists in a Time of Crisis*, Glencoe 1958, s. 79.

otrzymania jednostkowej informacji, której w odniesieniu do ważnych jego ogniów nie może uzyskać w przypadku stosowania niektórych innych technik, na przykład w pełni anonimowej ankiety rozsyłanej respondentom.

5. Ostatnia cecha omawianych badań, która nas tu interesuje, wiąże się z faktem, iż niektóre uzyskiwane w nich jednostkowe informacje można zdobyć również innymi drogami, na przykład za pomocą obserwacji, zwłaszcza uczestniczącej (w tym także informacje tego typu, czy ktoś jest religijny), następnie ankiety, wykorzystania istniejących dokumentów, organizowania specjalnych badań psychologicznych i innych. Te inne drogi bardzo często można z góry uznać za lepsze. W niektórych przypadkach wywiad kwestionariuszowy jest więc techniką zastępczą, używaną nie dlatego, że przynosi bardziej pewne rezultaty, lecz dlatego, że przy jego zastosowaniu można zdobyć dużo różnorodnych informacji, odnoszących się do dużej, dowolnie wybranej zbiorowości, w stosunkowo krótkim czasie.

Analiza cech wymienionych powyżej w pierwszych trzech punktach skłania do przyjęcia wniosku, że w odniesieniu do badań prowadzonych przy użyciu wywiadu kwestionariuszowego nie mogą być zastosowane — przynajmniej bez daleko idących zmian — procedury przyjęte w innych naukach. W rzeczy samej procedur tych nie stosuje się w przeprowadzanych analizach weryfikacyjnych, mimo że związane z nimi pojęcia (trafność, rzetelność, błąd losowy i systematyczny) są omawiane w podręcznikach badań. Uwagi powyższe sugerują także, że w tych analizach, podobnie jak w samych badaniach z użyciem wywiadu kwestionariuszowego podstawową rolę powinny pełnić sądy o jednostkach badania. Poza tym końcowa uwaga zawarta w punkcie 3, głosząca, że w jednej analizie nie można poddać ocenie całego kwestionariusza, nasuwa wniosek, iż przedmiotem takiej oceny mogą być tylko poszczególne informacje i odpowiadające im pytania.

Określone wnioski, najważniejsze ze względu na cele niniejszego artykułu, wiążą się z ostatnimi punktami przedstawionych uwag. Sprowadzają się one do przekonania, że w analizach weryfikacyjnych można stosować dwa zasadnicze sposoby postępo-

wania, wyróżnione ze względu na podstawę oceny. Jeden z nich polega na tym, iż ocenia się wynik, opierając się na ocenie jednostkowych procesów otrzymywania danych, które do niego doprowadziły. Przy zastosowaniu sposobu następnego wynik ocenia się na podstawie zestawienia ze sobą rezultatów procesów otrzymywania danych, tj. informacji, które one przyniosły, osiągniętych różnymi drogami. Sposób ten posiada specjalne znaczenie w badaniach socjologicznych ze względu na to, o czym była mowa w ostatnim punkcie powyższych uwag, tj. w związku z tym, iż wywiad kwestionariuszowy stanowi niekiedy technikę zastępczą w stosunku do trudnych do zastosowania, doskonalszych technik uzyskiwania tych samych informacji, względnie informacji pozostających z nimi w bliskim związku. Sposób wymieniony poprzednio może być zastosowany w analizach oceniających wyniki otrzymane przy zastosowaniu wywiadu kwestionariuszowego, ponieważ, jak była o tym mowa w punkcie 4, procesy w rezultacie których pojawiają się jednostkowe informacje, są ujednolicone i badacz może zdobyć o nich wiedzę. Podstawowe dla tego sposobu jest założenie, że o jakości każdego produktu — a jednostkowe informacje i wynik końcowy stanowią niewątpliwie taki produkt — decyduje to, co dzieje się w procesie jego formowania. Jest to założenie, dla którego trudno znaleźć kryteria obalenia⁸.

Biorąc za podstawę obydwie sposoby można wyróżnić dwa zasadnicze typy analiz weryfikacyjnych: weryfikację wewnętrzną i zewnętrzną. Do weryfikacji wewnętrznej zaliczamy analizy, w których jednostkowe informacje i wynik w całości ocenia się głównie lub wyłącznie na podstawie danych o procesach, których rezultaty stanowią otrzymane informacje. Są to z reguły dane, które posiadają znaczenie tylko dla analizy metodologicznej i nie mogą być wykorzystane w merytorycznym opracowaniu. Przykład takich, częściowych zresztą, analiz stanowi każda kontrola postępowania ankietatorów, w wyniku której ocenia się negatywnie i odrzuca informacje dostarczane przez kwestionariusz wy-

⁸ Stwierdzenie o podobnym charakterze w ten właśnie sposób określa J. Giedymin, *Problemy, założenia, rozstrzygnięcia. Studia nad logicznymi podstawami nauk społecznych*, Poznań 1964, s. 108.

pełniony nie na podstawie rozmowy z respondentem. Innym przykładem weryfikacji wewnętrznej będzie analiza, w której ustalono, że duży odsetek respondentów nie rozumie tak samo jak badacz jakiegoś terminu, kluczowego do zrozumienia pytania i stąd nie można zaakceptować informacji otrzymanych za jego pomocą. Warto tutaj zresztą od razu zauważyć, że w niektórych analizach to, że proces miał przebieg nieprawidłowy, ustala się raczej pośrednio, na przykład na podstawie informacji, iż respondent, który udzielał odpowiedzi, był nietrzeźwy lub był bardzo zmęczony i nastawiony tylko na to, aby wywiad szybko zakończyć itp.

Analizy, w których ocenę wyniku ustala się — wyłącznie lub przede wszystkim — na podstawie zestawienia rezultatów osiągniętych różnymi drogami, zaliczamy do typu określonego tutaj jako weryfikacja zewnętrzna. Zewnętrzna — ponieważ w przeciwstawieniu do poprzedniej sięga się tutaj nie do informacji o samych procesach dostarczających informacji, lecz do danych zewnętrznych. W przeciwstawieniu do danych wykorzystywanych w analizach poprzedniego typu, informacje te posiadają — lub mogą posiadać — znaczenie dla rozwiązania merytorycznych problemów interesujących badacza. Przykład takich analiz stanowi weryfikacja danych dotyczących wysokości zarobków, uzyskanych w wywiadzie kwestionariuszowym, które porównuje się z danymi z dokumentów finansowych zakładów pracy. W tym przypadku, podobnie jak w wielu innych, dany rezultat zestawia się z innym rezultatem, osiągniętym inną drogą. Dzieje się tak także wtedy, gdy badacz ocenia wynik opierając się na ugruntowanej, choćby zdroworozsądkowej wiedzy w tym zakresie.

To zestawienie ocenianego wyniku z innym, nie ocenianym, osiągniętym przy zastosowaniu innego sposobu badania, nie jest jednak dla analiz omawianego obecnie typu konieczne. Zaliczyć do niego można również i takie analizy, w których ocenia się dany wynik, zestawiając ze sobą rezultaty częściowe, otrzymane różnymi drogami, nawet za pomocą jednego w zasadzie sposobu. Przykład stanowią tu studia, w których porównuje się rozkłady odpowiedzi uzyskane przez różnych ankietatorów w celu oceny ca-

tego wyniku. Niezależnie od tych różnic zestawienie ze sobą różnych rezultatów otrzymanych na różnych drogach stanowi podstawę ich oceny.

Mimo że posługiwanie się terminami „weryfikacja wewnętrzna” i „zewnętrzna” w przyjętym tu rozumieniu posiada uzasadnienie, nie są to bynajmniej znaczenia jedyne, które można im nadać⁹. Trzeba także podkreślić, iż oznaczają one typy analiz, a nie ich kategorie wyraźnie od siebie oddzielone i tworzące razem ich poprawny podział. Można bowiem wskazać analizy, zresztą raczej o marginesowym znaczeniu, które nie należą do żadnego z tych typów lub w których badacz wykorzystuje zarówno stwierdzenia dotyczące charakteru procesów formowania informacji, jak i rezultaty tych procesów¹⁰.

⁹ H. Hyman używa terminów „wewnętrzne i zewnętrzne sprawdzenie” (*check*) w innym znaczeniu. Podstawy odróżnienia stanowi to, czy informacje służące kontroli uzyskano w tym samym czy w innym, niezależnym badaniu (H. Hyman, *Survey Design and Analysis*, Glencoe 1955, s. 151, 163). W podobnym znaczeniu o weryfikacji wewnętrznej i zewnętrznej mówią P. F. Lazarsfeld i W. Thielens (*op. cit.*, s. 416).

W artykule moim pt. *Projekt analizy oceniającej wiarygodność wyników Międzynarodowych Badań Poznawczych nad Wykorzystaniem Opieki Zdrowotnej* weryfikacją wewnętrzną nazwałem analizę polegającą na sprawdzeniu, czy występuje zgodność rezultatu w przypadku odpowiedzi na różne pytania w tym samym kwestionariuszu, które miały przynieść tę samą informację (tzw. pytania kontrolujące się). Por. *Analizy i próby...*, t. IV, 1972, s. 265. Obecne zastosowanie omawianych terminów ma inne, szersze znaczenie.

¹⁰ W wyjątkowych przypadkach można ocenić wyniki nie odwołując się do żadnych dodatkowych danych. Tak jest na przykład w przypadku pytania skierowanego do studentów jakiegoś kierunku i roku studiów: „Czy należy Pan(i) do tej połowy studentów waszego roku, która ma więcej czy mniej nieobecności na wykładach?” przy poszukiwanej informacji zgodnej z treścią pytania. Jeśli więcej niż połowa odpowie: „Mniej” (lub „Więcej”), wówczas liczba błędnych informacji jest co najmniej taka, ile wynosi różnica między tym wynikiem a połową liczebności grupy. Pytanie wydaje się interesujące dla rozważań metodologicznych również z innych powodów. Analiza wyników, które ono przynosi, może doprowadzić do oceny siły tendencji do zniekształcenia — a więc i błędów — w danym zakresie i w danej grupie.

I. Weryfikacja zewnętrzna, jej rodzaje i możliwości

1. Zasadnicze typy weryfikacji zewnętrznej

Omówienie poszczególnych typów analiz weryfikacyjnych rozpoczniemy od weryfikacji zewnętrznej. Analizy te są zróżnicowane i można wyróżnić wśród nich grupy ze względu na różne zasady¹¹. Mając na myśli możliwość uzyskania sformalizowanej oceny wyników badań, celowe wydaje się wyróżnienie kilku typów. Przede wszystkim trzeba wyodrębnić takie analizy, w których przedmiot oceny stanowi początkowo każda jednostkowa informacja, a potem wynik odnoszący się do zbiorowości. Przykładowo — rezultaty badania kwestionariuszowego, w którym ustalono, jaki jest odsetek osób piśmiennych i niepiśmiennych w danej zbiorowości przy użyciu pytania: „Czy umie Pan(i) pisać?”, można zweryfikować na poziomie jednostkowym, prosząc badanych o napisanie kilku wybranych zdań. Porównując rezultaty otrzymane dwiema różnymi drogami ocenia się wyniki odnoszące się do zbiorowości.

Załóżmy jednak, iż zdarzyło się tak, że w analogicznym badaniu nad piśmiennością nie można zestawić zaklasyfikowań jednostek na podstawie odpowiedzi w wywiadzie z zaklasyfikowaniami w analizie sprawdzającej. Wówczas weryfikacja — przy wykorzystaniu informacji jednostkowych — dotyczyłaby tylko wyniku odnoszącego się do zbiorowości. Taka sytuacja występuje na przykład wtedy, gdy wyniki wywiadu kwestionariuszowego zestawiamy z wynikami ankiety w pełni anonimowej, na którą odpowiedzi udzielili ci sami, jednak nie zidentyfikowani personalnie respondenci.

¹¹ Por. scharakteryzowane w przypisie 9 stanowisko w tej sprawie Hymana. Jego podział, bynajmniej zresztą nie w pełni konsekwentny, krzyżuje się z przyjętym tutaj rozgraniczeniem według typów. Zasada podziału wprowadzona przez Hymana uwzględnia silnie przez niego akcentowaną okoliczność, iż w analizach wykorzystujących dane z tego samego badania, a zwłaszcza kwestionariusza, prawdopodobieństwo zniekształceń płynących z tego samego źródła jest większe.

Nie jest to jednak przypadek typowy. Typowa jest natomiast taka sytuacja, gdy poszukując tej samej informacji inną drogą nie można zbadać tych samych respondentów, lecz inną ich grupę. Ocena wyniku jest wtedy możliwa pod warunkiem, iż badacz wie, jaki dokładnie powinien być stosunek pierwszego wyniku do drugiego. W praktyce chodzi tu o przypadki, gdy obydwa wyniki powinny być identyczne (czy raczej bardzo zbliżone) ze względu na porównywalność grup. Pod tym określeniem rozumie się zazwyczaj sytuacje, kiedy obydwie grupy posiadają identyczne, czy też co najmniej bardzo zbliżone cechy i jedną można uznać za substytut drugiej z punktu widzenia danego wyniku badania. Wówczas wynik otrzymany od grupy kontrolnej powinien być identyczny (bardzo zbliżony). Podobieństwo lub różnice wyników decydują wówczas o pozytywnej względnie negatywnej ocenie obydwóch lub jednego z nich. Przykład takiej analizy stanowi ocena wyników uzyskanych od grup respondentów, z którymi wywiady przeprowadzali różni ankieterzy, przy czym przydział do nich respondentów był losowy. Zgodność lub rozbieżność wyników świadczy o ich wartości.

Dwa wskazane dotychczas ogólne typy analiz weryfikacyjnych, zresztą bardzo wewnętrznie zróżnicowane, o czym będzie jeszcze mowa, występują najczęściej. Obok nich należy wyróżnić jeszcze trzeci, także stosowany w badaniach z użyciem wywiadu kwestionariuszowego. Przytoczymy przykład analizy tego typu z pracy P. F. Lazarsfelda i W. Thielensa pt. *The Academic Mind*¹². Badano w niej wykładowców nauk społecznych na wyższych uczelniach w Stanach Zjednoczonych w połowie lat pięćdziesiątych, a więc jeszcze w okresie „zimnej wojny” i maccarthyzmu. Lazarsfeld podzielił wykładowców na kilka grup ze względu na stopień obawy (*apprehension*). Postawa obawy, zgod-

¹² Jak wspominaliśmy, książka ta zaopatrzona jest w aneksy, w których między innymi przedstawiono analizy weryfikacyjne dotyczące innych wyników z badań. Zawiera także opracowanie D. Riesmana pt. *Some Observations on Interviewing in the Teacher Apprehension Study*, poświęcone ocenie wartości wywiadów.

nie z koncepcją autorów, obejmowała uczucie zaniepokojenia (*worry*) z powodu politycznego nacisku, występującego zresztą w różnych formach, oraz uczucie ostrożności (*caution*), wyrażającą się w różnych zachowaniach (np. powstrzymywanie się od wyrażania w szerszym gronie tych politycznych opinii, które mogą być uważane za kontrowersyjne, odpowiednie przystosowanie tego, co się mówi do studentów i pisać do druku itp.). Podziału tego dokonano na podstawie analizy odpowiedzi na kilka pytań kwestionariusza.

Pragnąc się upewnić, czy podział według stopnia obawy odpowiada rzeczywistości, autorzy wykonali kilka dodatkowych analiz. Jedną z nich oparta była na następującym rozumowaniu. Wyniki (na skali obawy) powinny być wyższe wśród nauczycieli akademickich, którzy mieli rzeczywiste powody do obaw. Mieli je ci, którzy byli rzeczywiście atakowani lub których poprzednia przynależność organizacyjna mogła być uważana za kontrowersyjną. Stopień narażenia na niebezpieczeństwo ujęty w tych kategoriach powinien więc korelować ze stopniem obawy¹³. Ponieważ autorzy mieli do dyspozycji dane odnoszące się do ataków na profesorów i ich poprzedniego politycznego zaangażowania, podzieleni ich na cztery grupy według tych dwóch kryteriów i zbadali, czy istnieje statystyczny związek między stopniem narażenia na niebezpieczeństwo a stopniem obawy. Związek ten okazał się bardzo wyraźny i stąd autorzy stwierdzają: „Obydwie miary weryfikują nasze oczekiwania i implikują, że indeks obawy stanowi sensowny i czuły instrument (pomiaru)”¹⁴. Tym samym autorzy uważają, że przeprowadzona analiza wskazuje na wartość otrzymanego wyniku, tj. podziału zbiorowości na grupy według stopnia obawy. Podobny sposób postępowania weryfikacyjnego autorzy stosują w odniesieniu do podziału wykładowców na tolerancyjnych (*permissive*) i konserwatywnych¹⁵. Lazarsfeld uważa go za ogólnie wskazany przynajmniej w badaniach tego

¹³ *Op. cit.*, s. 88.

¹⁴ *Op. cit.*, s. 89.

¹⁵ *Op. cit.*, s. 130 - 144.

typu¹⁶. Podobne analizy są zresztą stosowane w wielu badaniach¹⁷.

Omawiany typ analizy odznacza się następującymi właściwościami. Weryfikowany wynik odnosi się do rozkładu cechy, o której wiemy z góry, że jest skorelowana w określony sposób z inną. Informacje o występowaniu tej ostatniej uzyskujemy osobno, inną drogą, np. postawieniem innego pytania w tym samym wywiadzie. Sprawdzamy, czy taka korelacja występuje w naszym badaniu. Jeśli tak, wówczas upewniamy się, iż weryfikowany wynik jest wartościowy. W przeciwnym przypadku jego wartość można zakwestionować.

Jak się wydaje, do scharakteryzowanych wyżej typów można zaliczyć znakomitą większość przeprowadzonych faktycznie analiz, w których wartość wyniku ocenia się zestawiając ze sobą informacje uzyskane różnymi drogami, przy czym nie są to informacje odnoszące się do samego procesu formowania odpowiedzi. Te trzy typy dadzą się uporządkować w następujący sposób.

Analizy weryfikacyjne, należące do pierwszego typu, można przeprowadzić, gdy spełnione są dwa warunki:

1. Informacje uzyskane różnymi drogami odnoszą się do jednostek badania z tej samej zbiorowości, zidentyfikowanych ponadto przez badacza.

2. Są to informacje analogiczne pod względem treści, tj. stanowią one rozstrzygnięcie tych samych pytań (oczywiście odpowiedzi udzielone na te pytania w odniesieniu do danej jednostki mogą być różne).

Analizy typu drugiego nie zakładają, iż zrealizowany jest warunek pierwszy, zaś analizy należące do typu trzeciego nie

¹⁶ Por. P. F. Lazarsfeld i M. Rosenberg, *The Language of Social Research*, Glencoe 1955, s. 15 (Introduction). Por. także P. F. Lazarsfeld, *Evidence and Inference in Social Research*, [w:] *Evidence and Inference*, ed. D. Lerner, Glencoe 1959.

¹⁷ Por. analizy omawiane przez H. Hymana, *op. cit.*, s. 155-159. Hyman uważa je za lepsze od wielu innych, ponieważ bardzo różnorodne dane, które się ze sobą koreluje, mniej są narażone na zniekształcenia płynące z tego samego źródła.

wymagają, aby spełniony był warunek drugi. Zakładają one jednak co innego: porównywalność dwóch zbiorowości, od których pochodzą wyniki (typ II) lub wysoki stopień skorelowania cech, których dotyczą informacje uzyskiwane różnymi drogami (typ III).

2. Analizy weryfikacyjne z zastosowaniem pewnego sposobu otrzymywania danych

Analiz należących do wyróżnionych typów nie można ocenić jednolicie ze względu na możliwości weryfikacyjne, których one dostarczają. Trzeba jeszcze uwzględnić co najmniej jeden czynnik, mianowicie stopień pewności, jaki badacz wiąże z wynikami uzyskanymi różnymi drogami. Może to być — w przypadku drogi innej niż zastosowanie wywiadu kwestionariuszowego — stopień maksymalny, tj. badacz uważa te wyniki za praktycznie pewne. Niekiedy nie sądzi on, że są one pewne, jednak przekonany jest, iż są bardziej pewne niż wyniki uzyskane drogą wywiadu kwestionariuszowego. Często także badacz nie umie określić stopnia pewności obydwóch wyników. Dzieje się tak między innymi wtedy, gdy ową drugą drogą stanowi także wywiad kwestionariuszowy, a zwłaszcza gdy polega ona na powtórnym zadaniu tych samych pytań.

Ten trójczłonowy podział należy zastosować do wyróżnionych wyżej typów. Otrzymamy w rezultacie dziewięć typów analiz weryfikacyjnych, które łatwiej jest wartościować ze względu na możliwości uzyskania za ich pomocą dokładnej oceny wyniku badań. Najbardziej cenne z tego punktu widzenia, jak łatwo jest się domyślić, są analizy, w których uzyskuje się te same informacje odnoszące się do członków tej samej zbiorowości za pomocą innego, pewnego sposobu. Przykład takiej analizy stanowić może badanie weryfikacyjne, w którym zeznania o zażywaniu narkotyków korelowano drogą analizy chemicznej moczu¹⁸. Do

¹⁸ J. C. Ball, *The Reliability and Validity of Interview Data Obtained From 59 Narcotic Drug Addicts*, [w:] *Stages of Social Research Contemporary Perspectives*, ed. D. P. Forcésse i S. Richer. Artykuł J. C. Balla ukazał się najpierw w „American Journal of Sociology”, t. 72, 1967.

analiz tego typu można zaliczyć także wymienione wyżej badania nad piśmiennością¹⁹, weryfikację danych odnoszących się do wysokości zarobków, w których informacje inną drogą uzyskuje się z list płacy²⁰, badania, w których dane dotyczące zachowania, uzyskane w wywiadzie, kontroluje się obserwacyjnie itp.²¹

Wartość oceny wyników badań z zastosowaniem wywiadu kwestionariuszowego przy zastosowaniu analiz tego typu zależy od tego, czy dane zdobyte inną drogą są całkowicie pewne, ewentualnie czy wartość tę można w jakiś sposób oszacować. Pomijając tę kwestię, której rozważenie wychodzi poza zadania artykułu, należy stwierdzić, że ocena ta może być bardzo wielostronna. Przede wszystkim na podstawie analiz omawianego typu można stwierdzić, ile jednostek badania zostało błędnie zaklasyfikowanych, tj. ile jest jednostkowych błędów, i jeśli ich klasyfikacji dokonano przy użyciu odpowiedniej skali, jak duży jest każdy indywidualny błąd²². Oczywiście błędy te można policzyć. Po podsumowaniu, w którym nie bierze się pod uwagę kierunku odchylenia, otrzymujemy miarę tzw. błędu brutto (*gross error*, *total response error*), popełnionego w odniesieniu do danej zbiorowości. Otrzymana wielkość błędu brutto obejmuje dwa składniki tzw. błędu netto (*net response error*, *response bias*, *constant error*, *systematic error*) oraz różnicę między wielkością błędu brutto i netto (*response variance*, *variable error*), którą można

¹⁹ Były one przeprowadzane w Jugosławii w ramach kontroli wyników spisu powszechnego z 1953 r. Por. A. Sztandynger-Kaliszewicz, *Badania wiarygodności zeznań w spisie powszechnym*, [w:] *Analizy i próby...*, t. I, 1966, s. 266.

²⁰ Por. w niniejszym tomie artykuły: K. Lutyńskiej, *Weryfikacja danych uzyskanych z odpowiedzi respondentów na pytania o wysokość zarobków* i J. Koniarka, *Weryfikacja zewnętrzna informacji uzyskanych w wywiadzie kwestionariuszowym*.

²¹ Por. w niniejszym tomie artykuł M. Stefanowskiej, *Weryfikacja technik badania częstotliwości i intensywności oglądania telewizji w oparciu o dane z obserwacji uczestniczącej ukrytej*.

²² Zagadnieniem wielkości błędów jednostkowych nie będziemy się zajmowali w dalszym ciągu artykułu.

nazwać błędem wyrównanym²³. Pojęcia te — podobnie jak i pojęcie błędu brutto — odnoszą się zawsze do zbiorowości. Rozmiar błędu netto stanowi rezultat podsumowania błędów indywidualnych, w których uwzględnia się kierunek (znak) odchylenia. Wówczas jednostkowe błędy o różnym kierunku znoszą się. Suma tych ostatnich stanowi błąd wyrównany.

Dla jasności przytoczmy przykład. Błędne jednostkowe informacje o przynależności do danej organizacji uzyskano od 100 respondentów na 1000, spośród których 500 zgodnie z dobrze prowadzoną ewidencją należało do tej organizacji. Błąd brutto wynosi więc 100 przypadków, czyli 10%. Spośród nich w 70 przypadkach jednostkowy błąd polegał na tym, iż na podstawie relacji respondenta zaliczono go do członków organizacji, podczas gdy nim nie był, zaś w 30 przypadkach na tym, iż zaliczono go do nieczłonków, podczas gdy faktycznie do niej należał. Błąd netto wyniósł więc 40 przypadków (70-30), czyli 4%, zaś błąd wyrównany 60, czyli 6%.

Udział dwóch składników błędu brutto bywa różny. W krańcowych przypadkach może zdarzyć się tak, że błąd brutto jest równy błędowi netto bądź błędowi wyrównanemu. W pierwszym przypadku wszystkie odchylenia, tj. błędy indywidualne, są jednokierunkowe, a wielkość błędu wyrównanego wynosi 0. W drugim przypadku liczebności odchylenia w różnych kierunkach są sobie równe, w podsumowaniu z uwzględnieniem znaku znoszą się i błąd netto równy jest 0.

Odróżnienie dwóch składników błędu brutto jest bardzo potrzebne ze względu na to, iż tylko wielkość jednego z nich, a

²³ Wszystkie te pojęcia w odniesieniu do wyników otrzymanych przy użyciu wywiadu kwestionariuszowego stosują: M. H. Hansen, W. N. Hurwitz i W. G. Madow, *Sample Survey Methods and Theory*, t. I, s. 89 i nast., t. II, s. 280-289; H. Hyman i inni, *Interviewing in Social Research*, s. 20-22, 226-232, 321 i inne; C. A. Moser, G. Kalton, *op. cit.*, s. 378-381, 392, 398-399, i inni. Por. także K. Lutyńska i N. Szeżenia, *Wstępne wyniki zewnętrznej weryfikacji odpowiedzi respondentów w wywiadzie na temat zachowań zdrowotnych*, [w:] *Analizy i próby...*, t. IV, 1972, s. 344-347.

mianowicie, błędu netto, wpływa na liczebności, które charakteryzują rozkład danej cechy w populacji. Na te liczebności nie ma natomiast wpływu rozmiar błędu wyrównanego. Te właściwości znajdują właśnie wyraz w nazwach błędu netto i błędu wyrównanego. W podanym wyżej przykładzie rozmiar błędu netto wynosi 40. Odpowiadać mu będzie taka sama różnica między liczebnością członków organizacji według stanu faktycznego, ustalonego na podstawie wiarygodnych dokumentów, a ich liczebnością według danych z wywiadu kwestionariuszowego (500 i 540). Gdyby rozmiar błędu netto był równy 0 lub 200, wówczas ta różnica wyniesie także 0 lub 200. Taka sama, oczywiście, będzie rozbieżność w zakresie liczebności osób nie będących członkami organizacji. Natomiast przy różnych wartościach błędu wyrównanego różnice między tymi liczebnościami pozostaną bez zmian, jeśli nie ulegnie zmianie błąd netto. W przedstawionym powyżej przykładzie przy błędzie netto = 40 i liczebności członków organizacji według danych z dokumentów 500, zaś według danych z wywiadów — 500, błąd wyrównany może wynieść zarówno 0 (wówczas błąd brutto będzie równy błędowi netto), jak i 920 (wówczas błąd brutto będzie równy 960).

Przytoczone przykłady ilustrują przypadki krańcowe. Należy do nich zaliczyć jeszcze i taką sytuację, kiedy błąd wyrównany, a w konsekwencji i błąd brutto, obejmuje 100% przypadków, zaś błąd netto = 0. W podanym przykładzie sytuacja taka wystąpiłaby wówczas, kiedy wszyscy członkowie organizacji, stanowiący 50% zbiorowości, stwierdzą, że nimi nie są, zaś wszyscy nieczłonkowie podadzą, że do niej należą. Jest to oczywiście sytuacja paradoksalna. Paradoks polega na tym, iż na podstawie samych błędnych jednostkowych informacji dochodzi się do ogólnego wyniku zgodnego z prawdą.

Mogłoby się wydawać, że wielkość błędu wyrównanego nie posiada znaczenia dla badacza danego merytorycznego zagadnienia ze względu na to, iż nie wpływa on na liczebności, które charakteryzują rozkład danej cechy w zbiorze. Tak jednak nie jest. Wystąpienie błędnych zaklasyfikowań — wówczas gdy liczebności ich znoszą się — wpływa na obraz zależności danej cechy

z innymi, na to, jak będzie wyglądać ich siła, a nawet kierunek. Tak więc zmiana kierunku zależności wystąpi w ostatnio podanym przykładzie (błąd wyrównany = 100%), jeśli przynależność do danej organizacji będzie faktycznie silnie skorelowana z płcią. Gdyby w rzeczywistości przynależność ta była silnie pozytywnie związana z płcią męską, to związek ten przy tej wysokości błędu wyrównanego przekształci się w negatywny, pozostając tak samo silny. Oczywiście wynik taki będzie w tej sytuacji błędny²⁴.

Błędy, polegające na niewłaściwym zakwalifikowaniu jednostek badania do jednej z wyróżnionych klas badanych zjawisk, podczas gdy należą one do innej z wyróżnionych klas, są jedynymi błędami popełnionymi wtedy, gdy zrealizowany jest warunek, polegający na tym, iż wszystkie poszukiwane informacje można, przynajmniej teoretycznie, uzyskać w odniesieniu do wszystkich jednostek, gdy dotyczą one każdej z nich. Innymi słowami, gdy we wszystkich przypadkach spełnione są założenia jednostkowego pytania badacza, odpowiadającej poszukiwanej informacji, na przykład założenie głoszące, że respondent podziela jedną z wskazanych w pytaniu w kwestionariuszu opinii, czego dotyczy taka poszukiwana informacja, lub że rzeczywiście pracuje zarobkowo, jeśli badacz chce się dowiedzieć, ile on zarabia, i zleca zadanie mu odpowiedniego pytania.

W rzeczywistych badaniach niemal zawsze zdarza się tak, że w odniesieniu do pewnej, choćby bardzo małej ilości respondentów, którym zadano odpowiednie pytania, założenia te nie są spełnione, często wbrew oczekiwaniom badacza. Pytania te nie były więc dostosowane do rzeczywistości. W tych przypadkach idealny respondent udzielił odpowiedzi nieistotnej, wskazując, że fakt, o który go pytają, nie występuje (odpowiedzi tzw. korek-

²⁴ Wielkość błędu netto w niektórych przypadkach nie wpływa na wysokość korelacji, jak stwierdza H. Hym an (*Survey Design...*, s. 150, 160, 167). Są to jednak przypadki specjalne. W najczęściej występujących w socjologii typach analiz błąd tego rodzaju przyczynia się do obniżenia wartości uzyskanych współczynników. Pojęcia: „błąd brutto”, „netto” i „błąd wyrównany” można zastosować także wtedy, gdy w klasyfikacjach uwzględnia się wiele kategorii, tj. cecha występuje w wielu odmianach. Dla uproszczenia w przykładach ograniczono się tylko do dychotomii.

cyjne). Jednak może zdarzyć się i tak, że wskutek różnych przyczyn (złe rozumienie pytania, tendencja do udzielania odpowiedzi „byłe jakich”) respondent udzieli odpowiedzi istotnej, tj. takiej, która przynosi poszukiwaną informację, na podstawie której dana jednostka zostanie zaliczona do jednej z wyróżnionych klas badanych zjawisk. Odpowiedź ta nie jest jednak trafna, gdyż pociąga za sobą błąd, polegający na zaklasyfikowaniu do niewłaściwej kategorii kogoś, kto do niej nie należy. Jest to jednak błąd inny niż poprzednio wyróżniony, ponieważ w tym przypadku zaklasyfikowanie objęło przypadek, który w ogóle nie mieści się w danym podziale. Cechą wspólną obydwu rodzajów błędów jest jednak to, że w jednym i drugim przypadku otrzymana informacja nie odpowiada rzeczywistości.

W dalszych rozważaniach nie będziemy zajmować się wyróżnionymi ostatnio błędami i przypadkami, kiedy badacza spotkało niepowodzenie, tj. nie uzyskał on poszukiwanej informacji, choć dotyczyła ona danych jednostek badania. Należy jednak zauważyć, że w wyniku weryfikacji prowadzonej przy zastosowaniu sposobu pewnego — a właściwie doskonałego — można wskazać ilość rzeczywistych niepowodzeń oraz przypadków, kiedy odpowiedź jest prawidłowa, tj. wskazuje ona na konieczne ograniczenie, ponieważ poszukiwana informacja nie dotyczy danych jednostek badania²⁵.

Odróżnienie błędu brutto i jego składników na poziomie całej zbiorowości jest użyteczne przy omawianiu następnego typu analiz weryfikacyjnych, polegających na zastosowaniu sposobu pewnego do dwóch porównywalnych zbiorowości. Przykład takiej

²⁵ Pojęcie odpowiedzi istotnej, tj. odpowiedzi respondenta, która stanowi podstawę dla rozstrzygnięcia pytania badacza, przynosi poszukiwaną informację, jak również odpowiedzi trafnej, tj. odpowiedzi respondenta, która przynosi prawdziwą informację, oraz związane z nimi pojęcia błędu i niepowodzenia omówiono w cytowanym artykule *Analiza procesu otrzymywania informacji...* Pewnej zmianie w porównaniu z tamtym artykułem uległa jednak terminologia. Mianowicie termin „odpowiedź istotna” zastępuje tutaj poprzednią nazwę „odpowiedź pożądana”, zaś terminu „odpowiedź trafna” użyto obecnie zamiast nazwy „odpowiedź adekwatna”.

analizy stanowić mogą badania dotyczące zgodności postępowania ludzi z przepisem zakazującym przechodzenia przez jezdnię przy czerwonym świetle. Informacje uzyskano dwiema drogami: przy zastosowaniu pytań wprost w wywiadzie kwestionariuszowym oraz drogą obserwacji, którą traktowano jako drogę pewną. Badane zbiorowości uznano za porównywalne w wysokim stopniu. Ponieważ w obydwóch przypadkach otrzymano zbliżone wyniki, wynik dotyczący liczebności osób, które nie przestrzegają danego zakazu, uzyskany przy zastosowaniu wywiadu kwestionariuszowego, uznano w zasadzie za pozytywnie zweryfikowany²⁶.

Wartość tej oceny, w danym przypadku nie budząca zresztą zastrzeżeń, jest zależna od dwóch okoliczności. Po pierwsze, podobnie jak w przypadku poprzednich analiz zależy ona od stopnia pewności wyniku otrzymanego przy zastosowaniu innego sposobu. Po drugie, zależy od stopnia porównywalności obydwu grup. Przy zwiększonych wymaganiach stopień tej pewności i porównywalności należałoby oszacować. Niekiedy takie oszacowanie jest możliwe w odniesieniu do porównywalności grup, mianowicie wtedy, gdy zbiorowość kontrolna stanowi próbkę statystyczną zasadniczej próby. Wówczas gdy zbiorowość jest jedna i ta sama i nie można było tylko zestawić ze sobą rezultatów na poziomie jednostek, porównywalność wyników jest oczywiście zapewniona. Najczęściej jednak badacz musi uzasadnić inną drogą — czasem *ex post* — że zbiorowości, które badał są porównywalne, przy czym stosować tu może różne zabiegi.

Ale omawiany obecnie typ analiz różni się od poprzedniego także czymś innym, mianowicie zakresem oceny. Ponieważ w tych analizach nie zestawia się ze sobą zaklasyfikowań poszczególnych jednostek, lecz tylko wyniki globalne, odnoszące się do całej zbiorowości, ocena ta nie dotyczy błędu brutto, lecz błędu netto. Gdyby w podanym wyżej przykładzie liczba przechodzących przez jezdnię przy czerwonym świetle w wyniku pocho-

²⁶ A. Podgórecki i inni, *Poczucie prawne społeczeństwa polskiego*, Warszawa 1968, Ośrodek Badania Opinii Publicznej i Studiów Programowych przy Polskim Radio i Telewizji, nr 83, s. 50 - 51.

dzącym z obserwacji była większa — dajmy na to — o 10%, od liczby uzyskanej w wyniku badań z użyciem wywiadu kwestionariuszowego, wówczas jeśli traktujemy ten pierwszy wynik jako pewny, możemy przyjąć, iż w drugim błąd netto wynosi 10%. (Posługiwanie się danymi względnymi jest tu potrzebne, ponieważ liczba osób badanych przy zastosowaniu jednego i drugiego sposobu może być różna.) Tyle samo wyniosłby błąd netto, gdybyśmy jedną z tych zbiorowości poddali dwukrotnemu badaniu przy zastosowaniu tych samych sposobów. W tym ostatnim badaniu otrzymalibyśmy jednak także informacje o błędzie wyrównanym. Może bowiem zdarzyć się, że respondent podał błędnie w wywiadzie, że przechodzi przy czerwonym świetle przez jezdnię, chociaż tego faktycznie nie robi. Przypadki takie byłyby zapewne bardzo nieliczne w omawianym wyżej badaniu, natomiast w innych badaniach przy danym błędzie netto błąd wyrównany może być duży. W rezultacie zastosowania analiz należących do obecnie omawianego typu otrzymujemy więc informacje tylko o błędzie netto. W konsekwencji nawet najbardziej pozytywna ocena wyniku, otrzymanego przy zastosowaniu wywiadu kwestionariuszowego, uzyskana w rezultacie analizy tego typu, nie gwarantuje poprawności wyników odnoszących się do zależności między cechą, której dotyczy ten wynik, a innymi cechami. Ocena wyniku nie uwzględnia bowiem ewentualnego błędu wyrównanego, którego wpływ na korelację może być duży w zależności od jego rozmiarów. Okoliczność tę należy brać pod uwagę przy wykorzystaniu analiz omawianego typu.

3. Analizy wykorzystujące korelacje danych

Stwierdziliśmy, że między analizami należącymi do omawianych dotychczas typów występują znaczne różnice, zarówno ze względu na to, co za ich pomocą można ocenić (występowanie i zakres błędu danego rodzaju), jak i ze względu na założenia, których wymaga ich przeprowadzenie. Niezależnie od różnic powyższe analizy pozwalają na sformułowanie jakiejś oceny całego wyniku, znajdującej wyraz w liczbach. Z tego punktu wi-

dzenia należy je przeciwstawić analizom należącym do pozostałych typów. W ich rezultacie nie można bowiem sformułować żadnej liczbowej oceny, bądź też ocena tego rodzaju odnosi się do takiego aspektu wyniku, którego stosunek do większości błędów nie da się określić, przynajmniej w obecnym stanie metodologicznej refleksji.

W pierwszej sytuacji znajdują się wszystkie analizy, których wynik odnoszący się do występowania jakiejś cechy w dużej zbiorowości ocenia się drogą sprawdzenia, czy jest ona skorelowana z jakąś inną cechą. Występowania takiej korelacji można oczekiwać zgodnie z posiadaną już wiedzą. Ponieważ nie jest to zwykle wiedza zbyt dokładna, na podstawie analiz tego typu nie można stwierdzić, że w badaniu popełniono tyle a tyle jednostkowych błędów, czy na przykład błąd netto o danej wielkości. Można jedynie powiedzieć, że ponieważ wystąpiła oczekiwana korelacja, to prawdopodobnie znaczna część zaklasyfikowań jednostek badania, na których podstawie przyjęto weryfikowany wynik, jest właściwa. Prawdopodobnie — ponieważ rozumowanie, którego rezultatem jest ten wniosek, jest zawodne i może się zdarzyć tak, choć jest to mało prawdopodobne, że omawiana korelacja wystąpi i przy dużej ilości błędnych zaklasyfikowań. Gdyby się natomiast okazało, że korelacja taka nie wystąpiła, wówczas znaczna część zaklasyfikowań jednostek ze względu na jedną lub obydwie korelowane cechy jest błędna.

Ostatni wniosek posiada znaczenie dla badacza, ale tylko wtedy, gdy interesuje go w równej mierze ocena obydwóch wyników²⁷. Jeśli jednak chce on ocenić tylko jeden z nich, co zdarza się daleko częściej, musi umieć ocenić stopień pewności zaklasyfikowań dokonanych inną drogą, tj. przy zastosowaniu innego sposobu badania niż wywiad kwestionariuszowy. Jeśli uważa je za bardziej pewne, to wówczas, gdy oczekiwana korelacja nie wystąpi, może uznać za bardziej prawdopodobne, iż więcej błędów popełniono w klasyfikacji jednostek badania, przeprowadzo-

²⁷ Analizę, którą można zaliczyć do tego typu, przedstawia H. H y m a n w *Survey Design* ..., s. 156.

nej przy uwzględnieniu rezultatów wywiadu kwestionariuszowego, niż w pozostałej. Gdyby natomiast uważał je za pewne, wówczas co najmniej znaczną część informacji uzyskanych za pomocą tego wywiadu powinien uznać za błędne. W obydwóch ostatnich przypadkach badacz może także formułować przypuszczenia co do kierunku popełnionych błędów²⁸. Należy jeszcze zaznaczyć, że wartość wszystkich analiz omawianego typu zależy od tego, jak silna jest korelacja, której występowanie stanowi podstawy analizy. Musi być ona w ogóle bardzo silna, aby te analizy miały sens. Gdyby była równa 1, to wówczas można by formułować w nich te same wnioski, co w rezultacie analiz pierwszego z zasadniczych typów.

4. Analizy oparte na porównywaniu wyników niepewnych

Ostatnie wnioski dotyczyły różnych możliwości oceny wyniku badań, jakie stwarzają analizy bardziej szczegółowych typów, należących do trzeciego spośród wymienionych zasadniczych typów. Przejdźmy teraz do omówienia analiz weryfikacyjnych, należących do pozostałych czterech szczegółowych typów. We wszystkich tych analizach porównuje się wyniki uzyskane dwiema drogami, podobnie jak w dwóch omówionych najpierw typach. Typy tych analiz są następujące:

I. Badacz opiera się na informacjach, które zdobył dwiema drogami i które uważa za równie niepewne. Informacje te odnoszą się do wszystkich jednostek badania w zbiorowości. Przykład takich analiz stanowią studia, w których wykorzystuje się odpowiedzi dotyczące tych samych faktów na pytania zadane różnym osobom (np. małżonkom)²⁹ lub kategoriom osób, następnie odpowiedzi uzyskane na pytania kontrolujące się z tego same-

²⁸ Przykłady takich analiz można znaleźć u H. Hymana, *op. cit.*, s. 155 - 159.

²⁹ Klasyczny przykład stanowią tu analizy Kinseya, który różne pytania dotyczące życia seksualnego zadawał osobno mężowi i żonie (A. C. Kinsey, W. B. Pomeroy, C. E. Martin, *Sexual Behavior of the Human Male*, Saunders, 1948). Podobny sposób zastosowała E. Nowicka

go kwestionariusza, tj. na różne pytania dostarczające tych samych informacji³⁰, a także wszelkie powtórne badania przeprowadzone za pomocą identycznych lub różnych narzędzi w stosunku do tych samych osób.

II. Badacz wykorzystuje informacje uzyskane różnymi drogami, które uważa za równie niepewne. Informacje te odnoszą się do różnych, porównywalnych zbiorowości, przy czym każda zbiorowość badana jest tylko jedną drogą. W praktyce każda z różnych dróg stanowi najczęściej rodzaj wywiadu kwestionariuszowego. Zazwyczaj także obydwie badane grupy są częściami zbiorowości generalnej, której dotyczą wnioski we właściwym, merytorycznym badaniu. Przykład stanowią liczne analizy efektu ankierskiego³¹ lub analizy, w których badacz stara się usta-

w pracy pt. *Niektóre uwarunkowania procesu uzyskiwania informacji w wywiadzie drażliwym. Postawy wobec zagadnień seksualnych a wywiad o wychowaniu seksualnym* (nie publikowana praca magisterska z 1973 r., przygotowana w Zakładzie Metod i Technik Badań Społecznych Instytutu Socjologii Uniwersytetu Łódzkiego). Wywiady przeprowadzone z matkami i córkami dotyczyły głównie kwestii, jak córki były uświadamiane przez matki w sprawach seksualnych. Badania wykazały, że zgodne informacje otrzymywano najczęściej w przypadkach, gdy matka i córka odznaczały się tolerancją. Mniejsze znaczenie miała tutaj zgodność postaw wtedy, gdy matkę i córkę cechował rygorizm.

³⁰ Por. mój artykuł *Projekt weryfikacji...* oraz artykuły: T. Pawłowskiego, *Logiczne podstawy weryfikacji wewnętrznej*, i P. Boskiego, *Związki logiczne między odpowiedziami na pytania kwestionariusza*, [w:] *Analizy i próby...*, z. IV, 1972. Wykonane na podstawie założeń sformułowanych w tych artykułach prace J. Kijewskiej i M. Repiejki wykazały, że ilość niezgodności występujących w odpowiedziach na 49 pytań kwestionariusza „Międzynarodowych badań porównawczych nad wykorzystaniem opieki zdrowotnej” jest bardzo mała (niezgodnych jest ok. 2% odpowiedzi) i że niezgodności te mają przypadkowy charakter. Por. prace magisterskie tych osób w Bibliotece Instytutu Socjologii UŁ.

³¹ Badania te obszernie omawia H. Hyman w cytowanej pracy *Interviewing in Social Research*. Por. także D. Riesman, *Some Observations...*, s. 343 - 363. Riesman badał wpływ w stylu przeprowadzenia wywiadu na wyniki. W literaturze polskiej por. m. in. artykuł M. Jastrzęb, *Wpływ płci ankietera na wyniki uzyskiwane w badaniach*, [w:] *Analizy i próby...*, t. I, 1966.

lić wpływ na wyniki poszczególnych elementów sytuacji wywiadu³² bądź sformułowania pytań³³. Założenie porównywalności badanych grup jest w tych analizach przyjmowane hipotetycznie lub też jego słuszność jest uzasadniona w różny sposób³⁴.

III. Badacz opiera się na informacjach odnoszących się do tych samych jednostek i uzyskanych dwiema różnymi drogami, przy czym jedna z tych dróg i jej rezultaty na początku lub w wyniku interpretacji danych uważa się za bardziej pewne i stąd do informacji zdobytych przy jej wykorzystaniu ma większe zaufanie. Przykład stanowić może tutaj analiza zgodności i niezgodności danych o poglądach tych samych osób, uzyskanych za pomocą standardowego wywiadu kwestionariuszowego i wywia-

³² Por. np. K. Lutyńska, *Osoby trzecie w wywiadzie socjologicznym i ich wpływ na odpowiedzi respondentów*, [w:] *Analizy i próby...*, t. II, 1968, tejże, *Miejsce przeprowadzenia wywiadu i jego oddziaływanie*, [w:] *Analizy i próby...*, t. IV, 1972.

³³ Por. H. Cantril, *Gauging Public Opinion*, Princeton 1944, cz. I. Praca Cantrila, napisana wspólnie z jego współpracownikami, stanowi klasyczną już dzisiaj pozycję z tego zakresu. Dość liczne są analizy poświęcone różnicom w odpowiedziach na pytania otwarte i zamknięte, które miały przynieść te same informacje. Por. u nas M. Stefanowska, *Próba porównania odpowiedzi na pytania prekategoryzowane i otwarte w badaniach nad konsumpcją kulturalną*, [w:] *Analizy i próby...*, t. IV, 1972, oraz tejże, *Próba analizy porównawczej dwóch sposobów pytań o konsumpcję kulturalną*, [w:] *Z metodologii i metodyki socjologicznych badań terenowych*, pod red. Z. Gostkowskiego, Warszawa 1972.

³⁴ Najczęstsze sposoby zapewnienia porównywalności są następujące: 1. Jedną zbiorowość stanowi próbkę statystyczną, najczęściej losową, drugiej. 2. Stosuje się losowy przydział jednostek do dwóch grup tworzących razem większą zbiorowość. Bardzo często wykorzystuje się tutaj tzw. *split ballot technique* oraz próbki przenikające się (*interpenetrating samples*). Por. C. A. Moser, G. Kalton, *op. cit.*, s. 124. 3. Przydział taki może być także celowy np. drogą formowania grup posiadających zbliżone cechy. 4. *Ex post* wykazuje się, że badane zbiorowości były ze względu na ważne cechy bardzo zbliżone. 5. *Ex post* wykazuje się, że ustalone zależności między przynależnością do grup a różnicami w wynikach nie były pozorne. W tym przypadku wnioski dotyczą porównywalności jedynie w danym zakresie.

du swobodnego, przeprowadzonego przez osobę znajomą. Wywiad swobodny badacz uważał za sposób lepszy³⁵.

IV. Badacz wykorzystuje rezultaty odnoszące się do dwóch porównywalnych zbiorowości, z których jedną bada drogą zastosowania wywiadu kwestionariuszowego, a drugą za pomocą sposobu bardziej pewnego. Tę, która z tych dróg jest bardziej pewna, ustala na początku analizy lub po interpretacji danych. Przykładem studium tego typu jest analiza, w której porównuje się wyniki uzyskane za pomocą wywiadu kwestionariuszowego z wynikami pogłębionego wywiadu weryfikacyjnego (*follow-up interview, quality check*), zastosowanych do próbki uprzednio badanych osób³⁶, następnie analiza porównawcza wyników standardowego wywiadu kwestionariuszowego z wynikami anonimowej ankiety oraz wywiadu ze znajomymi, zastosowanych do porównywalnych zbiorowości³⁷. Scharakteryzowane obecnie

³⁵ Por. K. M. Siomczyński, *Wpływ oficjalnej i prywatnej sytuacji wywiadu na wypowiedzi respondentów w środowisku inteligencji*, [w:] *Analizy i próby...*, t. I, 1966. Por. także A. Kubiak, *Analiza weryfikacyjna pytań globalnych w wywiadzie kwestionariuszowym* (nie opublikowana praca magisterska, wykonana w Zakładzie Metod i Technik Badań Społecznych IS UŁ).

³⁶ Por. H. Hyman, *op. cit.*, s. 167-168, C. A. Moser, G. Kalton, *op. cit.*, s. 395-398. Jeśli analiza tego rodzaju przeprowadzona jest z próbką zbiorowości generalnej, to wówczas w zastosowaniu do tej próbki oraz przy identyfikacji respondentów należy ona do typu poprzedniego.

³⁷ Por. dwa artykuły w *Analizach i próbach...*, t. III, 1970: L. Sołoma, *Analiza porównawcza wartości danych, zdobytych za pomocą różnych technik*, oraz G. Carceles-Breis, *Dwie techniki pomiaru poziomu piśmienności w krajach słabo rozwiniętych*. Por. także analizy tego typu, omawiane przez H. Hymana (*op. cit.*, s. 162-163, 167), a także cytowaną pracę H. Cantrila, *Gauging Public Opinion*, rozdz. V: *Gauging „Secret Versus Nonsecret Ballots”*, s. 77-83.

Jak zaznaczyliśmy, ocena, iż druga droga jest bardziej pewna, w analizach III i IV typu może być sformułowana po interpretacji otrzymanych danych. Jest to najczęściej interpretacja psychologiczna. Za bardziej pewne przyjmuje się na przykład wyniki otrzymane przy zastosowaniu anonimowej ankiety od różnych od nich wyników wywiadu kwestionariuszowego, jeśli w ankiecie na te same, co w wywiadzie, drażliwe pytania otrzy-

szczegółowe typy I i III należą do pierwszego zasadniczego typu, zaś II i IV — do zasadniczego typu drugiego.

W efekcie analiz, należących do scharakteryzowanych typów, uzyskuje się wyniki liczbowe dotyczące różnic między rezultatami uzyskiwanymi różnymi drogami. Są one pod wieloma względami analogiczne do tych wyników, które otrzymuje się przy weryfikacji z użyciem sposobów pewnych w odniesieniu do tej samej lub porównywalnej zbiorowości. Stąd też omawiane różnice liczbowe można nazwać podobnie — różnicami brutto, netto i wyrównanymi. Rezultaty analiz należących do I i III typu charakteryzują różnice brutto, netto i różnice wyrównane. W rezultacie analiz typu II i IV można oszacować różnicę netto. Jeśli na przykład różnica wyników osiągniętych drogą wywiadu kwestionariuszowego i ankiety, zastosowanych do różnych, porównywalnych zbiorowości, wynosi 20% (np. drogą wywiadu ustalono, że do kina stale chodzi 40% osób, zaś drogą ankiety, że takich osób jest 60%), to tyle samo wyniosłaby różnica netto, gdybyśmy obydwie sposoby zastosowali do wszystkich jednostek w jednej zbiorowości. Wynik dotyczący zbiorowości porównywalnej zastępuje tutaj wynik otrzymany przy użyciu drugiej drogi w badaniu tej samej zbiorowości. Różnica netto w liczbach względnych jest więc stała dla danych sposobów, niezależnie od ilości osób, które zbadano przy zastosowaniu każdego z nich.

Jak należy interpretować te różnice? Ponieważ pojawiły się one wskutek zastosowania odmiennych dróg w badaniu jednej lub porównywalnych zbiorowości, można uznać je za efekt różnic występujących między drogami. Źródłem ich są więc odmienne wartości cech (zmiennych), które charakteryzują te drogi. Jeśli więc cechą różniącą dwa rodzaje wywiadu, które zastosowano do tej samej lub porównywalnych zbiorowości było to, iż

mano więcej odpowiedzi wymagających większego stopnia szczerości. Przykłady takich interpretacji przedstawiają cytowane wyżej prace. Por. także H. Hyman, *Interviewing in Social Research*, s. 141-142. Trzeba tu zaznaczyć, że w tych przypadkach autorzy wykorzystują w analizie sposób charakterystyczny dla weryfikacji wewnętrznej, ponieważ opierają się na wnioskach dotyczących procesów formowania informacji.

w jednym przypadku wywiad przeprowadzała osoba dobrze znana respondentowi, zaś w drugim — osoba nieznajoma, to ewentualną różnicę w wynikach należy wiązać z faktem uprzedniej znajomości między respondentem a ankierem. Oczywiście może zdarzyć się tak, że te same różnice, a więc i związane z nimi błędy, o czym niżej, można powiązać nie tylko z tym, lecz i z innym źródłem.

Występowanie różnic zawsze świadczy o błędach. Jeśli bowiem jakaś jednostka badania została zaklasyfikowana niejednokrotnie przy użyciu dwóch dróg zdobywania informacji, wówczas obydwie zaklasyfikowania nie mogą być zgodne z faktycznym stanem rzeczy, i co najmniej jedno z nich musi być błędne. Natomiast zgodność zaklasyfikowań nie świadczy o braku błędu w ogóle. Świadczy jednak o tym, że w danym przypadku nie występuje błąd związany z różnicą (lub różnicami) między sposobami otrzymywania danych.

Twierdzenia te pozwalają na sformułowanie następujących zasad interpretacji różnic między wynikami, otrzymywanymi przy zastosowaniu różnych dróg zdobywania informacji. Jeśli tych różnic nie ma, należy przyjąć, że obydwie wyniki nie są obciążone błędem, którego źródłem są różnice między tymi drogami. Mogą być jedynie obciążone błędami o innej genezie, innego rodzaju. Zasada ta odnosi się do wszystkich czterech typów omówionych ostatnio analiz, z tym zastrzeżeniem, że wnioski formułowane w rezultacie analiz typu III i IV dotyczą tylko braku błędów związanych z różnicami w zaklasyfikowaniach, które złożyły się na różnicę netto. Nie uwzględniają więc one wszystkich ewentualnych błędów o danej genezie³⁸.

³⁸ Wydawać się może, że w przypadku analiz należących do typu III i IV nie powinno się otrzymywać takich samych rezultatów przy zastosowaniu dwóch dróg otrzymywania danych, ponieważ jedna z tych dróg, właśnie ze względu na te rezultaty, została uznana za lepszą. Otrzymanie takich samych rezultatów nie świadczy jednak, że to przekonanie jest niesłuszne. Różnica w ocenie może być uzasadniona w stosunku do innych zastosowań omawianych dróg, nie wywarła ona jedynie wpływu w danej sytuacji. Warto także zaznaczyć, że zróżnicowana ocena dróg otrzymywania danych nie może pojawić się dopiero po interpretacji rezultatów ana-

Gdy między wynikami otrzymanymi różnymi drogami występują różnice, to zawsze jeden z tych wyników obciążony jest błędem, którego źródłem są różnice między drogami. Niezależnie od tego obydwa wyniki mogą być obciążone błędami o innej genezie, innego rodzaju. W przypadku analiz typu II i IV są to ponadto tylko błędy związane z różnicą netto, a więc nie wszystkie błędy o danej genezie³⁹.

Jaka ilość błędów jednostkowych — czyli jaki rozmiar błędu na poziomie całej zbiorowości — związana jest z różnicami w rezultatach otrzymanych różnymi drogami? W analizach typu I i III wielkość błędu na poziomie zbiorowości ustalić jest bardzo łatwo. Każda bowiem jednostka klasyfikowana jest w nich dwukrotnie i jeśli została zaklasyfikowana niejednakowo, wówczas co najmniej jedno z tych zaklasyfikowań jest błędne. W konsekwencji rozmiar tego błędu równy jest wielkości różnicy między rezultatami. Obliczenie wielkości błędu w tych analizach przynosi jednak bardzo mało. Po pierwsze, trzeba wziąć pod uwagę, że jest to tylko błąd, którego źródła tkwią w różnicach między sposobami otrzymywania danych. Po drugie, nie wiadomo, do czego należy odnieść te omyłki. Błędy popełniono bowiem w jednej lub drugiej klasyfikacji jednostek badania, związanych z różnymi drogami otrzymywania informacji. Bez dodatkowych więc informacji, o czym będzie jeszcze mowa niżej, nie można ich — a więc i odpowiadającego im błędu popełnionego na poziomie zbiorowości — odnieść tylko do jednego z porównywanych wyników i uznać, że stanowi on rezultat zastosowania jednej z dróg badania.

liz weryfikacyjnych tych typów w przypadku, gdy wyniki uzyskane przy zastosowaniu tych dróg są jednakowe. Musi być ona wtedy sformułowana wcześniej.

³⁹ Nie oznacza to, że błąd wiążący się z różnicą netto wchodzi w skład ogólnego błędu netto, którego wielkość można określić porównując oceniany wynik z wynikiem pewnym. Jest tak dlatego, że ewentualne błędy z różnych źródeł mogą znosić się i błąd wiążący się z różnicą netto danego rodzaju wejdzie w skład błędu wyrównanego. Natomiast błąd związany z różnicami wyrównanymi o danej genezie zawsze wejdzie w skład ogólnego błędu wyrównanego.

Pewne wnioski co do wielkości błędu netto omawianego rodzaju można sformułować wówczas, kiedy porównujemy wyniki odnoszące się do dwóch zbiorowości badanych przy wykorzystaniu różnych dróg, tj. w analizach typu II i IV. Wielkości tego błędu w większości przypadków nie da się wówczas ustalić dokładnie. Można jedynie wskazać granice, w których on się mieści. Wyznaczenie tych granic wymaga obliczeń opartych na następującym rozumowaniu.

Gdy wszystkie jednostki, wchodzące w skład obydwóch zbiorowości, były klasyfikowane dwa razy — każda przy zastosowaniu dwóch dróg, jak to się dzieje w analizach typu I i III — wówczas rozmiar błędu omawianego rodzaju związanego z różnicą netto byłby równy wielkości tej różnicy. Jednak w omawianych obecnie analizach każda jednostka klasyfikowana jest tylko raz. Z tego powodu wielkość błędu, który popełnia się przy tej klasyfikacji, musi być mniejsza od wielkości różnicy netto, obliczonej w stosunku do całej połączonej zbiorowości. Byłaby ona równa połowie tej wielkości, w przypadku gdyby wszystkie błędy popełniono tylko jedną drogą, przy zastosowaniu której zbadano połowę wszystkich jednostek. Natomiast gdyby tą drogą zbadano nie połowę, lecz $\frac{3}{4}$ wszystkich jednostek, wówczas w całości popełniono by ilość błędów, równą $\frac{3}{4}$ wielkości różnicy netto, obliczonej dla całej zbiorowości.

Może jednak zdarzyć się tak, że wszystkie błędy popełniono nie przy zastosowaniu tej drogi, lecz drogi drugiej, za pomocą której zbadano $\frac{1}{4}$ wszystkich jednostek. Wówczas ilość błędów wyniosłaby $\frac{1}{4}$ wielkości różnicy netto, obliczonej dla całej zbiorowości. Przy tych liczebnościach osób, zbadanych przy zastosowaniu każdej drogi, obydwie wielkości stanowią górną i dolną granicę rozmiaru błędu danego rodzaju w całej zbiorowości, ponieważ jeśli zwiększy się proporcja błędów netto popełnionych przy zastosowaniu jednej drogi, to odpowiednio musi zmniejszyć się ich proporcja wśród rezultatów otrzymanych drugą, ze względu na stałą wartość różnicy netto przy zastosowaniu tych dróg. Wielkość błędów badanego rodzaju maksymalnego i minimalnego zbliżają się do siebie tym bardziej, im bardziej zbliżają się do

siebie ze względu na swoją liczebność obydwie zbiorowości składowe. Gdy są one równoliczne, wówczas różnica między błędem maksymalnym i minimalnym zanika i rozmiar błędu jest zawsze równy połowie wielkości różnicy netto, obliczonej dla całej zbiorowości⁴⁰.

Czy i ewentualnie jaką wartość dla oceny wyniku otrzymane-

⁴⁰ Oto przykład liczbowy odnoszący się do porównywalnych zbiorowości. Jeśli jedną drogą, dajmy na to za pomocą wywiadu kwestionariuszowego, przeprowadzonego w domu, otrzymano rezultat, iż 180 respondentów na 300, czyli 0,6 zbadanych, stale czyta książki, zaś inną drogą, kiedy wywiad przeprowadzono w miejscu pracy, uzyskano wynik, iż takich osób było 40 na 100, czyli 0,4 zbadanych, to różnica netto między tymi rezultatami wynosi $0,6 - 0,4 = 0,2$. Błąd danego rodzaju dla całej zbiorowości (400) będzie maksymalny wtedy, gdy błędne informacje pochodzą tylko z wywiadów przeprowadzonych w domu, ponieważ tam zbadano 0,75 osób z całej zbiorowości. Jego wielkość jest następująca: $0,2 \times 0,75 \times 1 = 0,15$, czyli 60 przypadków (1 oznacza, że wszystkie błędy popełniono przy zastosowaniu tej drogi). Natomiast wielkość błędu danego rodzaju będzie minimalna, jeśli błędne informacje uzyskiwano tylko w miejscu pracy. Jego wielkość jest następująca: $0,2 \times 0,25 \times 1 = 0,05$, czyli 20 przypadków. Błąd związany z różnicą miejsca przeprowadzenia wywiadu mieści się więc w granicach 0,05 - 0,15, czyli 20 - 60 przypadków. Gdyby połowę (200) respondentów zbadano w domu, połowę zaś w miejscu pracy, wówczas różnica między błędem minimalnym i maksymalnym ginie, ponieważ zarówno dla jednej, jak i dla drugiej drogi otrzymujemy wynik: $0,2 \times 0,5 \times 1 = 0,10 = 20$ przypadków.

Analogiczny sposób obliczania wielkości błędu (*bias*) związanego z różnicą wyników otrzymanych od różnych ankietowanych występuje *implicite* u F. Mostellera. W związku z obliczeniem wielkości zniekształceń w całej zbiorowości Mosteller radził, aby liczebności ankietowanych wykazujących sympatie, np. konserwatywne i liberalne, które mogą wpłynąć na rezultaty wywiadów, były równe. W tym bowiem przypadku, przy różnych rzeczywistych a nie znanych badaczowi z góry rozkładach poglądów liberalnych i konserwatywnych wśród respondentów liczebność błędów będzie przeciętnie najmniejsza. Oczywiście wszyscy ankietujący badają jednakowe ilości osób. Por. F. Mosteller, *Correcting for Interviewer Bias*, [w:] H. Cantril, *Gauging Public Opinion*, s. 315 - 317. Por. także H. Hyman, *Interviewing in Social Research*, s. 315 - 317, którego sformułowania wyraźnie dotyczą sprawy wielkości błędu, nie uwypuklonej przez Mostellera.

go we właściwym badaniu posiada określenie lub oszacowanie błędu omawianego rodzaju? Zależy to od typu analizy, a w przypadku analiz typu II, także od tego, co badacz traktuje jako ów wynik.

W analizach typu III i IV, kiedy badacz wie, która droga badania jest lepsza, z góry, co w tym przypadku potwierdza fakt zróżnicowania wyników, bądź nawet tylko *ex post* w rezultacie interpretacji różnic — może on z powodzeniem wykorzystać miarę tego błędu. W rezultacie analiz typu III stwierdza on wtedy, że błąd (danego rodzaju) o takiej wielkości obciąża w większym stopniu wyniki otrzymane drogą wywiadu niż inną drogą. Czyli że błąd obciążający wyniki wywiadu zapewne zbliża się do takiej granicy. W rezultacie analizy typu IV wniosek badacza musi być jeszcze bardziej ostrożny. Sprowadza się on do przekonania, że błąd danego rodzaju zbliżony jest bardziej do danej wielkości niż do określonej wielkości innej i obciąża w większej części wynik otrzymany przy zastosowaniu tej, a nie innej drogi. Przykładowo, jeśli różnica netto między wynikami otrzymanymi za pomocą wywiadu i wynikiem otrzymanym drogą anonimowej ankiety wynosi 0,20 (20%), zaś za pomocą ankiety, którą uznano za sposób lepszy, zbadano 0,4 całej zbiorowości, wówczas można sądzić, że rozmiar błędu (danego rodzaju) zbliża się do wielkości $0,12 (0,20 \times 0,6 \times 1)$ niż do wielkości $0,08 (0,20 \times 0,4 \times 1)$ i że obciąża on — co jest zrozumiałe — w większej części wynik otrzymany drogą wywiadu. W obydwu analizach badacz może także wskazać kierunek błędu danego rodzaju jako jedyny (analizy typu IV) lub dominujący (analizy typu III, w których uwzględnia się także różnice wyrównane). Stwierdzenia powyższe nie są precyzyjne, mogą być jednak użyteczne dla badacza. Byłyby one, oczywiście, daleko bardziej dokładne, gdyby badacz z góry mógł określić, o ile dana droga jest lepsza od pozostałej.

Oszacowane wielkości błędu danego rodzaju są użyteczne także w niektórych analizach typu II, mianowicie w tych, kiedy badaczowi chodzi o ocenę wyników otrzymanych przy zastosowaniu obydwóch dróg łącznie i o błąd popełniony przy badaniu całej, połączonej zbiorowości. Dotyczy to zwłaszcza analiz, w których

z różnych przyczyn badana zbiorowość dzieli się na grupy badane przy zastosowaniu różnych sposobów, np. wywiadu i ankiety. Mogą to być także sposoby mało różniące się od siebie, jak np. wywiad przeprowadzany w domu i miejscu pracy przez różnych ankierów czy różne ich grupy. (Oczywiście warunek zasadniczy, który musi być tutaj spełniony, dotyczy porównywalności grup.) Jeśli badacz wie, jak liczne były te grupy, może ustalić, w jakich granicach mieści się błąd obciążający wynik dla całej zbiorowości, błąd, którego źródła związane są z różnicami między drogami. Jednak w odróżnieniu od sytuacji, kiedy jest w stanie ocenić, który sposób jest lepszy, w tych przypadkach nie może powiedzieć o kierunku tego błędu.

Cechą charakterystyczną rodzajów analiz, o których teraz mowa, jest to, że ocenia się w nich wynik dotyczący całej zbiorowości na podstawie porównywania wyników odnoszących się do jej części składowych. Takimi częściami składowymi mogą być zwłaszcza wyniki odnoszące się do grup badanych przez różnych ankierów. W przypadku grup wyróżnionych na tej podstawie najłatwiej jest bowiem zapewnić ich porównywalność, ponieważ badacz może wpływać na to, kto z kim będzie przeprowadzać wywiad. Nic więc dziwnego, że wielu badaczy koncentruje się na analizach rozkładów odpowiedzi uzyskanych przez różnych ankierów. W analizach tych wykorzystuje się różne miary zmienności wyników, spowodowanej przez ankierów, a więc różnic między uzyskanymi przez nich wynikami netto⁴¹. Ponieważ zmienność ta jest czymś niepożądanym, w analizach tych formuluje się niekiedy wnioski praktyczne, zmierzające do jej zmniejszenia. Za jeden ze sposobów prowadzących do zmniejszenia zmienności tego rodzaju uważa się zwiększenie ilości ankierów. Ponieważ jednak z tym zabiegiem może łączyć się obniżenie poziomu ankierów, zawsze zaś wiąże się zwiększenie kosztów, w niektó-

⁴¹ Studia tego rodzaju omawia m. in. H. Hyman, *op. cit.*, s. 327-339. W studiach tych badacze posługują się pojęciem „błąd różnicowy netto” (*net differential effect*) i pojęciem międzyankierskiej zmienności (*inter-interviewer variation*). Pierwsze z nich odpowiada pojęciu „błąd związanego z różnicą netto”, drugie stanowi pewną jego modyfikację.

rych analizach opracowano formuły pozwalające ustalić, jaka ilość ankierów będzie optymalna dla danego badania przy uwzględnieniu jego kosztów⁴².

Przejdźmy na koniec do omówienia sytuacji, kiedy badacz nie jest w stanie określić, która z zastosowanych dróg jest lepsza, nie interesuje go także lub nie może interesować ocena wyniku łącznego. Sytuacja tego rodzaju występuje zawsze w analizach typu I, może także wystąpić w analizach typu II. W tych ostatnich ma miejsce zawsze, kiedy zbiorowości badane różnymi drogami nie stanowią łącznie jednej całości, np. kiedy są to różne zbiorowości badane w różnym czasie.

Możliwości oceny wyniku w tych sytuacjach są bardzo małe — oczywiście jeśli ograniczymy się do tych przypadków, kiedy rezultaty otrzymane różnymi drogami są od siebie różne. Jednak i tutaj trzeba wprowadzić pewne zróżnicowanie, biorąc mianowicie pod uwagę, czy dla badacza ma znaczenie ocena tylko jednego, czy też — osobno — dwóch rezultatów.

Jeśli badacza interesuje tylko wynik otrzymany drogą zastosowania wywiadu kwestionariuszowego lub danego typu wywiadu, ocena zaś pozostałych rezultatów nie jest dla niego ważna, wówczas jest w stanie o tym wyniku powiedzieć tylko tyle, że ewentualny błąd związany z różnicami netto i wyrównanymi (typ I) lub tylko netto (typ II) zbliżyć się będzie do granicy, którą wyznaczają te różnice. Oczywiście chodzi tu tylko o błąd danego rodzaju, tj. taki, którego źródła tkwią w różnicach cech między dwiema drogami. Takie bardzo skromne wnioski, nie dostarczające odpowiedzi na najważniejszy dla badacza problem, można wyciągnąć na przykład z zestawienia różniących się wyników powtórnych badań, zastosowanych do tej samej lub odmiennych, porównywalnych zbiorowości⁴³.

⁴² *Op. cit.*, s. 339.

⁴³ Badania takie omawia M. Parten, *Surveys Polls and Samples*, Nowy Jork 1950, s. 496-498. W rozdziale XVI pt. „Evaluation of the Data and Sample” charakteryzuje ona analizy weryfikacyjne, należące do różnych typów, podobnie jak Moser i Hyman w poprzednio cytowanych pracach.

Oczywiście dla danego wyniku jest lepiej, jeśli omawiane różnice są małe i, co można stwierdzić w analizach typu I, wyrównane. Ewentualne błędy danego rodzaju, obciążające wynik otrzymany przy zastosowaniu wywiadu, mogą być wówczas potraktowane jako przypadkowe. Nieco tylko lepsza jest sytuacja badacza wówczas, gdy w równej mierze interesują go obydwa wyniki. Sytuacja ta występuje najczęściej wtedy, gdy stosuje on dwukrotnie identyczne sposoby badania lub sposoby, które uważa za ściśle równoległe (np. pytania, na które odpowiedzi kontrolują się wzajemnie, ponieważ mają dostarczyć tej samej informacji). Jest ona lepsza o tyle, że w tym przypadku różnice między rezultatami, zwłaszcza duże, obciążają wynik ważny dla badacza niezależnie od tego, że nie wie on, o który wynik tu chodzi. Jeśli są to różnice małe i wyrównane, co może stwierdzić w analizach typu I, wtedy błędy można potraktować jako przypadkowe⁴⁴. Natomiast występowanie dużych różnic świadczy o tym, że zastosowanie identycznego sposobu było odmienne, lub też, gdy miał to być sposób ściśle równoległy, że się takim nie okazał. To ostatnie stwierdzenie ważne jest m. in. w przypadku budowania skali złożonej z wielu pytań na wzór skal psychologicznych i analizy jej rzetelności oraz w analizach, w których wykazuje się nierównoważność różnego sposobu zadawania pytań przy zastosowaniu techniki *split-ballot*⁴⁵.

⁴⁴ Tego rodzaju wnioski można było sformułować w analizach wspomnianych w przypisie 30, w których badano sprzeczności występujące w odpowiedziach na pytania wzajemnie się kontrolujące. W analizach tych stwierdzono ponadto, że różne kategorie respondentów w bardzo małym stopniu różnią się ze względu na częstotliwość sprzeczności występujących w ich wypowiedziach. Nie ma także powiązania między występowaniem sprzeczności w kwestionariuszu danego respondenta a popełnionymi przez niego błędami ustalonymi drogą weryfikacji przy zastosowaniu sposobu pewnego. Przemawia to za tym, iż błędy związane z tymi sprzecznościami mają charakter przypadkowy.

⁴⁵ Por. M. Parten, *op. cit.*, s. 495, H. Hyman, *Survey Design...*, s. 165. Technika *split ballot*, wprowadzona przez Cantrila, polega na tym, że respondentów dzieli się losowo na grupy, którym zadaje się pytania w różny sposób.

5. Wnioski dotyczące weryfikacji zewnętrznej

Dotychczasowe rozważania prowadzą do kilku wniosków. Pierwszy z nich dotyczy głównego problemu niniejszego artykułu, mianowicie kwestii oceny wartości poznawczej rezultatów otrzymanych w badaniach z zastosowaniem wywiadu kwestionariuszowego. Tylko analizy, w których wykorzystuje się drogę pewną, pozwalającą na sformułowanie ścisłej oceny wyrażonej w liczbach, zresztą przy założeniu, które trzeba uzasadnić, że jest ona pewna, oraz w przypadku analiz odnoszących się do dwóch zbiorowości, że są one porównywalne. We wszystkich innych analizach, należących do różnych typów i ich odmian, bądź nie można wcale sformułować oceny wyniku wyrażonej w liczbach, bądź też w takiej ocenie, zazwyczaj bardzo niedokładnej, uwzględnione być może występowanie błędu jednego tylko rodzaju.

W tej sytuacji celowe są poszukiwania nowych sposobów weryfikacji przy zastosowaniu drogi prawnej. Nie wydaje się bowiem, aby obecnie wykorzystywane były już wszystkie możliwości stosowania analiz omawianego rodzaju. Tak więc opierając się na opiniach kompetentnych sędziów można ustalić, jaki jest stopień religijności pewnej ilości znanych im dobrze osób, traktując te wyniki jako pewne. Te same osoby badamy także przy zastosowaniu wywiadu kwestionariuszowego. Porównanie obydwóch wyników pozwala ocenić rezultat badań. Wnioski z takiej analizy można rozszerzyć na inne badania, traktując ocenę liczbową tego wyniku jako ocenę metody. Analogiczne sposoby, nawiązujące m. in. do sposobów znanych w psychologii, były już dawno zalecane przez metodologów, są jednak bardzo rzadko wykorzystywane⁴⁶. Oczywiście zastosowanie powyższego sposobu wymaga bardzo dokładnego spreżycowania oraz konkretyzacji pojęcia religijności i jej stopni.

Przytoczony przykład nasuwa jeszcze inną uwagę, która dotyczy także praktyki badawczego postępowania w omawianym zakresie. Ocena kompetentnych sędziów nie jest drogą całkowi-

⁴⁶ G. Lundberg, *Social Research*, Longmans Green, 1941, s. 243.

cie pewną. Ale całkowicie pewny nie jest także wynik testu. piśmienności, polegający na tym, iż badany ma napisać kilka specjalnie dobranych zdań, przy sprecyzowaniu kryteriów poprawności napisanego tekstu. Jest bowiem możliwe, że ktoś, kto umie pisać, obojętnie z jakich powodów nie napisze tych zdań i powie, że nie umie tego zrobić, bądź też celowo napisze je niepoprawnie. Całkowicie pewne nie są z reguły także wszelkie dokumenty. W związku z tym badacze mają także na uwadze fakt, iż sformułowanie całościowej liczbowej oceny wyników może mieć bardzo duże znaczenie w konkretnej weryfikacyjnej analizie, nieraz dość liberalnie traktując wymagania dotyczące pewności wyniku osiągniętego drugą kontrolną drogą i stanowiącego kryterium oceny. Niejednokrotnie bowiem wyniki te, jak stwierdza Moser, „są prawdopodobnie pewne w tym stopniu, jaki występuje w ogóle w życiu. Stąd też trzeba się nimi zadowolić”⁴⁷. Takie stanowisko, ogólnie rzecz biorąc, wydaje się słuszne — przynajmniej w obecnym etapie rozwoju metodologicznej refleksji i badań z tego zakresu.

Tak więc jest niewątpliwie celowe rozbudowanie analiz, w których wykorzystuje się dane pewne i wzbogacanie stosowanych w nich sposobów. Powinno jednak iść to w parze z rozbudowywaniem analiz innych typów. Przemawia za tym szereg okoliczności. Po pierwsze, zakres analiz, w których wykorzystuje się dane pewne, mimo możliwości jego rozszerzenia jest i będzie zawsze ograniczony. Nie zawsze bowiem można uzyskać dane pewne, nawet kosztem dużych wysiłków, odnoszące się do wielu zjawisk, zwłaszcza do przeżyć psychicznych. Po drugie, analizy przy zastosowaniu drogi pewnej wymagają na ogół podjęcia nowych badań, z wyjątkiem sytuacji, kiedy korzysta się z ogólnie dostępnych danych urzędowych i podobnych — oczywiście jeśli

⁴⁷ C. A. Moser, G. Kalton, *op. cit.*, s. 397. Zacytowana opinia Mosera dotyczy rezultatów pogłębionych wywiadów weryfikacyjnych, przeprowadzonych regularnie z pewną ilością respondentów po każdym spisie powszechnym w St. Zjednoczonych. Są to wspomniane wyżej w tekście *quality checks* (sprawdzian jakości). Moser uważa, że organizowanie takich wywiadów powinno stać się regułą w dużych badaniach (s. 482).

są one pewne. Tymczasem w analizach, w których podstawę oceny stanowi fakt występowania lub niewystępowania oczekiwanych korelacji, można wykorzystać informacje zebrane w tym samym badaniu. Podobnie przedstawia się sprawa z analizami należącymi do pierwszego, szczegółowego typu. Po trzecie, zwłaszcza analizy ostatnich typów (III-IV) dostarczają danych, które rzucają światło na źródła błędów. Ustalenie źródeł błędów nie jest równoznaczne z oceną wyniku, jednak jest bardzo ważne ze względu na inne cele metodologiczne studiów.

Wszystkie te okoliczności wskazują na potrzebę przeprowadzenia i rozbudowania także tych analiz, które nie dostarczają podstaw do liczbowej całościowej oceny wyników badań. W związku z tym wydaje się celowe zwrócenie uwagi na sprawę, o której dotychczas nie było mowy. Mianowicie we wszystkich analizach wynik poddaje się pewnej próbie, pewnemu testowi. Próby, które stanowią owe „gorsze” analizy, odznaczają się między innymi tym, iż nie dotyczą one wyniku w całości. Sprawdzeniu podlega najczęściej jakiś jego aspekt. Z tej racji, jak zawsze w takich przypadkach, lepiej jest, gdy prób takich podejmuje się więcej, gdyż punkty widzenia, które one uwzględniają, mogą się uzupełniać. Przykładowo, jeśli badacz sprawdzi, że rezultaty otrzymywane przez lepszych i gorszych ankierów nie różnią się, wówczas nabiera pewności, iż w tym przypadku te ważne cechy procedury badawczej nie wpłynęły na wynik⁴⁸. Jeśli uzupełnić tę analizę innymi analizami, takimi, w których ten sam wynik poddany będzie innej próbie, polegającej np. na sprawdzeniu, czy cecha, której on dotyczy, jest skorelowana zgodnie z ugruntowaną już wiedzą z innymi cechami, wówczas jego za-

⁴⁸ Taką weryfikację przeprowadził m. in. A. Borucki w odniesieniu do niektórych wyników z badań nad percepcją różnicowania społecznego w wybranych miastach w Polsce. Różnice między rezultatami uzyskanymi przez różnych ankierów były minimalne. Upoważnia to do wniosku, że w danym przypadku wskazane różnice między ankierami nie odegrały roli prawdopodobnie ze względu na zastosowane szkolenie i kontrolę ankierów. (Wyniki tej weryfikacji nie były jeszcze opublikowane, opieram się tutaj na ustnej informacji A. Boruckiego.)

ufanie do danych wzrośnie w daleko większym stopniu — oczywiście jeśli te próby dadzą pozytywne rezultaty. W przeciwnym wypadku będzie mógł zrezygnować w swoim badaniu z tych danych, do których z określonych względów nie może mieć zaufania⁴⁹. Trzeba zresztą zauważyć, że na pewno nie umiemy jeszcze wykorzystać wszystkich możliwości, które stwarzają różne analizy. Stąd potrzeba rozwijania nad nimi teoretycznej refleksji równoległe do rozbudowanej praktyki w zakresie weryfikacji.

II. Weryfikacja wewnętrzna

1. Stosowania i konieczne założenia weryfikacji wewnętrznej

Gdy w rezultacie zewnętrznej weryfikacji negatywnie oceniamy wynik jakiegoś badania, ocenę tę można przenieść i na to badanie, na dany proces badawczy. W weryfikacji wewnętrznej odwraca się ten kierunek postępowania. Ocenia się najpierw proces, a potem rezultat. Rezultatem są zawsze informacje, które przynoszą odpowiedź na pytania kwestionariuszowe. Procesem jest ciąg zdarzeń od momentu, gdy badacz nadał określoną postać pytaniu kwestionariuszowemu, do zaklasyfikowania jednostki badania do jednej z kategorii badanych zjawisk. Są stwierdzający, że dana jednostka należy do tej, a nie innej kategorii, stanowi odpowiedź na pytanie badacza, zmniejsza jego niepewność i przenosi poszukiwaną informację. Stąd też oceniany proces można nazwać procesem formowania informacji.

W każdym badaniu, jak już wspominaliśmy, występuje wiele takich procesów, tyle, ile informacji chce się uzyskać od wszystkich respondentów. Są to procesy bardzo zróżnicowane. W ich skład wchodzi zarówno przejawy aktywności badacza, jak i ankietera (zadanie pytania, zapis odpowiedzi) oraz respondenta i kodera, a więc wszystkich osób, których postępowanie, a w

⁴⁹ Przykłady zastosowania wielu prób weryfikacji do jednych i tych samych, licznych zresztą wyników można znaleźć w cytowanej już pracy Lazarsfelda i Thielensa. Należy dodać, że wybór rodzaju tych prób może — i powinien — wiązać się z przewidywaniami badacza dotyczącymi źródeł i kierunków odchyleń.

pewnym zakresie i przeżycia psychiczne, są niezbędne do pojawienia się poszukiwanych informacji i bezpośrednio lub za pośrednictwem innych decydują o ich rodzaju bądź nieuzyskaniu⁵⁰.

Jakość tych procesów ze względu na wartość poszukiwanych lub otrzymanych informacji jest przez badacza oceniana co najmniej dwa razy: wówczas gdy decyduje on, że należy przystąpić do przeprowadzania wywiadów, i wtedy, kiedy po ich zakończeniu i co najmniej wstępnej analizie rezultatów ocenia metodę i przebieg badań w ich fazie terenowej. Z reguły jednak, jak już wspominaliśmy, że nie są to oceny znajdujące wyraz w liczbach choćby dlatego, że nie znamy sposobów umożliwiających przygotowanie takich ocen. Bardzo często są to przy tym oceny powierzchowne, które nie stanowią rezultatu głębszych rozważań i studiów. Pozostaje to w związku z faktem, że w socjologii — przynajmniej w Polsce — nie ma dotąd ogólnie przyjętego systemu kryteriów poprawności metodologicznej badania, w tym i poprawności jego zasad, drogą porównywania wyników niezależnych zaklasyfikowań itp.

W związku z tymi ocenami, a także niezależnie od nich podejmuje się liczne próby analizy procesu formowania informacji w konkretnych badaniach. Taką analizę badań Lazarsfelda i Thielensa, na którą już powoływaliśmy się, przeprowadził D. Riesman⁵¹. Jest ona najbardziej kompletna spośród znanych autorowi niniejszego studium. Obejmuje próbę przedstawienia wyczerpującej i zarazem oceniającej charakterystyki działalności ankieterów, między innymi przy użyciu koncepcji „stylu ankieterskiego”, a także reakcji na badania i pytania z kwestionariusza respondentów, w tym unikalne, jak się wydaje, w literaturze rozważania, poświęcone wpływowi na odpowiedzi ich obawy i chęci podwyższenia lub zachowania własnego prestiżu. W swej analizie Riesman wykorzystał własne doświadczenia respondenta

⁵⁰ Ciąg ten był szczegółowo omawiany przeze mnie w cytowanym artykule pt. *Analiza procesu*...

⁵¹ D. Riesman, *Some Observations on Interviewing in the Teacher Apprehension Study*, [w:] P. F. Lazarsfeld i W. Thielens, op. cit.

i doświadczenia tego rodzaju swoich współpracowników, oraz głównie materiały takie, jak raporty ankietatorów, sprawozdania z wywiadów przeprowadzonych z ankietarami, ankietę zorganizowaną wśród respondentów i wywiady z nimi, następnie nagrania magnetofonowe wywiadów i oczywiście teksty wypełnionych kwestionariuszy, wszelkiego rodzaju instrukcje i inne dokumenty dotyczące badań.

Można wymienić wiele analiz, także przeprowadzonych w Polsce, które poświęcono charakterystyce procesów otrzymywania informacji w konkretnych badaniach z zastosowaniem wywiadu kwestionariuszowego, a zwłaszcza samym wywiadem, i w których porusza się podobne zagadnienia, co w studium Riesmana, oraz posługuje się zbliżonymi materiałami⁵². W większości przypadków są to studia poświęcone różnym fragmentom i aspektom omawianych procesów, takich jak postępowanie ankietatorów i ich pomyłki, następnie sprawie szczerości responden-

⁵² W Polsce tego rodzaju studiów opublikowano dość dużo, głównie w kolejnych tomach *Analiz i prób...* Przykładowo wymienić tu można studia z pierwszego i trzeciego tomu tej serii: t. I: K. M. Słomczyński, *Wpływ oficjalnej i prywatnej sytuacji wywiadu na wypowiedzi respondentów w środowisku inteligencji* (tylko w zakresie samooceny szczerości wypowiedzi respondentów); K. Lutyńska, *Refleksje metodyczne w wywiadach z urzędnikami w Łodzi w latach 1960 - 1961*; t. III: S. Szostkiewicz, *Modelowa i rzeczywista sytuacja wywiadów prowadzonych przez ankietatorów społecznych w zbiorowościach lokalnych*; Z. Gostkowski, *Kontrola procesu badawczego w małej społeczności*; K. Brzozowska-Ejme, Z. Kawka, *Analiza błędów pracy ankietera w wywiadzie kwestionariuszowym na podstawie zapisu magnetofonowego*; M. D. Pełka-Sługocka, *Przeprowadzenie wywiadów z więźniami w zakładach karnych*; P. Pradervand, *Problemy społecznej sytuacji wywiadów w badaniach nad stosunkiem ludności do planowania rodziny w krajach słabo rozwiniętych*. W Polsce, podobnie jak w innych krajach, często wykorzystuje się w tych studiach tzw. ankietę do ankietatorów, tworząc niekiedy części kwestionariusza; por. w tej sprawie S. Szostkiewicz, *Ankietę do ankietatorów (próba typologii)*, oraz K. Lutyńska, *Ankietę do ankietera jako źródło wiedzy o wywiadzie kwestionariuszowym*, [w:] *Analizy i próby...*, t. II, 1968. Wymienione ankietę do ankietatorów niektórych rodzajów odpowiadają raportom ankietatorów wykorzystywanym przez Riesmana.

tów, roli ankietera i jej akceptacji u respondentów, sytuacji, w której przebiegają wywiady, atmosferze wywiadów i ich długości itp. Obok nich do analiz poświęconych tym procesom zaliczyć należy także analizy samych pytań ze względu na ich poprawność, następnie studia mające za przedmiot problem, jak respondenci rozumieją pytania i odpowiedzi oraz jak można globalnie ocenić wiarygodność respondentów.

Kwestie związane z rozumieniem pytań analizują najczęściej sami badacze, wykorzystując samorzutne komentarze respondentów lub odpowiedzi na celowo zadane pytania w kwestionariuszu, a zwłaszcza pytania otwarte. Niekiedy bada się ten problem również zadając odpowiednie pytania po wywiadzie. Wiarygodność respondentów usiłuje się ustalić drogą analizy ich odpowiedzi na specjalne pytania, tzw. puste. Są to pytania dotyczące na przykład rzeczy nie istniejących, na które prawidłowa reakcja polega na zakwestionowaniu ich założeń. Jeśli badacz udzieli odpowiedzi na to pytanie zgodnie z zawartym w nich żądaniem, wówczas jego wiarygodność, a w każdym razie przydatność jako respondenta jest kwestionowana. Negatywnie można ocenić wówczas i procesy formowania informacji, w których on uczestniczy⁵³.

Analizy, o których mowa, powiększają naszą wiedzę o funk-

⁵³ Analizę problemu, jak respondenci rozumieli poszczególne kluczowe terminy występujące w pytaniach i odpowiedziach na podstawie ich samorzutnych komentarzy, przeprowadzali między innymi P. F. Lazarsfeld i W. Thielens w cytowanej pracy (por. s. 86, 99 i inne). Por. także H. Hyman, *Survey Design*, s. 153, oraz K. Lutyńska, *Pilotaż pogłębiony. Koncepcja, realizacja i opracowanie materiałów*, w niniejszym tomie *Analiz i prób...*, oraz W. E. Szlapientoch, *Problemy dostawności statystycznej informacji w socjologicznych badaniach*, Moskwa 1973, s. 79-80. Problem wykorzystania pytań pustych (*sleepy questions*) omawia m. in. H. Hyman (op. cit., s. 161). Por. także P. Daniłowicz, *Próba zastosowania pytań pustych w badaniach pilotażowych*, w niniejszym tomie *Analiz i prób...*, oraz W. Bielecka, *Pytania puste w wywiadzie kwestionariuszowym. Wybrane zagadnienie*, 1974 (praca magisterska przygotowana w Zakładzie Metod i Technik Badań Społecznych Uniwersytetu Łódzkiego).

cjonowaniu wywiadu kwestionariuszowego, przynoszą lepsze zrozumienie tej techniki w jej różnych zastosowaniach. Pozwalają niejednokrotnie wskazać na jej wady i słabości i umożliwiają sformułowanie wniosków, które zmierzają do ich usunięcia. W niektórych analizach spośród obecnie rozpatrywanych, ze studium Riesemana na czele, formułuje się także oceny wyników badań.

Oceny te rzadko znajdują jednak wyraz w liczbach. Dzieje się tak najczęściej dlatego, że badacz-metodolog często wykorzystuje dane niepełne lub w ogóle opiera się na informacjach, które dotyczą zjawisk nieprzeliczalnych. Ale i w tych analizach, w których podaje się wyniki liczbowe, z reguły nie są to wyniki odnoszące się do liczebności informacji zaakceptowanych czy błędnych. Co najwyżej badacz przedstawia liczbowe dane, na których podstawie można przypuszczać, że ilość błędów jest mała czy duża. Tak więc Rieseman, który do sprawy ilości błędów przywiązuje dużą wagę, na zakończenie części analizy poświęconej ogólnej ocenie pisze: „Przypuszczam (*I would guess*), że nie więcej niż 6 lub 7 procent wywiadów zawiera poważne błędy w tym, co przekazują...”⁵⁴.

Ostrożność Riesmana jest uzasadniona. Istnieją co najmniej dwa powody, dla których formułowanie wniosków o ilości zaakceptowanych czy błędnych informacji na podstawie analiz omawianego rodzaju jest bardzo ryzykowne. Po pierwsze, co w mniejszym stopniu odnosi się do jego analizy, lecz dotyczy wszystkich analiz cząstkowych, badacz nie uzyskuje zazwyczaj danych o wszystkich ogniwach każdego procesu otrzymywania informacji, lecz co najwyżej o niektórych z nich. Nie dostarcza to podstaw do oceny informacji w całym badaniu. Jeśli bowiem w całym procesie otrzymywania informacji zachowanie ankietera było w pełni poprawne, nie oznacza to przecież, że respondent nie sformułował swojej odpowiedzi pod wpływem chęci pokaza-

nia się lepszym, niż jest w rzeczywistości. Po drugie, występowanie w procesach formowania informacji negatywnie ocenianych zjawisk nie zawsze musi prowadzić do błędów w tych informacjach. Tak więc, przykładowo, niewłaściwe postępowanie ankietera, polegające na niedosłownym zapisie odpowiedzi lub nieprzedstawieniu się respondentowi z nazwiska, czego wymagał od niego badacz, niekoniecznie musi pociągnąć za sobą występowanie błędów w otrzymanych informacjach. Podobnie o błędach tego rodzaju nie świadczą także oceny ankieterów, zgodnie z którymi dane pytanie czy kwestionariusze w ogóle były dla respondentów trudne lub drażliwe, czy też, że atmosfera wywiadu była napięta. Taką samą niepełną wagę w omawianej kwestii mają negatywne oceny respondentów, na przykład takie, w których stwierdza się, iż ankieter prowadził rozmowę jak biurokratyzowany urzędnik. Nie stanowią one wystarczających świadectw w tym zakresie, chociaż mogą być bardzo użyteczne przy formułowaniu ogólnej, jakościowej oceny badań. Podobnie twierdząca odpowiedź na pytanie puste: „Czy czytuje Pan «Przegląd Kulturalny»?” nie oznacza, że dany respondent nie czyta żadnego z wychodzących obecnie czasopism, które wymienił ponadto w swojej odpowiedzi⁵⁵.

Czy jest wobec tego możliwe otrzymanie jakiegokolwiek liczbowej oceny wartości otrzymanych informacji drogą weryfikacji wewnętrznej? Wydaje się, że tak. Analizy weryfikacyjne, które miałyby do tego doprowadzić, muszą mieć jednak specjalny charakter. Ocena każdego procesu otrzymywania informacji musi być w nich dokonana po zestawieniu go z wyraźnie sprecyzowanym modelem tego procesu, odnoszącym się do prawidłowego przebiegu zdarzeń, które się nań składają. Byłby to więc model normatywny. Zgodność z nim lub brak zgodności w jakimkolwiek ogniwie — decydowałaby o pozytywnej lub negatywnej ocenie rzeczywistych procesów, a tym samym i otrzymanych dzięki nim informacji. Model taki byłby w zasadzie odrębny dla każdego pytania w związku z daną poszukiwaną informacją, choć dla

⁵⁴ D. Riesman, *op. cit.*, s. 269. Dane liczbowe, które przytacza Riesman, odnoszą się na przykład do ilości osób, które uważały, że zapis w kwestionariuszu nie przedstawia wiernie ich poglądów, które krytykowały kwestionariusz i jego poszczególne pytania itp.

⁵⁵ For. H. H y m a n, *op. cit.*

poszczególnych rodzajów pytań można tworzyć modele bardziej generalne, wymagające jednak uszczegółowienia w przypadku każdego pytania.

2. Sposoby budowania normatywnego modelu formowania informacji

W jaki sposób można dojść do sprecyzowania takich modeli? Trudno jest w tym zakresie odwołać się wyłącznie do dyrektyw zawartych w podręcznikach metodologii i metodyki badań. Mają one bowiem na ogół charakter reguł, które zalecają pewne sposoby postępowania lub na odwrót — nie zalecają ich, zależnie od różnych, często bliżej nie sprecyzowanych warunków. Rzadko są to natomiast wyraźne nakazy i zakazy, odnoszące się do dokładnie sprecyzowanych okoliczności. Tylko te ostatnie mogą natomiast stanowić podstawę przy konstrukcji modelu. Poza tym, jak wiadomo, dyrektywy te nie tworzą zwartego zespołu, których stosowanie gwarantuje z jakimś określonym prawdopodobieństwem wartość otrzymywanych informacji. W związku z tymi okolicznościami dyrektywy zawarte w podręcznikach, które oczywiście należy wykorzystywać w konstrukcji modelu, nie mogą stanowić wyłącznej jego podstawy.

Nasuwa się możliwość innej drogi budowania modeli, o które chodzi, do pewnego stopnia analogicznej do tej, którą przy ocenie wiarygodności informatora proponuje J. Giedymina i która według jego interpretacji jest stosowana przez historyków przy akceptacji twierdzeń o faktach historycznych, zawartych w źródłach⁵⁶. Wykorzystanie tej drogi wymaga, aby badacz, z jednej strony, posiadał dokładne dane dotyczące przebiegu pewnej, dużej ilości procesów otrzymywania analogicznych informacji, z drugiej zaś, wiedział z góry, czy te informacje są prawdziwe. Musi on więc oprzeć się na wynikach weryfikacji przeprowadzonej przy użyciu sposobu pewnego w odniesieniu do tej samej zbiorowości. Następnie zestawiając ze sobą procesy, których rezultaty były pozytywne oraz negatywnie zweryfikowane, można

⁵⁶ J. Giedymina, *Problemy, założenia, rozstrzygnięcia...*, rozdział VIII i IX. W koncepcjach Giedymina podstawową rolę pełni pojęcie wiarygodnego informatora.

dojść do obrazu i modelu danego procesu, który gwarantuje pomysłny, a także niepomysłny rezultat, zawsze czy z określonym prawdopodobieństwem. Z tym modelem należy w dalszym ciągu analizy zestawiać inne procesy i na tej podstawie ustalać, czy miały one przebieg prawidłowy, czy nie. To z kolei upoważnia do oceny, czy — ewentualnie z jakim prawdopodobieństwem — można przyjąć, że informacje będące rezultatem tych innych procesów są prawdziwe.

Zbudowane w ten sposób modele posiadałyby wiele zalet. Byłyby to modele w pełni empiryczne. Włączając do nich twierdzenia, dotyczące poszczególnych ogniw i aspektów procesów formowania informacji, można także starać się o to, aby wśród nich było możliwie najwięcej takich twierdzeń, które odnoszą się do zachowań ankietera, a więc do zjawisk w pełni kontrolowanych przez badacza. W ten sposób budowa modelu byłaby połączona z udoskonaleniem techniki wywiadu⁵⁷.

Trzeba jednak zauważyć, że wskazana metoda budowania modelu procesów formowania informacji byłaby niesłychanie pracochłonna, a w przypadku wielu pytań i poszukiwanych informacji jest po prostu nie do zastosowania ze względu na niemożliwość znalezienia pewnych dróg uzyskania danych. Stąd też wątpliwe jest, aby można ją było wykorzystać do tych celów.

⁵⁷ Do pewnego stopnia analogiczny sposób wprowadzania udoskonaleń do techniki wywiadu stosuje się w Ośrodku Badań Surveyowych Uniwersytetu w Michigan, kierowanym przez Ch. F. Cannella. Sposób ten, polega na tym, iż w odniesieniu do jakiejś zbiorowości respondentów uzyskuje się dane z dokumentów (chodzi tu niemal wyłącznie o informacje o takich zjawiskach, jak przebyte choroby, pobyt w szpitalu, wizyty u lekarzy itp.), a następnie stosuje się różne sposoby uzyskiwania informacji, zwłaszcza wywiad kwestionariuszowy w różnych formach, celem ustalenia, który z tych sposobów daje najlepsze rezultaty. Por. np. między innymi następującą pozycję z serii pt. *Metodologia Wywiadu i Ankiety*, wydawaną przez ten ośrodek: Ch. F. Cannel, K. M. Marquis, A. Laurent, *An Experimental Study of the Effects of Reinforcement, Questions Length and Reinterviews on Reporting Selected Chronic Conditions in Household Interviews*, Michigan 1969. Charakteryzując ten sposób opieram się także na ustnej informacji prof. Cannella, dotyczącej badań metodologicznych w jego ośrodku.

Jeśli więc chcemy w rezultacie weryfikacji wewnętrznej otrzymać dokładną liczbową ocenę wyniku badań z zastosowaniem wywiadu kwestionariuszowego, to trzeba znaleźć inną drogę zbudowania modeli procesów formowania informacji. Najbardziej właściwa wydaje się droga, w której wykorzystuje się niektóre ustalenia rozbudowanej w Polsce refleksji nad wskaźnikami oraz nad pytaniami, przeprowadzonej z zastosowaniem aparatu logiki. Można ją w związku z tym nazwać metodą „logiczną”, używając cudzysłowu choćby dla zaznaczenia, że nie należy ona do logiki.

3. „Metoda „logiczna” budowania modelu prawidłowego przebiegu procesu otrzymywania informacji

Punktem wyjścia przy budowaniu omawianego modelu tą metodą może być ogólny schemat wnioskowania ze wskaźników o indicatum, czyli badanym zjawisku, przedstawiony przez E. Mokrzyckiego⁵⁸. Schemat ten należy przystosować do procesu otrzymywania informacji, związanego z jednym pytaniem w kwe-

⁵⁸ E. Mokrzycki, *Założenia socjologii humanistycznej*, Warszawa 1971, s. 71–72. E. Mokrzycki wykorzystuje w swoich rozważaniach refleksje i ustalenia T. Pawłowskiego i S. Nowaka. Opierając się na wynikach rozważań Mokrzyckiego, muszę jednocześnie zauważyć, że nie podpisuję się pod jednym z jego stwierdzeń, odnoszącym się do rozumowań na podstawie wskaźników. Tak więc zgodnie z tym, o czym była mowa poprzednio, uważam za niesłuszne w odniesieniu do badań kwestionariuszowych jego następujące sformułowanie: „Socjolog ma [...] znacznie więcej do czynienia z wnioskowaniem o jakieś własności na podstawie szeregu jej wskaźników (bądź raczej na podstawie szeregu jej domniemanych wskaźników) [...] niż z wnioskowaniem o indicatum na podstawie jednego wskaźnika”, s. 72. Jeśli w badaniach kwestionariuszowych za wskaźnik uważać będziemy odpowiedź na pytanie (najczęściej jest to zresztą wskaźnik specjalnego typu, mianowicie zdanie opisujące dany stan rzeczy), wówczas za najczęstsze musimy uznać właśnie sytuację, kiedy z jednym pytaniem (wskaźnikiem) łączy się odpowiedź na jedno pytanie badacza, a więc i jedno indicatum. Można zgodzić się jednak z wyżej przytoczonym poglądem, jeśli odniesiemy go do stanu rzeczy pożądanego w socjologii.

stionariuszu, nie zmieniając jego założeń. Będzie miał on wówczas następującą postać.

Przesłanki:

I. Respondent R_1 na pytanie P w wywiadzie odpowiedział w sposób O_1 (względnie O_2, O_3 itd.).

II. Jeśli respondent z kategorii R na pytanie P w wywiadzie odpowiedział w sposób O_1 (O_2, O_3 itd.), to odpowiadająca mu jednostka badania należy do klasy K_1 (K_2, K_3 itd.).

Wniosek:

III. Jednostka badania odpowiadająca respondentowi R_1 należy do klasy K_1 (K_2, K_3 itd.).

Schemat powyższy to oczywiście zastosowanie trybu *ponendo ponens*. Odnosi się on do procesów otrzymywania informacji nie tylko w wywiadzie kwestionariuszowym, lecz przy wszelkim wypytywaniu. Schemat stanowi najbardziej ogólny, czysto logiczny model procesu otrzymywania informacji, o który nam chodzi. Wniosek (III) stanowi zarazem odpowiedź na pytanie badacza, a więc przenosi poszukiwaną informację o jednostce badania. Najczęściej bywa nią respondent z danej próbki oznaczonej literą R , ale jak już wspominaliśmy, nie jest tak zawsze i stąd potrzeba wprowadzenia do schematu pojęcia jednostki badania. Przesłanka pierwsza charakteryzuje najogólniej, co dzieje się w danym fragmencie wywiadu. Przesłankę drugą stanowi założenie, na podstawie którego dochodzi się do poszukiwanej informacji.

Gdy rozumowanie według tego schematu jest poprawne, wniosek, a więc poszukiwana informacja, będzie zawsze prawdziwy, jeśli prawdziwe są obydwie przesłanki. Pierwsza z nich orzeka, że w danym przypadku wystąpił jakiś stan rzeczy. To, czy przesłanka ta jest prawdziwa, należy stwierdzić empiryczną drogą. Druga przesłanka stwierdza pewną zależność ogólną. Nie zawsze jest ona twierdzeniem empirycznym, jak może się wydawać na pierwszy rzut oka. Może być i zdaniem analitycznym. W jej skład — również wtedy, gdy zawiera ona twierdzenia empiryczne — mogą wejść także definicje, a więc propozycje terminologiczne. Twierdzeniem analitycznym jest ona wtedy, gdy jednostkę badania stanowi respondent, a jego cechą, na podstawie któ-

rej zalicza się go do danej klasy badanych zjawisk, jest tylko to, iż odpowiedział w pewien sposób na pytanie kwestionariusza. W takich sytuacjach przesłanka druga jest zawsze prawdziwa na mocy zwyczaju językowego. Kiedy jest ona twierdzeniem empirycznym, co zresztą zdarza się daleko częściej, wówczas jej prawdziwość należy stwierdzić odwołując się do doświadczenia. O tych różnych sytuacjach będzie zresztą mowa później.

Jakie możliwości w zakresie budowania interesującego nas modelu wiążą się z tym schematem? Pewne takie możliwości ujawnia zestawienie schematu i jego przesłanek z listą kolejnych ogniw procesu otrzymywania informacji. Model musi składać się z twierdzeń, dotyczących prawidłowego przebiegu tego procesu. Mogą to być także twierdzenia odnoszące się do jego ogniw i cech tych ogniw. Twierdzenia te, a więc i model, powinny odznaczać się tym, że jeśli w zastosowaniu do jakiegoś procesu wszystkie one okażą się prawdziwe, prawdziwa będzie i poszukiwana informacja. Chcąc, aby tak było rzeczywiście, należy dobierać je w ten sposób, żeby ich prawdziwość pociągała za sobą prawdziwość obydwóch przesłanek lub przynajmniej pierwszej z nich w tych przypadkach, gdy druga nie wymaga empirycznego uzasadnienia. Wtedy gdy twierdzenia te okażą się prawdziwe, prawdziwe będą i przesłanki oraz poprawnie wyprowadzony z nich wniosek. Oznacza to, że prawdziwa będzie identyczna z tym wnioskiem poszukiwana informacja, odpowiedź na pytanie badacza. Wówczas cel, dla którego zbudowany jest model, będzie osiągnięty.

Konstruowanie modelu według tych zasad należy rozpocząć od wymienienia kolejnych ogniw procesu otrzymywania informacji. Ponieważ realizuje się on w toku wzajemnego komunikowania badacza, ankietera, respondenta i kodera, takie naturalnie odrębne od siebie ogniwa stanowią przekazy, kierowane kolejno przez jedne z tych osób do innych. Są to przekazy następujące: I. pytanie sformułowane przez badacza jako pytanie kwestionariuszowe, II. pytanie wydrukowane (napisane) w tekście kwestionariusza, III. pytanie zadane respondentowi, IV. odpowiedź (reakcja słowna lub milczenie) respondenta, V. zapis odpowiedzi res-

pondenta, VI. symbol klasyfikacyjny kodera wraz z interpretacją. Ostatni, VII przekaz stanowi otrzymana informacja⁵⁹.

Które z tych przekazów odznaczać się powinny określonymi cechami ze względu na pierwszą przesłankę i jakie są to cechy, jeśli ta przesłanka ma być prawdziwa? Głosi ona, że respondent R_1 na pytanie P odpowiedział w sposób O_1 . Jeśli tak ma być rzeczywiście, to poszczególne przekazy — osobno czy w powiązaniu z innymi — charakteryzować powinny następujące cechy:

1. Pytanie w kwestionariuszu jest identyczne z pytaniem, które przygotował badacz (przekazy I i II).

2. Pytanie zadane respondentowi jest równoważne pod względem treści pytaniu w kwestionariuszu (przekazy II i III).

3. Odpowiedź respondenta ujęta w zapisie jest autentyczna, więc jest odpowiedzią tego respondenta na to pytanie (przekaz IV).

4. Zapis odpowiedzi w kwestionariuszu uwzględnia wszystkie jej właściwości, które służą za podstawę jej zaliczenia do jednej z klas przez kodera (przekaz V).

5. Klasyfikacja odpowiedzi, którą wykorzystuje koder, stanowi podział logicznie poprawny (przekaz VI).

6. Zaklasyfikowanie odpowiedzi przez kodera na podstawie zapisu kodera jest poprawne, tj. wpisany symbol kodowy odpowiada klasie, do której należy ta odpowiedź ze względu na swój zapis (przekaz VI).

Cecha wymieniona w twierdzeniu 5 posiada tylko pośrednie znaczenie dla oceny prawdziwości omawianej przesłanki ze względu na twierdzenie 6. Poprawne zaliczenie jednostek do różnych kategorii jest bowiem możliwe tylko wtedy, gdy te kategorie

⁵⁹ Por. cytowany artykuł *Analiza procesu otrzymywania informacji...* W artykule tym przekazy 1 i 2 utożsamiono ze sobą. Odróżnienie ich wydaje się potrzebne nie tylko ze względu na to, że w druku kwestionariusza występują pomyłki, lecz i z tego powodu, że badacze niekiedy wprowadzają zmiany do pytań już po ich wydrukowaniu (powieleniu), co często nie jest już brane pod uwagę przez ankieterów w pracy terenowej. Natomiast w interpretacji wyników badacz uwzględnia na ogół to nowe sformułowanie, co może prowadzić do błędów.

tworzą poprawnie zbudowany podział. (Kategorie te określają właśnie sposoby odpowiadania, które to sformułowanie występuje w ogólnym schemacie.) Pozostałe twierdzenia mają bezpośrednie znaczenie dla prawdziwości pierwszej przesłanki.

Jeśli odpowiednie przekazy w procesie otrzymywania informacji, związane z pierwszą przesłanką w schemacie. Należy przy dzeniach, wówczas na podstawie wiedzy o tym, jaki symbol klasyfikacyjny wpisano przy odpowiedzi, można prawdziwie orzec, że dany respondent na dane pytanie odpowiedział w taki, a nie inny sposób. Wymieniając te cechy wskazuje się więc wszystkie elementy prawidłowego przebiegu procesu otrzymywania informacji, związane z pierwszą przesłanką w schemacie. Należy przy tym podkreślić, iż te wszystkie elementy, związane z daną przesłanką, są zawsze jednakowe, niezależnie od rodzaju pytania w kwestionariuszu i poszukiwanej informacji. To ostatnie przekonanie uzasadnia choćby fakt, że przy dobieraniu twierdzeń nie potrzebowaliśmy odwoływać się do treści tych pytań i rodzaju informacji. Staraliśmy się jednak przy rozważaniu kolejnych przekazów i ich cech, tak dobrać te twierdzenia, aby łącznie uzasadniały one przekonanie o prawdziwości pierwszej przesłanki.

Przesłanka druga, jak wspomnieliśmy, jest zdaniem ogólnym. Odnosi się więc ona także do każdego procesu otrzymywania informacji osobno. Uzależnia ona rodzaj jednostkowej informacji, tj. zaliczenie jednostki badania do danej klasy, od przynależności odpowiedzi respondenta do jednej kategorii (sposobu) odpowiedzi. Zestawienie tej przesłanki z przedstawioną poprzednio listą przekazów, bez odwołania się do pytań i poszukiwanych informacji, pozwala wskazać tylko na dwie cechy, charakteryzujące prawidłowy przebieg procesu otrzymywania informacji. Pierwsza z nich odnosi się do przekazu VII, a więc do samej poszukiwanej informacji. Przekaz ten zakłada, że jednostki badania zalicza się do poszczególnych klas w podziale. Zaliczenie to może być poprawne tylko wtedy, gdy podział jest poprawny pod względem logicznym. Twierdzenie dotyczące tej cechy jest więc analogiczne jak twierdzenie 5, tj. twierdzenie o poprawności klasyfikacji wykorzystywanej przez koder (por. listę na s. 373).

Druga cecha to odpowiedniość między tymi dwiema klasyfikacjami. Jest to odpowiedniość tego rodzaju, iż przynajmniej grupy klas w klasyfikacji koder a poszczególne klasy badanych zjawisk (jednostek badania) są sobie podporządkowane wzajemnie jednoznacznie⁶⁰. Najczęściej zresztą to wzajemne jednoznaczne podporządkowanie dotyczy poszczególnych klas w obydwóch podziałach. Omawiana cecha i odpowiadające jej twierdzenie odnoszą się do przekazu VI i VII. Obydwa przedstawione ostatnio twierdzenia uzupełniają listę twierdzeń, podaną na s. 373, zwiększając ich liczbę do 8.

W procesie otrzymywania informacji, którego przebieg jest prawidłowy, obydwie wymienione wyżej cechy muszą występować zawsze, podobnie jak cechy wyróżnione w związku z przesłanką pierwszą. W przeciwstawieniu do tych ostatnich ich występowanie nie gwarantuje jednak prawdziwości odpowiadającej im przesłanki. Może przecież zdarzyć się tak, że zaklasyfikowanie danej jednostki badania — a więc otrzymana informacja — okaże się błędna mimo logicznej poprawności klasyfikacji i jej zgodności z klasyfikacją odpowiedzi. Prawdziwość twierdzeń dotyczących występowania tych cech posiada jednak znaczenie dla omawianej przesłanki, a w związku z tym jej prawdziwości. Jeśli cechy te nie występują, nie można jej poprawnie sformułować w zastosowaniu do danego procesu, a właściwie wszystkich procesów związanych z danym pytaniem z kwestionariusza i pytaniem badacza. Nie można bowiem sensownie mówić o zależności między przynależnością do klas w różnych klasyfikacjach, jeśli klasy te nie są poprawnie wyróżnione i jeśli nie ma między nimi odpowiedniości o wspomnianym charakterze. Niepoprawnie sformułowana przesłanka nie może być uznana za prawdziwą niezależnie od jej tautologicznego czy empirycznego charakteru. Ta uwaga odnosi się zresztą i do przesłanki pierwszej ze względu na klasyfikację, którą posługuje się koder przy zaliczaniu odpowiedzi respondenta do jednej z klas (przekaz V). Warto jeszcze dodać, że

⁶⁰ Mam tu na myśli takie sytuacje, gdy przy budowaniu tabel badacz łączy ze sobą różne kategorie wyróżnione przy kodowaniu.

twierdzenie odnoszące się do odpowiedniości wspomnianych klasyfikacji łączy jeszcze jedna cecha. Ich prawdziwość lub fałszywość można wykazać bez odwoływania się do przebiegu poszczególnych procesów otrzymywania informacji, co jest konieczne w odniesieniu do twierdzeń pozostałych, drogą analizy różnych dokumentów dotyczących opracowania odpowiedzi na dane pytanie. W tych bowiem dokumentach, włączając i zestawienia liczbowe, zwłaszcza wstępne, oraz ich interpretacje, przedstawia się wspomniane klasyfikacje, znajdują w nich także odzwierciedlenie stosunki między nimi w ujęciu badacza.

Występowanie wymienionych ośmiu cech procesu otrzymywania informacji stanowi więc warunek konieczny, ale nie wystarczający, aby były prawdziwe obydwie przesłanki, co — przy poprawnym wyprowadzeniu wniosku — gwarantuje jego prawdziwość, a w związku z tym i prawdziwość jednostkowej informacji. W związku z tym mogłoby się wydawać, że twierdzenia dotyczące owych ośmiu cech nie mogą nigdy tworzyć całości poszukiwanego modelu. Przypuszczenie takie nie jest jednak słuszne. Tworzą one model wystarczający dla oceny procesu otrzymywania informacji wówczas, gdy twierdzenie lub twierdzenia, składające się na przesłankę drugą, mają charakter analityczny, gdy jest ona tautologią, co, jak wiemy, może wystąpić. Czy tak jest rzeczywiście w przypadku danego procesu, zależy od tego, jaki jest stosunek pytania z kwestionariusza i rodzaju odpowiedzi na nie do pytania badacza wraz z odpowiedzią.

Omawiana przesłanka posiada charakter tautologiczny wówczas, gdy — jak wspominaliśmy — pytanie badacza dotyczy tylko kwestii, w jaki sposób dany respondent odpowiedział na postawione mu pytanie. Takie pytanie badacza jest identyczne z pytaniem, które w odniesieniu do każdej odpowiedzi respondenta stawia sobie koder, gdy oznacza ją odpowiednim symbolem. W związku z tym można je nazwać pytaniem kodera.

Czy sytuacje tego rodzaju występują rzeczywiście w badaniach, czy też jest to tylko sytuacja teoretyczna? Wbrew temu, co można przypuszczać, zdarza się ona bynajmniej nierzadko. O jej występowaniu świadczy sposób, w jaki badacz przedstawia

liczbowe wyniki swoich poszukiwań, często ujęte w tabelę. Ogranicza się on wówczas do stwierdzeń następujących rodzajów: „w odpowiedzi na pytanie X tylu respondentów odpowiedziało w ten, a tylu w ten sposób” lub „spośród respondentów tylu wymieniło w odpowiedzi to a to, a tylu coś innego” lub „w odpowiedzi na pytanie X tylu respondentów wskazało na taką opinię, a tylu na inną” itp. W takich wypowiedziach badacz stwierdza w odniesieniu do całej zbiorowości tylko to, co w stosunku do każdej jednostki stwierdził koder. Nie wiąże z tymi odpowiedziami żadnej dalej idącej interpretacji, a otrzymaną przez niego informację można uznać za skrajnie minimalistyczną pod względem poznawczym. Badacz wypowiada się tylko o tym, co sam wywołał w swoich poszukiwaniach, nie mówi nic o rzeczywistości zewnętrznej w stosunku do nich.

Dla takich sytuacji wystarczający jest model prawidłowego przebiegu procesu formowania informacji, obejmujący tylko twierdzenia o występowaniu wskazanych wyżej cech. Gdy te twierdzenia są prawdziwe, wówczas obydwie przesłanki są sformułowane poprawnie, zagwarantowana jest prawdziwość przesłanki pierwszej, a druga jest prawdziwa ze względu na swój tautologiczny charakter.

Okazuje się więc, że model prawidłowego przebiegu procesu otrzymywania informacji, składający się z ośmiu przedstawionych wyżej twierdzeń, jest wystarczający w niektórych sytuacjach. Obok nich należy wspomnieć o sytuacjach, kiedy do takiego modelu należy wprowadzić tylko minimalne zmiany.

Są to sytuacje, gdy pytanie badacza różni się od pytania, które zadaje sobie koder w odniesieniu do odpowiedzi respondenta, tylko terminologią. To znaczy, że w pytaniu badacza i w odpowiedziach na nie używa się terminów innych niż w pytaniu kodera i odpowiedziach na nie — a więc i w pytaniu z kwestionariusza wraz z odpowiedziami — lecz te terminy zdefiniowano odwołując się tylko do rodzaju odpowiedzi respondenta. Powiedzmy, że za człowieka, którego cechuje dana postawa, uznaje się na mocy definicji takiego i tylko takiego respondenta (jest on tutaj także jednostką badania), który na dane pytanie odpowiedział

twierdząco, np. zaaprobował wymienioną w pytaniu opinię. Został na przykład uznany za rygorystę, ponieważ zaaprobował odpowiedź: „Należy stosować karę fizyczną jako środek wychowawczy”. Sąd o rygoryzmie respondenta oznacza przy tym tutaj tylko tyle, że respondent na określone pytanie udzielił takiej, a nie innej odpowiedzi. Mimo odmiennej terminologii otrzymana informacja w tym przypadku jest także skrajnie minimalistyczna. Jednocześnie jednak model prawidłowego przebiegu procesu otrzymywania informacji, w skład którego wchodzi wszystkie wymienione poprzednio twierdzenia, należy uzupełnić o twierdzenie nowe, o określenie pojęcia rygoryzmu. Jest to definicja częściowa, mająca postać zdania redukcyjnego. Jako ustalenie terminologiczne nie podlega ona ocenie z punktu widzenia prawdy i fałszu, chociaż dlatego, aby rozumowanie, w którym się ją wykorzystuje, było poprawne, musi spełniać ona inne wymagania.

To nowe twierdzenie dostarcza uzasadnienia dla przesłanki drugiej z pierwotnego schematu. Posługując się poprzednio używanymi oznaczeniami można mu nadać następującą postać: Na mocy definicji jednostka badania, odpowiadająca respondentowi z kategorii R , należy do klasy K_1 (K_2 itd.), jeśli ten respondent na pytanie P w wywiadzie odpowiedział w sposób O_1 (O_2 itd.). Należy jeszcze dodać, że jednostkę badania stanowi w tych przypadkach z reguły sam respondent.

Przedstawiona sytuacja zdarza się w badaniach socjologicznych, chociaż najczęściej trudno jest stwierdzić z pewnością jej występowanie. Utrudnia to okoliczność, że nie zawsze wiadomo, czy badacz ustalając, kiedy będzie się posługiwał danym terminem, przedstawia jego definicję, a kiedy to ustalenie traktuje tylko jako zabieg techniczny, podczas gdy właściwie przez dany termin (np. „rygoryzm”) rozumie co innego. W tym ostatnim przypadku badacz liczy na występowanie zależności empirycznej między udzieleniem odpowiedzi określonego rodzaju a posiadaniem przez respondenta cech wskazanych przez niezależnie sformułowaną definicję. W związku z tym omawiana sytuacja przekształca się w inną, o której będzie mowa później.

Sytuacje, kiedy w porównaniu z pierwotnym schematem mo-

del trzeba uzupełnić o twierdzenie definicyjne wskazanego wyżej typu, zdarzają się zapewne nieczęsto. Niezależnie jednak od częstości występowania zasługują one na uwagę z tego powodu, że wskazują sposób koniecznego uzupełnienia modelu wówczas, gdy przesłanka druga nie jest tautologią. Podobnie jak w przypadku przesłanki poprzedniej polega on na tym, że do modelu wprowadza się twierdzenia, które uzasadniają drugą przesłankę. Te twierdzenia mają charakter definicyjny, mogą to być jednak i twierdzenia empiryczne. Powinny one wówczas nadawać się do uzasadnienia drogą odwołania się do dostępnych danych. Jeśli te twierdzenia będą prawdziwe, wówczas prawdziwa będzie i przesłanka druga, a więc, przy założeniu prawdziwości przesłanki pierwszej, poprawnie wyprowadzany z nich wniosek, czyli odpowiedź na pytanie dotyczące jednostki badania. Oczywiście wszystkie twierdzenia, uzasadniające obydwie przesłanki, wchodzi w skład modelu. W ten sposób staje się on przydatny jako kryterium oceny procesów otrzymywania informacji z punktu widzenia ich prawidłowego przebiegu również i w takich przypadkach, kiedy pytanie badacza zasadniczo różni się od pytania kodera.

Twierdzenia empiryczne, o które należy uzupełnić model, powinny nadawać się do empirycznego uzasadnienia. Nie zawsze musi być to jednak uzasadnienie bezpośrednie. Może być ono pośrednie za pomocą innych twierdzeń, które z kolei — bezpośrednio lub nawet pośrednio — nadają się do empirycznego uzasadnienia. Oczywiście obok twierdzeń empirycznych w całym układzie, tworzącym taki rozbudowany model, występować mogą i twierdzenia definicyjne.

Sposób, w jaki można uzupełnić model, szcharakteryzujemy jeszcze inaczej, odwołując się do koncepcji pytań zależnych⁶¹. Jeśli pytanie badacza i pytanie dotyczące rodzaju odpowiedzi respondenta, a więc pytanie kodera, różnią się od siebie, stanowić powinny parę pytań zależnych, tj. takich, na które odpowiedzi

⁶¹ Por. J. Giedymina, *op. cit.*, s. 19 i następne. Por. także T. Pawłowski, *Metodologiczne zagadnienia humanistyki*, Warszawa 1969. Pawłowski wprowadza pewne modyfikacje do określeń Giedymina.

łączy stosunek wynikania. Odpowiedź na pytanie kodera jest przy tym odpowiedzią całkowitą, pośrednią na pytanie badacza, ponieważ — jak stwierdza się w przesłance drugiej — wynika z niej ta ostatnia odpowiedź. Jest to jednak zawsze wynikanie ze względu na jakieś inne twierdzenie względnie twierdzenie o charakterze definicyjnym lub empirycznym. Twierdzenia te należy wyszukać, sformułować i sprawdzić, czy w całości, którą stanowią wraz z już sformułowanymi twierdzeniami z modelu, składają się na ciąg rozumowania, prowadzący do wniosku identycznego z odpowiedzią na pytanie badacza. Wówczas można je włączyć do modelu.

4. Struktura modelu i jej wykorzystanie w ocenie otrzymywania informacji

Jaka będzie struktura modelu zbudowanego w ten sposób? Jakie twierdzenia będą się na niego składać? Zgodnie z tym, o czym była już mowa, wyróżnić trzeba dwa zasadnicze rodzaje tych twierdzeń: stałe i zmienne, uzyskane drogą uzupełnienia rozumowania według przyjętego schematu. Jak wiemy, zmienna część modelu nie jest potrzebna w przypadku pewnych pytań i poszukiwanych informacji. Twierdzenia stałe, zawsze w jednakowej postaci, wchodzą w skład każdego procesu otrzymywania informacji. Należy wyróżnić wśród nich dwie grupy. Jedna związana jest z pierwszą przesłanką w rozumowaniu według przyjętego schematu. Są to twierdzenia opatrzone numerami 1-4 i 6 w liście przedstawionej na s. 373. Jeżeli są one prawdziwe, wówczas prawdziwa jest i ta przesłanka.

Druga grupa to twierdzenie 5, oraz dwa twierdzenia uzupełniające tę listę (por. s. 374-375). Odnoszą się one do klasyfikacji odpowiedzi na pytania kwestionariuszowe i na jednostkowe pytania badacza. Ich prawdziwość jest koniecznym warunkiem tego, aby można było poprawnie sformułować obydwie przesłanki występujące w schemacie. Model złożony tylko z tych twierdzeń jest wystarczający do oceny procesu otrzymywania informacji, gdy dotyczy ona rodzaju odpowiedzi na pytanie kwestionariuszowe. Wówczas klasyfikacja jednostek badania jest identyczna z klasyfikacją tych odpowiedzi, w związku z czym twierdzenie 8 jest

zawsze prawdziwe. Prawdziwa jest także przesłanka druga ze względu na tautologiczny charakter.

Pozostałe twierdzenia związane są tylko z przesłanką drugą w schemacie i zmieniają się w zależności od rodzaju pytania z kwestionariusza i odpowiedzi. Wyróżnić można wśród nich twierdzenia definicyjne i empiryczne. Definicje częściowe występują wtedy, gdy twierdzenie o definicyjnym charakterze stanowi jedyne uzupełnienie modelu. Jeśli definicje występują obok twierdzeń empirycznych, wówczas mogą mieć charakter definicji równościowych.

Różnorodność empirycznych twierdzeń, które mogą się znaleźć w modelach różnych procesów otrzymywania informacji, związanych z różnymi pytaniami, uniemożliwia dokonywanie ich pełnej systematyzacji w niniejszym studium. Aby wyróżnić choćby niektóre ich rodzaje, przytoczymy najpierw przykład, bynajmniej nie należący do skomplikowanych. Przypuśćmy, że w badaniach zadajemy pytanie zamknięte, w którym żąda się ujawnienia pozytywnej bądź negatywnej opinii respondenta w jakiejś sprawie. Jedna poszukiwana informacja związana z tym pytaniem musi dotyczyć kwestii, czy respondent posiada taką opinię, względnie czy potrafi ją po zastanowieniu sformułować. Drugie pytanie badacza, uzależnione zresztą od rodzaju odpowiedzi na pytanie poprzednie, odnosi się do tego, jaka to jest opinia, pozytywna czy negatywna. Dla uproszczenia weźmiemy pod uwagę tylko drugą z poszukiwanych informacji i sytuację, kiedy respondenci wybrali jakiś wariant odpowiedzi istotnej ze względu na tę informację. Założmy także, że respondenci nie mają skryzalizowanych opinii w danej kwestii. Można jednak liczyć na to, że potrafią ją sformułować.

Twierdzenia, o które należy uzupełnić model prawidłowego przebiegu omawianego procesu, mogą być w tym przypadku następujące. 1. Respondent usłyszał pytanie i dotarło ono do jego świadomości. 2. Respondent zrozumiał je tak, jak chce tego badacz. 3. Respondent zaakceptował pytanie, tj. uznał, że może na nie odpowiedzieć zgodnie z zawartym w nim żądaniem. 4. Respondent wykonał operacje myślowe, które umożliwiają uformo-

wanie stanowiska. 5. Respondent aprobuje to stanowisko wyrażone w wybranej odpowiedzi. Wszystkie te twierdzenia wymagają uszczegółowienia i komentarzy, których nie będziemy tutaj przedstawiać ze względu na to, że analogiczne pytania będą szczegółowo omówione w jednym z następnych artykułów⁶². Wymaga natomiast wyraźnego podkreślenia to, że te twierdzenia uległyby znacznej zmianie, gdyby dobierać je dla innych pytań kwestionariuszowych i innej informacji, na przykład pytań projekcyjnych lub hipotetycznych, czy choćby pytań typu egzaminacyjnego, sprawdzających wiedzę respondenta.

Twierdzenia powyższe dotyczą prawidłowego przebiegu procesów psychicznych respondenta, jego reakcji wewnętrznej na niektóre pytania. Jeśli są one prawdziwe, to możemy przyjąć, że odpowiedź zgodna jest z przekonaniem respondenta, do którego doszedł on właściwą drogą. Ale badacz może interesować nie tylko to, co respondenci myśleli w czasie wywiadów, ale i to, jakie jest ich stanowisko w danej sprawie w innych sytuacjach w życiu. Jeśli więc poszukiwana informacja dotyczy nie tylko przekonania respondenta podczas wywiadu, ale i jego poglądu w danej sprawie w różnych sytuacjach, to model powinien obejmować co najmniej jeszcze jedno twierdzenie, nie odnoszące się już do reakcji wewnętrznej respondenta, lecz do stosunku między sytuacją, w której respondent udzielał wywiadu, i sytuacjami życiowymi, w których zajmuje on, lub może zająć stanowisko w danej kwestii. Nie przedstawiając tutaj bardziej precyzyjnych sformułowań należy jedynie zauważyć, że ogólnie rzecz biorąc, mogą to być tylko sytuacje zbliżone pod istotnymi względami do tego, co z punktu widzenia respondenta występowało w wywiadzie. Gdy te sytuacje będą bardzo podobne ze względu na istotne cechy, wówczas można przyjąć, że zachowania i przeżycia respondenta w czasie wywiadu są reprezentatywne dla jego zachowań i przeżyć danego rodzaju gdzie indziej. Używając termi-

⁶² Por. artykuł B. Tuchańskiej, *Analiza modelu wewnętrznej reakcji respondenta na pytanie w wywiadzie kwestionariuszowym*, w niniejszym tomie *Analiz i prób...*

nologii z nauk przyrodniczych można tę myśl wyrazić słowami, że to, co stwierdzono w sytuacji zbliżonej do laboratoryjnej, *in vitro*, występuje także *in vivo*. Oczywiście, przekonanie o występowaniu wspomnianego wyżej podobieństwa powinno być uzasadnione. Trzeba także zauważyć, że może zdarzyć się tak, iż respondent zajmuje różne stanowiska w danej kwestii zależnie od różnych sytuacji, odznaczających się przeciwstawnymi cechami. Jeśli tak jest, to zachowanie i przeżycia respondenta w wywiadzie będą reprezentatywne tylko dla części spośród tych sytuacji.

Kiedy w badaniu można liczyć na to, że opinia respondentów w danej sprawie jest skryształizowana i stała i badacz chce uzyskać informację wyłącznie o niej, twierdzenie lub twierdzenia, dotyczące stosunku sytuacji wywiadu do sytuacji z życia, nie są w modelu potrzebne. Natomiast rozumienie pytania przez respondenta powinno obejmować wówczas i uświadomienie sobie tego, iż badaczowi chodzi właśnie o jego stałą opinię⁶³.

Obok twierdzeń dotyczących wewnętrznej reakcji respondenta oraz stosunku sytuacji wywiadu do sytuacji w życiu wyróżnić należy jeszcze jeden ich rodzaj. Twierdzenia te odnoszą się do zjawisk i ich prawidłowości, a zwłaszcza ich związków występujących poza wywiadem. Na przykład, badacz na podstawie odpowiedzi respondenta na pytanie dotyczące kwestii, na czytanie którego działu w gazecie poświęca najwięcej czasu, ustala, czy interesuje się on polityką. Opiera się wówczas na założeniu, że między ilością czasu poświęconą na czytanie działu w gazecie a zainteresowaniem tym, czemu poświęcony jest dział, występuje związek będący przejawem szerszych prawidłowości. Występowanie tego związku powoduje, że poświęcenie dużej ilości czasu na czytanie działu w gazecie może być uznane za wskaźnik zainteresowania przedmiotem tego działu. Zdanie o występowaniu tego związku, uzasadniające wprowadzenie tego wskaźnika, sta-

⁶³ Przeciwwstawienie „opinia skryształizowana i stała” i „pogląd sformułowany w wywiadzie”, używane w tekście, dotyczy typów krańcowych i w praktyce między tymi krańcami znaleźć można wiele stopni pośrednich. Posiada to oczywiście konsekwencje dla modelu, którym nie będziemy się już zajmowali.

nowi wówczas twierdzenie, o które trzeba uzupełnić model. Jest to twierdzenie odnoszące się do zjawisk poza wywiadem, dotyczy ono występowania ogólnej prawidłowości.

W niektórych przypadkach, kiedy jedno z empirycznych twierdzeń włączonych do modelu jest twierdzeniem powyższego typu, badacz wprowadza także nowe pojęcie. Na przykład, w powyższym przykładzie, za „czytelnika zainteresowanego polityką” można uznać takiego czytelnika, który na czytanie odpowiedniego działu w gazecie poświęca najwięcej czasu, zaś „czytelnikiem interesującym się sportem” można nazwać takiego, który najdłużej czyta wiadomości sportowe i komentarze z tego zakresu itp. Takie twierdzenia o charakterze definicji także powinny wejść do modelu. Warto zauważyć, że nie są to definicje częściowe, lecz równościowe, gdyż cechy, które decydują o zastosowaniu danego pojęcia, są niezależne od zadania pytania w wywiadzie.

Wśród twierdzeń definicyjnych, wchodzących w skład modelu, trzeba więc wyróżnić definicje różnego typu. Wśród twierdzeń empirycznych wyróżniliśmy trzy rodzaje: odnoszące się do respondenta i jego wewnętrznej reakcji na pytanie, dotyczące stosunku sytuacji w wywiadzie do sytuacji innych, do których odnosi się poszukiwana informacja, oraz twierdzenia, w których charakteryzuje się prawidłowości zjawisk poza wywiadem. To, które z tych różnych ze względu na swój rodzaj twierdzeń powinny się znaleźć w modelu, oraz to, jakie będą to twierdzenia, zależy od rodzaju i treści pytania oraz poszukiwanej informacji, o czym już zresztą wspominaliśmy. Nie oznacza to oczywiście, że w modelach wielu pytań nie występują twierdzenia pod wieloma względami analogiczne. Przeciwnie, twierdzenia takie występują i w związku z tym, jak się wydaje, można konstruować modele ogólne, które odnoszą się do grupy pytań. Aby je wykorzystać w ocenie procesu związanego z danym pytaniem, trzeba jednak taki ogólny model uzupełnić, zastępując twierdzenia ogólne bardziej szczegółowymi. Tego rodzaju sytuacja wystąpiła właśnie w badaniach metodologicznych, które będą omawiane w następnych artykułach.

Przy omawianiu struktury modelu prawidłowego procesu

otrzymywania informacji należy zwrócić uwagę także na możliwość zastępowania jednych twierdzeń empirycznych, najczęściej zresztą ich układów, przez inne. Budowanie modelu, jak wskazaaliśmy, polega na uzupełnianiu rozumowania, w którym dany jest jego punkt wyjścia i punkt dojścia, czyli wniosek o nowe twierdzenia. Tego rodzaju postępowanie nie zakłada, że w przypadku jednego rozumowania — a więc i jednego procesu otrzymywania informacji za pomocą danego pytania — model nie może być uzupełniony w różny sposób. Innymi słowami, zadanie budowy modelu może być rozwiązane w sposób poprawny różnie. W praktyce to zróżnicowanie dotyczy, jak się wydaje, głównie twierdzeń o wewnętrznej reakcji respondenta na pytanie. Twierdzenia te nie muszą na przykład dotyczyć tylko szczegółowych procesów psychicznych występujących u respondenta. Mogą odnosić się do jego cech dyspozycyjnych, które umożliwiają zajście tych procesów. Tak więc w modelu może znaleźć się twierdzenie dotyczące zasobu słów używanych przez respondenta, koniecznego do udzielenia w sposób prawidłowy odpowiedzi na dane pytanie, zamiast twierdzenia o tym, że respondent zrozumiał dane słowo lub go nie zrozumiał zgodnie z sensem przyjętym przez badacza. Zastosowanie takich twierdzeń oznaczałoby, że — przynajmniej w danym zakresie — model posiadałby charakter probabilistyczny. Umiejętność bowiem posługiwania się mniejszym czy większym zasobem słów nie przesądza jednoznacznie o tym, że respondent zrozumiał pytanie o określonym stopniu trudności terminologicznej lub że go nie zrozumiał.

Dobór uzupełniających, zmiennych twierdzeń empirycznych jest więc w pewnym zakresie dowolny. Oczywiście ta dowolność jest bardzo ograniczona. W każdym przypadku bowiem twierdzenia te muszą spełniać podstawowy wymóg, aby można było ustalić, czy są one prawdziwe w odniesieniu do przebiegu procesów otrzymywania informacji, które mają być weryfikowane tą drogą. Nie zawsze musi być to przy tym sprawdzenie bezpośrednie, może być i pośrednie za pomocą innych twierdzeń.

Ta ostatnia kwestia stanowi jeden z głównych problemów, które wiążą się z weryfikacją wewnętrzną. Rozwiązanie tej kwes-

tii musi być odmienne w przypadku twierdzeń różnych typów. Twierdzenia dotyczące prawidłowości w występowaniu zjawisk poza wywiadem powinny być wystarczająco uzasadnione wcześniej, przed ich wykorzystaniem w modelu. Są to najczęściej twierdzenia formułowane jako wnioski w badaniach należących do tej dziedziny socjologicznych — lub zbliżonych — poszukiwań, do których należy także problem weryfikowanych badań. W innej nieco sytuacji są twierdzenia głoszące, że między sytuacją wywiadu a określonymi sytuacjami życiowymi, w których respondent powinien zachować się analogicznie jak w wywiadzie, występują istotne podobieństwa. Aby zweryfikować te twierdzenia, badacz musi, z jednej strony, dysponować wiedzą dotyczącą tych ostatnich sytuacji, z drugiej, wiedzą o przebiegu wywiadów. Wiedzę w tym drugim zakresie musi zdobyć sam. Użyteczne mogą być pod tym względem szczegółowe dodatkowe informacje ankietów, między innymi uzyskiwane w ankietach do ankietów lub zawarte w specjalnych adnotacjach robionych w trakcie wywiadu. Obok tego wiele, i to najlepszych informacji tego rodzaju zawiera zapis magnetofonowy wywiadu. Wiedza dotycząca sytuacji poza wywiadem jest wiedzą specjalistyczną należącą do jakiejś dziedziny socjologii.

Twierdzenia odnoszące się do klasyfikacji odpowiedzi respondenta i jednostek badania można sprawdzić drogą analizy dokumentów takich, jak instrukcje i wstępne opracowanie wyników, od razu w zastosowaniu do wszystkich procesów związanych z danym pytaniem i informacją.

Sprawdzenie twierdzeń odnoszących się do zewnętrznej strony przebiegu procesów otrzymywania informacji, a więc związanych z przesłanką pierwszą schematu, wymaga szczegółowych danych o tym, co działo się w wywiadzie i przy kodowaniu. Użyteczne mogą być tutaj także — poza zapisem odpowiedzi z kwestionariusza — specjalne adnotacje ankietów i przede wszystkim zapis magnetofonowy wywiadu. Jeśli zapis ten ma dostarczyć obraz „normalnego” zachowania ankietera, powinien być przygotowany bez jego wiedzy. Prawidłowość zakodowania można sprawdzić drogą zwykłej kontroli, dokonanej przez badacza.

Najwięcej problemów wiąże się ze sprawą danych, dotyczących wewnętrznej reakcji respondenta i twierdzeń o jego procesach psychicznych. Gdyby twierdzenia te odnosiły się tylko do cech dyspozycyjnych respondenta, na podstawie których dałoby się ustalić z określonym prawdopodobieństwem, czy jego reakcja może być prawidłowa, wykorzystać tutaj można byłoby testy psychologiczne, istniejące i specjalnie przygotowane. Byłyby to testy określające przydatność respondenta do udzielania prawidłowej odpowiedzi ze względu na jego możliwości i nastawienia, także wywołane w toku wywiadu. Zastosowanie takich testów umożliwiłoby obiektywizację sposobu postępowania, skróciłoby również czas potrzebny do zdobycia informacji. Jest to jednak w chwili obecnej niemożliwe. W związku z tym jedyna droga, którą można obecnie wykorzystać, polega na wypytывaniu respondenta o to, jak odbierał pytanie, jakie wywołało ono w nim myśli, jak formułował odpowiedź itp. Będzie to więc zastosowanie introspekcji pytaniowej w wywiadzie o wywiadzie. Wywiad ten został szczegółowo omówiony w jednym z artykułów w tym tomie ⁶⁴.

Wyniki weryfikacji wewnętrznej w zestawieniu z wynikami weryfikacji zewnętrznej. Uwagi końcowe

Na podstawie dotychczasowych rozważań łatwo można sformułować wniosek, że zarówno zbudowanie modelu prawidłowego przebiegu procesu otrzymywania informacji, jak i zwłaszcza zastosowania go w ocenie konkretnych procesów jest trudne i bardzo pracochłonne. Należy w związku z tym rozpatrzyć sprawę, jakie wyniki przyniesie weryfikacja wewnętrzna, jeśli ją przeprowadzimy za pomocą takiego modelu.

Przed wszystkim warto zwrócić uwagę na to, że samo zbudowanie

⁶⁴ I. Przybyłowska, *Wywiad o wywiadzie jako metoda otrzymywania informacji o reakcji wewnętrznej respondenta na pytanie kwestionariusza*, w tym tomie *Analiz i prób...* Pojęcie introspekcji pytaniowej wprowadził M. Kreutz, *Metody współczesnej psychologii. Studium krytyczne*, Warszawa 1962, s. 109 - 132.

wanie modelu może przynieść pewne rezultaty. Stanowi on mianowicie środek kontroli poprawności pytań przy układaniu kwestionariusza. Gdy w odniesieniu do jakiegoś pytania w związku z daną poszukiwaną informacją która musi być jasno sformułowana, nie udało się skonstruować takiego spójnego logicznie modelu, wówczas pytanie to ma ewentualnie rację bytu tylko jako próba podjęta w braku innych możliwości. Trudno bowiem przywiązywać większą wagę do informacji, o których badacz nie wie, jaką drogą może je uzyskać — oczywiście jeśli nie ma zewnętrznych pozytywnych sprawdzianów ich wiarygodności.

Sformułowanie modelu w postaci szczegółowych twierdzeń dostarcza dalszych możliwości kontroli. W niektórych wypadkach może być jasne, że te twierdzenia nie mogą być pozytywnie zweryfikowane w przypadku bardzo dużej ilości procesów otrzymywania informacji. Najlepsze ilustracje w tym zakresie stanowią przykłady niezgodnego rozumienia różnych terminów występujących w pytaniach kwestionariuszowych, z jednej strony, przez badacza, a z drugiej, przez respondentów. Nie są to bynajmniej ilustracje jedyne. Wystarczy uzmysłować sobie, jakich umiejętności wymaga się często od respondentów, każąc im zestawiać ze sobą kilkanaście lub więcej możliwych odpowiedzi w kafeterii, niekoniecznie dla nich zrozumiale sformułowanych, w celu dokonania wyboru kilku najważniejszych możliwości.

Tych zastosowań omawianego modelu, aczkolwiek niewątpliwie istotnych, nie można jednak uznać za najważniejsze w studium poświęconym weryfikacji. Jakie możliwości w tym zakresie przedstawia model? Zakładając, że został on dobrze skonstruowany i że badacz dysponuje danymi, które pozwalają zestawiać z nim konkretne procesy otrzymywania informacji, możemy za jego pomocą ocenić każdy proces i uznać go za prawidłowy lub nieprawidłowy ze względu na różnorodne możliwe zakłócenia występujące w jego przebiegu. Gdy przebieg ten jest prawidłowy, otrzymana informacja będzie prawdziwa. Model określa więc warunek wystarczający prawdziwości jednostkowej informacji. Zawarte w nim twierdzenia dotyczą niezbędnych składników tego warunku.

Czy jest to także warunek konieczny prawdziwości tych informacji? Innymi słowy, zakładając pełną poprawność weryfikacji, czy może zdarzyć się tak, że dany proces nie wytrzymuje próby, którą stanowi zestawienie z modelem w zakresie jednego lub większej ilości twierdzeń, otrzymana zaś informacja będzie prawdziwa?

Sytuacja taka jest możliwa. Powiedzmy, że respondent nie zrozumiał pytania, lecz wybrał na chybił trafił którąś odpowiedź, co, jak wiadomo, zdarza się w wywiadach lub zrozumiał je błędnie. Nie znaczy to, że otrzymana informacja jest fałszywa. Dana odpowiedź mogła być przecież odpowiednia dla tego respondenta przypadkowo. Może także zdarzyć się i tak, że w jednym procesie wzajemnego komunikowania, w ramach którego formuje się potrzebna badaczowi informacja, wystąpiły dwa zniekształcenia w przeciwnych kierunkach. W rezultacie końcowa informacja nie jest obciążona błędem. Na przykład ankieter źle zapisał odpowiedź, lecz ponieważ pomylił się i koder, wpisany został symbol klasyfikacyjny tej odpowiedzi, której faktycznie udzielił respondent.

Zgodność z modelem nie stanowi więc warunku koniecznego, aby otrzymana informacja była prawdziwa. Może być ona prawdziwa przypadkowo. Ilość takich przypadków może być nawet relatywnie duża, na przykład w bardzo trudnych pytaniach o dwóch możliwych odpowiedziach kiedy szansa przypadkowego trafienia na właściwą odpowiedź wynosi 50% (przy założeniu, że pytanie jest dostosowane do respondenta). Ilość ta będzie duża także wówczas, gdy sam model nie będzie obejmował wszystkich różnych prawidłowych dróg otrzymywania informacji, które czasem występują. Taka sytuacja świadczy jednak już o błędzie w konstrukcji modelu. Ogólnie rzecz biorąc, ilość informacji prawdziwych przypadkowo nie może być zbyt duża i informacje, otrzymane w rezultacie procesów, w których stwierdzono zakłócenia, należy uznać za podejrzane o to, iż są błędne. Podejrzenie tego rodzaju w tym przypadku jest oparte na znacznie mocniejszych podstawach niż podejrzenie, które można sformułować wówczas,

gdy stwierdzi się na przykład jakąkolwiek omyłkę ankietera, o czym wspominaliśmy poprzednio.

Ze względu na to, czego dotyczą liczbowe dane, które można otrzymywać przy zastosowaniu weryfikacji wewnętrznej (oczywiście przy założeniu, że udało się ją dobrze przeprowadzić), jej rezultaty tylko w części mogą się pokrywać z rezultatem weryfikacji zewnętrznej, w której dla kontroli zastosowano sposób całkowicie pewny do tej samej zbiorowości. Zgodność tych rezultatów polegałaby na tym, iż wszystkie informacje, zaakceptowane w wyniku weryfikacji wewnętrznej, muszą okazać się prawdziwe w rezultacie zastosowania weryfikacji zewnętrznej tego rodzaju. Natomiast może być tak, że pewna ilość informacji nie zaakceptowanych w rezultacie weryfikacji wewnętrznej okaże się prawdziwa, gdy zostaną one zweryfikowane zewnętrznie.

Rezultaty weryfikacji wewnętrznej nie pokrywają się więc w pełni z wynikami weryfikacji zewnętrznej. Niezależnie od tego rezultaty te, gdyby udało się je osiągnąć, należy uznać za bardzo cenne, ponieważ przedstawiają one liczbową ocenę otrzymanych informacji. Niezmiernie ważne jest przy tym to, że ta ocena może dotyczyć i takich pytań, i poszukiwanych informacji, w odniesieniu do których nie można przeprowadzić weryfikacji zewnętrznej przy zastosowaniu sposobu pewnego jako kontroli, np. w stosunku do pytań o opinie. Trzeba jeszcze dodać, że weryfikacja wewnętrzna pozwala także wykryć bezpośrednie źródło zakłóceń w procesie otrzymywania informacji, co ma bardzo duże znaczenie i dla interpretacji danych w tym badaniu i w ogóle dla doskonalenia techniki wywiadu.

Podkreślając wagę rezultatów, które można by osiągnąć drogą weryfikacji wewnętrznej, trzeba jednak od razu wymienić jej ograniczenia, a także problemy wiążące się z jej zastosowaniem.

Jedno z takich ograniczeń łączy się z faktem, iż weryfikację wewnętrzną niezmiernie trudno byłoby zastosować w odniesieniu do przypadków, w których jedną informację otrzymuje się za pomocą kilku pytań. Należałoby wówczas skonstruować kilka modeli, tyle, ile pytań, ponadto zaś zbudować model, który obejmowałby twierdzenia służące jako dodatkowe przesłanki w rozumo-

waniu prowadzącym od informacji szczegółowych do informacji globalnej o jednostce badania. Te szczegółowe informacje mogą dotyczyć wówczas także tylko tego, jak respondent odpowiedział na dane pytanie.

Nie jest to jednak ograniczenie — czy raczej dodatkowy problem do rozwiązania — dużej wagi, zwłaszcza w porównaniu z innymi. Te inne problemy nasuwa sama konstrukcja modelu w jej końcowej fazie, a więc dobieranie twierdzeń empirycznych jako elementów poprawnego rozumowania, oraz sprawdzenie, czy rzeczywiste przebiegi procesów formowania informacji są zgodne z modelem. Trudne wydaje się zwłaszcza uzyskanie wystarczających i wystarczająco pewnych danych, które umożliwią sformułowanie oceny prawidłowości przebiegu wewnętrznej reakcji respondenta na pytanie z kwestionariusza.

Ustalenie, czy lub w jakim zakresie jest to trudność do przezwyciężenia, ewentualnie wymaga prób, które warto, jak się wydaje, podjąć ze względu na ewentualne rezultaty. Trzeba jednak z góry liczyć się z tym, iż weryfikacji wewnętrznej nie będzie można zastosować w masowych badaniach dla oceny ich wyników. Źródłem tego ograniczenia jest nie tylko jej pracochłonność, lecz i to, że połączenie jej z takimi badaniami nie da się realizować w praktyce. Wydaje się, że ten typ weryfikacji wymaga przeprowadzenia specjalnych metodologicznych badań. Mogą to być także rozbudowane badania pilotażowe⁶⁵. Zakres takiej weryfikacji z konieczności musi być niewielki, tj. może ona dotyczyć tylko niewielu pytań. Oparcie się na wynikach takich studiów pozwala z kolei formułować przypuszczenia dotyczące ilości zweryfikowanych pozytywnych lub negatywnych informacji w większych badaniach.

Na zakończenie części artykułu poświęconej weryfikacji zewnętrznej sformułowano szereg uwag ogólnych, dotyczących róż-

⁶⁵ Wiele elementów proponowanego tutaj sposobu postępowania występuje w pilotażach aktualnie prowadzonych przez Zespół Metodologii Badań Socjologicznych Instytutu Filozofii i Socjologii PAN w Łodzi. Por. artykuł K. Lutyńskiej, *Pilotaż pogłębiony. Koncepcja, realizacja i opracowanie materiałów*, w niniejszym tomie *Analiz i prób...*

nych jej rodzajów. Uwagi te wydają się słuszne, jeśli objąć nimi także weryfikację wewnętrzną. Wymaga ona prób i refleksji, aby można było ocenić jej możliwości i maksymalnie je wykorzystać. Celowe jest przy tym, jak się wydaje, równoległe stosowanie w nich analiz weryfikacyjnych różnych typów. Celowe dlatego, że w ten sposób można w pewnym zakresie ocenić ich wartość jako metod badawczego postępowania. Inny cel, którego realizację można wówczas przybliżyć, stanowi opracowanie rozbudowanych procedur weryfikacyjnych, w których wykorzystywano by łącznie analizy różnych typów. Analizy zawierające elementy takiego sposobu postępowania są znane w literaturze, jednak są one dostosowane do specjalnego rodzaju danych, przy czym nie dąży się w nich do sformułowania bardziej ścisłych liczbowych ocen⁶⁶. Być może w przyszłości przy łączeniu różnych analiz uda się uniknąć ograniczeń każdej z nich przy jednoczesnym wykorzystaniu ich walorów i możliwości.

⁶⁶ Por. Aneks E poświęcony weryfikacji w cytowanej pracy Lazarsfelda i Thielensa.

JAN LUTYŃSKI

CELE, PRZEBIEG I WSTĘPNE WYNIKI BADAŃ WERYFIKACYJNYCH WE WŁOCŁAWKU W 1973 R.

Artykuł niniejszy stanowi wprowadzenie do dalszych artykułów w tej części niniejszego tomu *Analiz i prób technik badawczych w socjologii*. Wszystkie te artykuły dotyczą specjalnych badań metodologicznych, prowadzonych w Zakładzie Techniki i Metod Badań Społecznych Instytutu Socjologii Uniwersytetu Łódzkiego. Badania rozpoczęto w końcu 1972 r., materiały empiryczne zebrano we Włocławku w lecie 1973. Badania nie są jeszcze zakończone. Wydaje się jednak, że zdobyte w nich doświadczenia i wnioski, które już można sformułować, są na tyle obfite, że warto je przedstawić w artykułach, których autorami są uczestnicy tych badań¹.

Cel omawianych badań związany jest z przemyśleniami, których wyniki przedstawiono w tym tomie w studium *Analizy weryfikacyjne, ich rodzaje i możliwości*. Cel bezpośredni stanowi wypróbowanie i ocena różnych metod weryfikacji danych zdobytych przy zastosowaniu wywiadu kwestionariuszowego, i, w miarę możliwości, zestawienie ich wyników. Celem dalszym, który nie został jeszcze zrealizowany, ma być opracowanie na podstawie

¹ W fazie terenowej badań nie uczestniczył R. Kietliński, brali w nich natomiast udział: P. Daniłowicz, K. Lutyńska i P. Sztabiński z Zespołu Metodologii Badań Socjologicznych Instytutu Filozofii i Socjologii PAN w Łodzi. Z Instytutu Filozofii i Socjologii uzyskano pomoc finansową na prowadzenie badań w terenie. Większość kosztów badań pokryto z kredytów Uniwersytetu Łódzkiego, przeznaczonych na praktyki studenckie i badania naukowe.