

dr Paweł Ostaszewski

Instytut Wymiaru Sprawiedliwości

Bezpośrednio, telefonicznie czy internetowo? Efekt techniki w badaniach wiktyimizacyjnych.

ABSTRAKT

Badania wiktyimizacyjne stanowią w kryminologii jedno z głównych źródeł rzetelnych danych ilościowych o rozmiarach wiktyimizacji, poziomie lęku przed przestępczością i społecznych ocenach organów ścigania i wymiaru sprawiedliwości. Badania te realizowane mogą być przy wykorzystaniu różnych technik badawczych, które z kolei mogą być obciążone różnego rodzaju błędami.

W artykule omówiono podstawowe rodzaje technik prowadzenia badań wiktyimizacyjnych oraz rodzaje błędów związanych z realizacją badań sondażowych. Opisano szczególne problemy, jakie może nieść za sobą wykorzystanie dwóch, alternatywnych do wywiadu bezpośredniego (PAPI lub CAPI), sposobów prowadzenia badań wiktyimizacyjnych, czyli sondaży telefonicznych (CATI) i kwestionariuszy internetowych (CAWI).

Wprowadzenie

Badania wiktyimizacyjne stanowią jedno z podstawowych źródeł kryminologicznych danych ilościowych. Są one formą sondażu społecznego i jako takie muszą spełniać trzy podstawowe założenia tej metody, a mianowicie wystarczająco dużą próbę do badania, losowy dobór badanych oraz wystandaryzowanie narzędzia i procesu realizacji badania¹. Sondaże wiktyimizacyjne zorientowane są na takich zagadnieniach, jak: „doświadczenia wiktyimizacyjne, lęk przed przestępczością, punitywność, stosunek do organów ścigania i wymiaru sprawiedliwości, postrzeganie dezorganizacji i więzi społecznych w zamieszkiwanej okolicy oraz szeroki zakres cech społeczno-demograficznych”². Dostarczają one wielu unikatowych danych o doświadczeniach i opiniach respondentów będących w szczególnym zainteresowaniu kryminologii. Ociążone są jednak także pewnymi ograniczeniami. Jak większość badań ilościowych w naukach społecznych zakładają bowiem możliwość obiektywnego opisu rzeczywistości społecznej, bez uwzględniania ewentualnego wpływu samego procesu badania na badane obiekty. Często operują też niedookreślonymi i nieprecyzyjnymi definicjami opisywanych zjawisk, z których z kolei mogą wynikać nieprecyzyjne wskaźniki konstruowane dla ich dokładnego zmierzenia. W końcu zaś, co jest przedmiotem niniejszego artykułu, problemów może narażać sama techniczna strona prowadzenia tych badań, czyli sposób zbierania i gromadzenia informacji o analizowanych zjawiskach³. W tym kontekście opisuje się sześć podstawowych rodzajów błędów w badaniach sondażowych:

1) błąd próby, związany z tym że wylosowane do badania jednostki nigdy nie stanowią idealnej reprezentacji populacji,

¹ Por. np. F. Sztabiński, T. Żmijewska-Jędrzejczyk, *Mixed Mode Survey Design: Problem Efektu Techniki*, *Mixed Mode Survey Design: problem efektu techniki*. “Przegląd Socjologiczny” nr 1/2012, s. 32.

² P. Ostaszewski, *Lęk przed przestępczością*, Warszawa 2015, s. 89.

³ Por. Tamże, s. 89-92.

2) błąd pokrycia, polegający na tym, że operat losowania (lista jednostek, z których losuje się badanych) może być niekompletny lub może zawierać powtórzenia,

3) błąd braku odpowiedzi, czyli fakt, że część wylosowanych osób pozostaje nieosiągalna, odmawia udziału w badaniu lub nie odpowiada na konkretne pytania,

4) błąd pomiaru, na który składa się m.in. nieprzewidziany wpływ na uzyskane wyniki sposobu sformułowania i formy zadawania pytań oraz osoby ankietera (tzw. wpływ ankieterski) czy respondenta,

5) błąd specyfikacji, czy inaczej niska trafność konstrukcyjna pytań, występująca gdy dany wskaźnik nie mierzy rzeczywiście tego, co ma w założeniu mierzyć,

6) błąd opracowania, a więc wszystkie nieprawidłowości mogące ujawnić się na etapie kodowania, wprowadzania i edytowania danych oraz uzupełniania braków danych⁴.

Techniki prowadzenia sondaży wiktyimizacyjnych można podzielić na trzy podstawowe rodzaje, zakładające: 1) samodzielne udzielanie odpowiedzi przez ankietowanego, 2) bezpośredni kontakt ankietera z respondentem oraz 3) pośredni kontakt ankietera z respondentem. W pierwszym z tych rodzajów wymienia się takie techniki, jak ankiet papierowa (*Paper Self Administered Questionnaires – PSAQ*), pocztowa, audytoryjna i wykorzystująca komputer techniką CASI (*Computer Assisted Self Interviewing*). Za jego specyficzną odmianę można też uznać sondaż internetowy (*Computer Aided Web Interviewing – CAWI*). W drugim, zależnie od formy rejestracji odpowiedzi respondenta, mamy do czynienia z techniką PAPI (*Paper and Pen Interviewing*), w której ankieter wypełnia papierowy kwestionariusz i techniką CAPI (*Computer Assisted Personal Interviewing*), czyli tak zwanym wywiadem bezpośrednim przy użyciu komputera. Podstawową techniką zakładającą pośredni kontakt ankietera z ankietowanym jest sondaż telefoniczny (*Computer Assisted Telephone Interviewing – CATI*)⁵.

W literaturze przedmiotu wyróżnia się trzy podstawowe problemy metodologiczne, które mogą występować przy porównaniu rozpatrywanych sposobów prowadzenia sondaży wiktyimizacyjnych. Są to mianowicie różnice w: 1) poziomie realizacji założonej próby (*response rate*), 2) strukturach zrealizowanych prób, czyli społeczno-demograficznym profilu badanych oraz 3) uzyskanych wynikach, a szczególnie we wskaźnikach wiktyimizacji. Różnice takie definiuje się jako specyficzny błąd pomiaru i określa jako tak zwany efekt lub wpływ techniki (*mode effect*). Błąd ten warto poddać bardziej szczegółowej analizie rozpatrując różnice w wynikach uzyskanych przy wykorzystaniu dwóch technik alternatywnych (CATI i CAWI) w porównaniu do sposobu prowadzenia sondaży wiktyimizacyjnych, uznawanego za podstawowy i preferowany, czyli badania bezpośredniego (PAPI, CAPI lub inaczej – *face-to-face*). Należy jednak od razu zauważyć, że o ile wpływ techniki CATI został poddany licznym analizom, to analogiczny wpływ sondażu internetowego, czyli metody stosowanej nieporównywalnie rzadziej, poddany został wielokrotnie mniejszej liczbie badań. Biorąc jednak pod uwagę stale obniżający się poziom odsetka realizacji (*response rate*) zarówno w kwestionariuszowych badaniach bezpośrednich, jak i telefonicznych, spadający odsetek gospodarstw domowych z telefonem stacjonarnym i rosnący odsetek Polaków korzystających z Internetu należy może zacząć patrzeć na sondaże internetowe, jak na przyszłość badań wiktyimizacyjnych. Szczególnie, że można już mówić o coraz większych grupach respondentów bardzo trudno dostępnych lub wręcz niedostępnych tradycyjnymi metodami badawczymi⁶.

⁴ F. Sztabiński, T. Żmijewska-Jędrzejczyk, *Mixed Mode Survey Design...*, s. 32-33.

⁵ Więcej na ten temat – por. np. P. Ostaszewski, *Lęk przed Przestępczością...*, s. 93-95.

⁶ F. Sztabiński, T. Żmijewska-Jędrzejczyk, *Mixed Mode Survey Design...*, s. 34.

1. Efekt techniki wywiadu telefonicznego (CATI)

Pierwsze badania wiktyimizacyjne, wprowadzone w latach 60-tych ubiegłego wieku realizowane były w formie wywiadu bezpośredniego, przy wykorzystaniu papierowego kwestionariusza. Nadal technika ta uchodzi za wzorcową dla tego typu badań. Wraz z rozwojem techniki komputerowej coraz częściej kwestionariusz papierowy zastępowany jest specjalnym programem komputerowym służącym do zapisywania odpowiedzi (technika CAPI). Oferuje on lepsze mechanizmy kontroli pracy ankietera, redukuje liczbę możliwych błędów wprowadzania, ale w pewnym stopniu „dehumanizuje” też kontakt respondenta z ankieterem (ten ostatni może bardziej skupiać się na obsłudze urzędnika niż na rozmowie z badanym).

W latach 80-tych i 90-tych dużą popularność (w USA i Europie Zachodniej) zyskały telefoniczne badania marketingowe i badania opinii publicznej, gdyż przewyższały sondaże prowadzone bezpośrednio pod względem szybkości, wysokości kosztów realizacji i możliwości kontroli pracy ankieterów. Jednak już pierwsze próby wykorzystania tej techniki ukazały znaczące różnice w uzyskiwanych w nich wynikach, związanych np. z niemożliwością dotarcia do pewnej grupy respondentów, którzy telefonu nie mają. Problem ten staje się szczególnie wyraźny, gdy takie niedostępne jednostki istotnie różnią się od pozostałych.

Następne lata przynoszą obserwacje kolejnych problemów z CATI. Metoda ta staje się mniej efektywna – systematycznie spada odsetek realizacji (*response rate*). Pojawiła się pewna grupa gospodarstw domowych posiadających więcej niż jedną linię telefoniczną, co prowadziło do tak zwanego nadmiernego pokrycia (*overcoverage*). Wzrost liczby takich urzędów, jak automatyczne sekretarki, poczty głosowe, rozpoznawanie głosu (*caller ID/Call screening*) utrudnia kontakt z respondentem i podnosi koszty badania, przez wzrost nieudanych prób dodzwonienia się. Zmienia się także stosunek badanych do takich badań – respondenci odbierają coraz więcej niechcianych telefonów i uznają je za coraz bardziej uciążliwe. Badania opinii publicznych zaczynają być postrzegane podobnie jak telemarketing, szczególnie gdy telemarketerzy podszywają się pod przeprowadzających badania społeczne. Potrzebne są coraz lepsze szkolenia ankieterów, także z technik perswazji (odwołanie do wartości uznanych społecznie, do autorytetów, przypodobanie się respondentowi, uwypuklenie jego wyjątkowości), by badany nie odmówił udziału w badaniu. Głównym problemem badania zaczyna być utrzymanie kontaktu z respondentem i przekonanie go do udziału. Ankieter musi mieć przygotowany zestaw argumentów do użycia w razie braku chęci kontynuowania rozmowy przez badanego, musi umieć słuchać i reagować na to, co mówi respondent, powinien unikać proszenia o uczestnictwo, czy wyrażenie zgody – respondentowi łatwiej bowiem wtedy odmówić. Ważne jest stałe utrzymywanie pozytywnej atmosfery badania, m.in. przez komunikaty typu: „dziękuję, bardzo mi Pan/Pani pomógł/pomogła”, „może jeszcze chwilę się Pan/Pani zastanowi, może jednak zdarzyła się taka sytuacja...”⁷.

Ostatnią znaczącą zmianą na tym polu jest w końcu ogromny wzrost znaczenia telefonów komórkowych i jednoczesny postępujący spadek znaczenia telefonów stacjonarnych. Można bowiem obecnie wyodrębnić trzy grupy osób: posiadające tylko telefon komórkowy (np. osoby młode), posiadające i telefon komórkowy i stacjonarny (osoby trochę starsze, mające samodzielne gospodarstwo domowe), posiadające tylko telefon stacjonarny (coraz mniejsza grupa osób

⁷ E. D. de Leeuw, J. M. Lepkowski, S.-W. Kim, *Have Telephone Surveys a future in the 21-th century?*, International Conference on Improving Surveys, Copenhagen 2002.

najstarszych). Standardowo, badania CATI realizowane były wyłącznie na próbie telefonów stacjonarnych, które można przypisać do konkretnych osób (gospodarstw domowych) i dokładnie zlokalizować terytorialnie. Dzwonienie na telefony komórkowe podnosi koszty badania i jest związane z niemożnością wstępnego terytorialnego zidentyfikowania respondenta. Wprowadzać może też problem nadmiernego pokrycia.

Ciekawy i szczególnie istotny dla badaczy wiktylizacji jest, w kontekście tematu niniejszego artykułu, raport Stevena Hope'a omawiający przeprowadzoną w 2005 r. próbę zmiany metodologii prowadzenia Szkockiego Badania Przystępczości i Wiktylizacji – z wywiadów bezpośrednich na CATI⁸. W Szkocji od roku 1993 robiono cykliczny sondaż wiktylizacyjny na 5-tysięcznej próbie respondentów. Dla umożliwienia prowadzenia analiz bardziej dokładnych uwarunkowań wiktylizacji i jej regionalnych różnic postanowiono zwiększyć tę próbę do 27 tysięcy, co jednak wymagało przejścia na technikę CATI. Jakkolwiek nie spodziewano się istotnych różnic we współczynnikach wiktylizacyjnych uzyskanych z obu metod, dla lepszej kontroli tego procesu obok próby telefonicznej skonstruowano dodatkowo około 3-tysięczną próbę badanych metodą *face-to-face*.

Porównanie procedury doboru i badania obu prób ujawniło m.in.: istotnie niższy odsetek realizacji badania CATI, dużo krótszy czas przeprowadzania wywiadu telefonicznego (ok. 20 minut w porównaniu do 30-40 w wywiadach bezpośrednich), słabszą kontrolę doboru próby w badaniu CATI (bazując na danych adresowych można lepiej kontrolować regiony geograficzne niż przy bazowaniu na numerach telefonów) oraz większe problemy z uzyskaniem odpowiedzi na kwestie wrażliwe w sondażu telefonicznym. Z drugiej strony jeden ankieter może wykonać nieporównywalnie więcej wywiadów CATI niż wywiadów bezpośrednich. W praktyce, przy konstruowaniu dużych prób potrzebnych w badaniu wiktylizacyjnym może nie być możliwości zebrania odpowiedniej liczby ankietów do badania *face-to-face*. CATI zmniejsza również możliwość wystąpienia ewentualnych błędów ankietów przez automatyzację procesu doboru pytań, automatyczne stosowanie filtrów, itp.

Badacze zauważyli również, że osoby niewiktylizowane, gdy uznają, że tego rodzaju wywiad ich nie dotyczy, łatwiej zrezygnują z kontaktu telefonicznego niż z bezpośredniego (tym samym zaburzając rozkład odpowiedzi). Badanie CATI stosuje zwykle metodę losowego doboru numeru telefonu – możliwe jest więc dodzwonienie się zarówno na prywatne numery, jak i np. firmowe lub numery stale przypisane, np. połączeniu komputerowemu, faxowi, itp., co może istotnie fałszować próbę. W badaniu CATI inaczej niż w przypadku wywiadu bezpośredniego oblicza się odsetek realizacji. Przyjmuje się zwykle, że gdy wielokrotnie (10 i więcej razy) nie udało się dodzwonić pod określony numer, to uznaje się go za niewybieralny (np. jest to linia na stałe podłączona do modemu komputera, może być to numer budki telefonicznej, numer faxu, itd.) i nie jest włączany do obliczeń *response rate*. W przypadku wywiadów *face-to-face* większość nieudanych – określoną liczbę razy – prób kontaktu uznaje się za brak kontaktu i do obliczeń odsetka realizacji się włącza (poza przypadkami ewidentnie błędnych lub nie-prywatnych adresów). Warto także zasygnalizować, że osoby nieposiadające telefonu lub posiadające jedynie telefon komórkowy mogą znacząco różnić się od reszty społeczeństwa (częściej są młodzi, bezrobotni, uzyskują niższy dochód, mieszkają w gorszych okolicach, w wynajętych mieszkaniach i samotnie), dlatego niedotarcie do nich podczas badania CATI może wywołać skrzywienie wyników. Wyniki

⁸ S. Hope, *Scottish Crime and Victimization Survey: Calibration Exercise: A Comparison of Survey Methodologies*, MORI Scotland 2005, artykuł dostępny na stronie internetowej: <http://www.scotland.gov.uk/Resource/Doc/47121/0020932.pdf> (dostęp październik 2015 r.).

szkockiego eksperymentu pokazują, że technika wywiadu bezpośredniego nadreprezentuje grupę starszych osób i niedoreprezentuje grupę osób młodszych (młodszych dorosłych). W badaniu telefonicznym zauważono odmienną zależność – nadreprezentację kobiet w wieku 35-64 i niedoreprezentację mężczyzn w wieku 20-24, 45-54 oraz osób z obu płci powyżej 65 lat. Zauważono także różnice w samych odpowiedziach. Przykładowo ankietowani przez telefon dwa razy częściej za poważny problem Szkocji uznali standard transportu publicznego (41% do 20%). Istotnie statystycznie różnice odnotowano także w przypadku lęku przed przestępczością – był on niższy dla badanych telefonicznie, ale jednocześnie istotnie częściej respondenci ci obawiali się bycia ofiarą konkretnych typów przestępstw i istotnie częściej podawali, że w okolicy ich zamieszkania dochodziło to konkretnych przestępstw niż badani *face-to-face*. Proponowanym wyjaśnieniem jest to, że badani telefonicznie mogą bardziej jednoznacznie rozróżniać poczucie bezpieczeństwa w swojej okolicy z ogólnymi obawami i ocenami zagrożenia, choć nie znalazło to potwierdzenia w treści pytań – w obu przypadkach zadano identyczne pytanie. Ankietowani przez telefon częściej podawali, że padli w ostatnim roku ofiarą przestępstwa i że zdarzyło się to większą liczbą razy. Różnice te były istotne statystycznie dla większości rozpatrywanych czynów. Ogólny wskaźnik wiktymizacji dla przestępstw przeciwko osobie był wśród tych respondentów wyższy o 60%.

Różnic tych nie udało się wyjaśnić odmiennością próby respondentów badanych telefonicznie i bezpośrednio. Przykładowo osoby badane techniką *face-to-face*, które podały, że nie mają telefonu stacjonarnego (a więc nie znalazłyby się w próbie telefonicznej) nie różniły się istotnie od pozostałych badanych bezpośrednio. Zadowolających efektów nie przyniosły również dwie próby ważenia wyników. Klasyczne ważenie oparte na wyrównaniu niedoreprezentacji i nadreprezentacji określonych grup socjo-demograficznych nie doprowadziło do większego zrównoważenia wyników. Z kolei ważenie oparte na wyliczonych różnicach w liczbie odmów udziału w badaniu poszczególnych respondentów zamiast minimalizować – zwiększało dysproporcje w strukturze socjo-demograficznej i nie wyrównywało wyników wiktymizacji. W podsumowaniu autorzy stwierdzają, że przeprowadzony sondaż telefoniczny dał tak odmienne wyniki, że nie mogą one wynikać jedynie z błędu próby, a raczej z błędu całego narzędzia (wyrażna tendencja do odmawiania udziału w badaniu przez osoby, które nie były ofiarami przestępstw). Nie znaleźli tym samym wystarczających dowodów na to, by telefoniczne badanie wiktymizacyjne (CATI) umożliwiało dokładne mierzenie wiktymizacji⁹.

Ciekawym uzupełnieniem do powyższych wyników Szkockiego Badania Przestępczości mogą być analizy Hubble'a i Wilder'a prowadzone na podstawie amerykańskiego narodowego sondażu wiktymizacyjnego (National Crime Victimization Survey). Autorzy ci wyjaśniali zauważone istotnie wyższe wyniki wiktymizacji dla badania CATI lepszą kontrolą ankierów – ci muszą przeprowadzić cały wywiad, nie mogą go skrócić przez wpisanie nieprawdziwych odpowiedzi o braku wiktymizacji (gdy respondent nie był wiktymizowany znacząco skraca to cały wywiad, a wynagrodzenie ankiera jest zwykle uzależnione od liczby wszystkich przeprowadzonych wywiadów – bez względu na ich długość)¹⁰.

Holbrook, Green i Krosnick podają dodatkowo, że w badaniu CATI częściej, w porównaniu do wywiadów bezpośrednich, występuje: brak zaangażowania respondenta w badanie, słabsza

⁹ Tamże.

¹⁰ D. L. Hubble, B. E. Wilder, *Preliminary results from the National Crime Survey (NCS) CATI experiment*, U.S. Bureau of the Census 1988, (za:) J. P. Lynch, *Secondary analysis of International Crime Survey Data*, (w:) *Understanding crime: Experiences of crime and crime control*, A. Alvazzi del Frate, U. Zvekic, J. M. van Dijk (red.), UNICRI publication No 49, Rome 1993.

współpraca z ankierem, skłonność do udzielania odpowiedzi „nie mam zdania”, nieuszczerławianie udzielanych odpowiedzi, przytakiwanie i akceptowanie proponowanych odpowiedzi, wyrażanie niechęci wobec długości badania (choć same badania są krótsze), podejrzliwość dotycząca procesu badania, chęć pokazania się w społecznie pożądanym świetle¹¹.

Weeks, Kulka i Lessler zwracają także uwagę, że ogólnie licząc przeprowadzenie wywiadu bezpośredniego jest dwukrotnie droższe niż wywiadu CATI, jednak porównanie takie nie uwzględnia dużo wyższych kosztów początkowych tej drugiej techniki, a mianowicie wyposażenia *call centre*¹².

W technice wywiadu bezpośredniego poszczególni ankierzy pracują zwykle na określonym terenie, co może przekładać efekt ankierski na wyniki terytorialne. Przy CATI nie powinno być takiego niebezpieczeństwa – ankierzy są przydzielani losowo do respondentów. Czasem jednak ośrodki badawcze mogą przydzielać swoich najlepszych ankierów do potencjalnie najtrudniejszych respondentów, albo ankierów pod jakimiś względami uznanych za najwłaściwszych dla danej grupy respondentów – np. posługujących się tym samym dialektem, mających określoną wiedzę o obszarze objętym badaniem, itp. Wtedy też efekt ankierski może być silniejszy. Przykładowo, Lipps zaobserwował, że wewnętrzna korelacja między wynikami skal zastosowanych w badaniu i odsetkiem wywiadów niezrealizowanych (*non-response*) uzyskanych przez tych samych ankierów jest wyższa w technice *face-to-face* niż w CATI, a więc efekt ankierski choć ogólnie słaby, to bardziej znaczący jest w badaniu bezpośrednim. Z drugiej strony, ponieważ tyle samo wywiadów CATI jest w stanie przeprowadzić o wiele mniejsza liczba ankierów, to potencjalnie negatywny wpływ efektu ankierskiego na wyniki może być jednak tam silniejszy (jeden „zły” ankier może „zepsuć” większą liczbę wywiadów)¹³.

Warto w tym miejscu przytoczyć także kilka innych ustaleń dotyczących stosowania techniki CATI. Farrell podaje, że ankierzy CATI deklarują większą satysfakcję z przeprowadzania wywiadów, które oferują im większą swobodę i więcej możliwości działań, które nie ograniczają się jedynie do podania treści pytania i instrukcji, ale które np. oferują możliwość swobodniejszego wybierania pytania, na które odpowiada respondent – zgodnie z naturalnym przebiegiem rozmowy (inne badania pokazują zaś, że brak satysfakcji ankiera z pracy i brak jego zaangażowania znacząco zwiększa poziom odmów). CATI umożliwia również lepszą niż kwestionariusz papierowy kontrolę spójności odpowiedzi (komputer może sygnalizować odpowiedzi niekonsekwentne), co jest przydatne np. w badaniach panelowych, w których wywiady przeprowadzane są z tymi samymi badanymi. Można wtedy na bieżąco śledzić stopień zmiany odpowiedzi i np. dopytać, czy rzeczywiście nastąpiła tak znaczna zmiana – gdy taką deklaruje respondent. W CATI treść samych pytań może być na bieżąco dostosowywana do poprzednio udzielonych odpowiedzi – zgodnie z algorytmem, a nie zgodnie z intuicją ankiera – co zmniejsza wpływ ankiera w sytuacji, gdy powinien on dostosować treść pytania. W CATI można również stworzyć algorytm „dopytywania”. Gdy ankier zaznacza odpowiedź „nie wiem”, może pojawić się dodatkowe pytanie, np. „proszę zastanowić się jeszcze chwilę i spróbować podać najtrafniejsze przypuszczenie”. W wywiadzie telefonicznym program może także losowo dobierać kolejność

¹¹ A. L. Holbrook, M. C. Green, J. A. Krosnick, *Telephone versus Face-to-Face Interviewing of National Probability Samples with Long Questionnaires: Comparisons of Respondent Satisficing and Social Desirability Response Bias*, „Public Opinion Quarterly” 2003, nr 67, s. 79-125.

¹² M. F. Weeks, R. A. Kulka, J. T. Lessler, R. W. Whitmore, *Personal versus telephone surveys for collecting household health data at the local level*, „American Journal of Public Health” 1983, nr 73, s. 1389-1394.

¹³ O. Lipps, *Interviewer and Respondent Survey Quality Effects in a CATI Panel*, „Bulletin de méthodologie sociologique” 2007, nr 95, artykuł dostępny na stronie internetowej <http://bms.revues.org/392> (dostęp październik 2015 r.).

czytanych przez ankietera propozycji odpowiedzi, przez co minimalizuje efekt pierwszeństwa i świeżości. W wywiadzie tym jest również łatwiej wrócić do zadanego wcześniej pytania, w celu poprawienia odpowiedzi na nie, np. gdy respondent coś sobie przypomniał – program wskaże wtedy też inne pytania, które trzeba zadać raz jeszcze. CATI umożliwia przywołanie okna pomocy, czy szczegółowych instrukcji, np. gdy ankieter ma jakieś problemy lub wątpliwości co do kodowania. Podobne dodatkowe informacje, z powodu braku miejsca, raczej nie znajdą się na kwestionariuszu papierowym. Ankieter jest wtedy zdany wyłącznie na siebie i swoją pamięć, co może prowadzić do błędów. CATI umożliwia w końcu automatyczny wybór respondenta. Minimalizuje to ewentualne błędy ankieterów, np. przeprowadzenie wywiadu pod złym adresem. Z drugiej strony ankieterzy odbierają metodę CATI jako bardziej dla nich stresującą i frustrującą, zauważono również więcej spóźnień i zwolnień chorobowych w porównaniu do wywiadów *face-to-face*¹⁴.

Interesujących, praktycznych wskazówek dla prowadzących badania CATI dostarcza również badanie Derr, Mitchell, Brannon, Smiciklas-Wright, Dixon i Shannon. Zakładało ono telefoniczne zbieranie informacji o dziennej diecie (ilości i rodzaju zjedzonego pożywienia przez dorosłych i dzieci chorych na hypercholesterolemię) od ustalonej wcześniej próby osób. Średnio, każdorazowe dodzwonienie się do respondenta wymagało od ankieterów około 3 prób, pomimo że lista numerów była wcześniej ustalona, a badani byli poinformowani o tym, że ankieterzy będą dzwonić (jednak bez informacji dokładnie kiedy). Najlepszym dniem na robienie takich wywiadów okazał się poniedziałek – wtedy to osiągnęto najwięcej zrealizowanych wywiadów, a dodzwonienie się wymagało najmniejszej liczby prób. W przypadku dorosłych dobrym dniem była także niedziela, ale już nie w przypadku dzieci. Najgorsze dni, to piątek i sobota w przypadku dzieci oraz czwartek w przypadku dorosłych. Najlepszą porą dnia na realizowanie wywiadów jest wieczór (ok. 60% zrealizowanych wywiadów). Łatwiej było zrealizować wywiady z dorosłymi niż z dziećmi. Wywiady z dziećmi trwały również dłużej i dłuższe były też okresy oczekiwania na połączenie. W związku z tym koszty zrealizowania badania wśród dzieci były prawie o połowę wyższe niż wśród dorosłych¹⁵.

Z kolei Crippa, de Lima Osório, Del-Ben, Filho, da Silva Freitag i Loureiro wykazali całkowity brak istotnych różnic w wynikach psychologicznego badania diagnostycznego przeprowadzonego metodą CATI i następnie powtórzonego, po okresie od 1 do 3 miesięcy, dla tych samych respondentów techniką *face-to-face*. Może to więc sugerować, że ewentualne różnice w wynikach uzyskiwanych obiema metodami nie są związane z samą procedurą wywiadu, a z procedurą doboru badanych lub rzetelnością narzędzi¹⁶.

Do podobnych wniosków doszła także Sherpenzeel po porównaniu wpływu techniki wywiadu bezpośredniego (CAPI) i telefonicznego (CATI) na wyniki panelowych badań sytuacji ekonomicznej gospodarstw domowych. Zastosowała ona podział próby metodą *split ballot* – połowa badanych jedną metodą, druga połowa drugą i po miesiącu dotychczas zbadani

¹⁴ E. Farrell, *Review of Computer Assisted Telephone Interviewing (CATI) Procedures in Overseas Statistical Agencies*, Methodology Division of Australian Bureau of Statistics 2000, artykuł dostępny na stronie internetowej

<http://www.nss.gov.au/nss/home.NSF/75427d7291fa0145ca2571340022a2ad/b75da801e21308dcca2571ab002470a7?OpenDocument> (dostęp październik 2015 r.).

¹⁵ J. A. Derr, D. C. Mitchell, D. Brannon, H. Smiciklas-Wright, L. B. Dixon, B. M. Shannon, *Time and cost analysis of a computer-assisted telephone interview system to collect dietary recalls*, "American Journal of Epidemiology" 1992, nr 136, s. 1386-1392.

¹⁶ J. A. Crippa, F. de Lima Osório, C. M. Del-Ben, A. S. Filho, M. C. da Silva Freitas, S. R. Loureiro, *Comparability Between Telephone and Face-to-Face Structured Clinical Interview for DSM-IV in Assessing Social Anxiety Disorder*, "Perspectives in Psychiatric Care" 2008, nr. 44, s. 241-247.

bezpośrednio byli ponownie badani tym razem przez telefon, a zbadani po raz pierwszy telefonicznie – teraz bezpośrednio. Generalnie nie zaobserwowano istotnych różnic w wynikach ze względu na stosowane metody, pod warunkiem formułowania tak samo brzmiących pytań. Badanie CATI pozwala oszczędzić czas i pieniądze (średni czas przeprowadzenia samego wywiadu był taki sam, ale oczywiście CAPI wymaga więcej czasu na dojazd, itd.). Trzeba podkreślić, że było to badanie panelowe, w którym te same osoby są badane rokrocznie, a ich adresy i numery telefonów są znane. Nie ma różnic w doborze badanych, a jedynie w sposobie ich pytania, dlatego też nie wystąpił tu omawiany wcześniej wpływ techniki na dobór próby (np. odsetki realizacji dla obu technik były bardzo podobne, list zapowiedni był wysłany na tydzień przed badaniem do wszystkich, z którymi zamierzano się skontaktować, zarówno bezpośrednio jak i telefonicznie). Zauważono jednak nieznaczne różnice w uzyskanych wynikach sugerujące większą tendencję do udzielania społecznie aprobowanych odpowiedzi w wywiadzie bezpośrednim. Przykładowo, na pytanie typu: „Proszę ocenić swój stan zdrowia. Czy jest on bardzo dobry, dobry...” wyższy był odsetek odpowiedzi „bardzo dobry” w badaniu CAPI. Co ciekawe, różnica ta znika przy połączeniu odpowiedzi „bardzo dobry” i „dobry”, czyli to społeczne oczekiwanie powodowało jedynie niewielką zmianę – z odpowiedzi „dobry” na „bardzo dobry”. Podobny, wyższy dla wywiadu bezpośredniego, wpływ postrzeganych przez badanego społecznych oczekiwań, co do udzielenia określonej odpowiedzi zaobserwowano jedynie w przypadku pytania typu: „czy zgadza się Pan/Pani, że ma kontrolę nad własnym życiem”. Przy tak sformułowanym pytaniu odpowiedzi „zgadzam się” częściej padały w badaniu CAPI niż w CATI. Jednak już w przypadku tego samego pytania, ale sformułowanego w formie: „jak Pan/Pani ocenia kontrolę nad własnym życiem...” żadne różnice nie występują. Nie potwierdzono również hipotezy, że w badaniu CATI, jako wymagającym szybszej odpowiedzi (zastanawianie się przy rozmowie telefonicznej jest mniej społecznie pożądane) respondenci będą podawać mniejszą liczbę zdarzeń, które musieliby sobie przypominać, np. podawana średnia liczba wizyt u lekarza była taka sama w obu badaniach¹⁷.

Z kolei następną analizą opartą na wynikach amerykańskiego narodowego sondażu wiktyimizacyjnego (*National Crime Victimization Survey* – NCVS) wykazała, że wszystkie różnice w odpowiedziach na konkretne pytania osób badanych metodą wywiadu bezpośredniego i wywiadu telefonicznego wynikają nie tyle z efektu metody, ile z różnic w poziomie zmęczenia respondentów. Zaobserwowane rozbieżności znikają bowiem w przypadku kontroli długości prowadzenia badania¹⁸.

Dirk Heerwegh i Geert Loosveldt potwierdzili zaś, że w porównaniu do ankiety pocztowej i internetowej to sondaż telefoniczny wprowadza skrzywienie wyników związane z silniejszą tendencją do udzielania odpowiedzi społecznie aprobowanych. CATI generuje przykładowo istotnie więcej pozytywnych ocen policji i różnice te pozostają istotne także w przypadku kontroli struktury społeczno-demograficznej próby respondentów¹⁹.

¹⁷ A. Sherpenzeel, *Mode effects in panel surveys: A comparison of CAPI and CATI*, Bundesamt für Statistik, Schweizer Haushalt-Panel 2001, artykuł dostępny na stronie internetowej <http://www.bfs.admin.ch/bfs/portal/en/index/themen/00/07/blank/02.html?publicationID=1521> (dostęp październik 2015 r.).

¹⁸ G. L. Couzens, M. Berzofsky, Ch. Krebs, *Analyzing Potential Mode Effects in the National Crime Victimization Survey*, Joint Statistical Meetings, JSM 2014 - Survey Research Methods Section, 2-7 Aug 2014, Boston, MA., artykuł dostępny na stronie: https://www.amstat.org/sections/srms/Proceedings/y2014/files/312744_89876.pdf (dostęp październik 2015 r.), s. 2921.

¹⁹ D. Heerwegh, G. Loosveldt, *Assessing Mode Effect in a National Crime Victimization Survey using Structural Equation Models: Social Desirability Bias and Acquiescence*, „Journal of Official Statistics” 2011, nr 27, s. 61.

2. Efekt techniki wywiadu internetowego (CAWI)

Jaki piszą Sztabiński i Żmijewska-Jędrzejczyk, „techniki ankiety ze wspomaganie komputerowym, a więc przede wszystkim CAWI łączą [...] wszystkie zalety technik ankiety i technik wspomaganych komputerowo. Oznacza to szybkość realizacji badania, niskie koszty przygotowania narzędzia i realizacji oraz możliwość wielokrotnego monitorowania. Podstawowym ograniczeniem jest jednak relatywnie niski i bardzo zróżnicowany geograficznie poziom wyposażenia gospodarstw domowych w Internet (tzw. poziom penetracji, który w Polsce wynosi ok. 50%) i związany z tym błąd pokrycia...”²⁰. Jak już wyżej sygnalizowano, nie prowadzono dotychczas wielu analiz efektu techniki badań internetowych na wyniki sondaży wiktyimizacyjnych. Należy jednak podkreślić, że przeprowadzone analizy sugerują, iż między badaniami internetowymi i tradycyjnymi pocztowymi występują w tym kontekście niewielkie i zwykle całkiem nieistotne różnice²¹.

Najnowsza analiza tego zagadnienia została przeprowadzona przez Nathalie Guzy i Heinz’a Leitgöb’a²². Zakładała ona porównanie wskaźników wiktyimizacyjnych uzyskanych w pilotażowym badaniu ICVS-2 za pomocą dwóch rodzajów kwestionariuszy – telefonicznego (CATI) i internetowego (CAWI). Badacze zakładali, że samodzielnie wypełniany kwestionariusz internetowy będzie wymagał więcej czasu na wypełnienie. Da on więc lepszą możliwość zapoznania się z treścią pytań, a pytania szczególnie długie i skomplikowane będą rozumiane bardziej prawidłowo, przez co odpowiedzi na nie będą bardziej rzetelne i trafne²³. Z kolei sondaż ankieterski, którego rodzajem jest wywiad telefoniczny (CATI) może wywoływać większą presję na społecznie aprobowane odpowiedzi²⁴.

Porównanie szczegółowych wyników uzyskanych za pomocą obu technik pokazało, że CAWI daje istotnie wyższe wskaźniki wiktyimizacji, także przy kontroli wpływu innych czynników, takich jak kraj przeprowadzenia badania, czy cechy społeczno-demograficzne badanych. Badanie stanowiło statystyczną analizę wielkości wpływu samej metody na te obserwowane różnice we wskaźnikach wiktyimizacji. Największy wpływ odnotowano w przypadku wskaźnika wiktyimizacji napaścią seksualną (*sexual assault*), a więc najbardziej wrażliwą z form wiktyimizacji. Większa skłonność do zaprzeczania takim czynom ujawniła w wywiadzie telefonicznym, czyli zakładającym chociaż pośredni kontakt z ankierem. Z kolei brak jakichkolwiek różnic zaobserwowano w przypadku najmniej poważnego zdarzenia – kradzieży roweru. Jednocześnie zaś było to proste i krótkie pytanie, a więc nie odgrywał w nim najpewniej roli problem ewentualnego niedostatecznego zrozumienia treści pytania w „szybszym” badaniu telefonicznym²⁵. Wykazano również pośredni wpływ błędu pokrycia (*coverage error*) i błędu braku odpowiedzi (*non-response error*) na wskaźniki wiktyimizacji. Po kontroli tego efektu techniki większość zaobserwowanych różnic znika, pozostaje jedynie w przypadku wiktyimizacji *sexual assault*. Badacze podkreślają, że pomimo, iż samodzielnie wypełniany kwestionariusz internetowy daje bardziej rzetelne wyniki

²⁰ F. Sztabiński, T. Żmijewska-Jędrzejczyk, *Mixed Mode Survey Design...*, s. 41.

²¹ M. Denscombe, *Web-Based Questionnaires and the Mode Effect. An Evaluation Based on Completion Rates and Data Contents of Near-Identical Questionnaires Delivered in Different Modes*, “Social Science Computer Review” 2006, vol. 24, (za:) D. Heerwegh, G. Loosveldt, *Assessing Mode Effect in a National Crime Victimization Survey using Structural Equation Models: Social Desirability Bias and Acquiescence*, “Journal of Official Statistics” 2011, vol. 27, s. 52.

²² N. Guzy, H. Leitgöb, *Assessing mode effects in online and telephone victimization surveys*, “International Review of Victimology” 2015, vol. 21.

²³ Tamże, s. 104.

²⁴ Tamże, s. 105-106.

²⁵ Tamże, s. 117-121.

odnośnie do wrażliwych pytań, to jednak problemy z właściwym doбором próby i jej reprezentatywnością nadal wykluczają pełną implementację tej metody do prowadzenia badań wiktyimizacyjnych²⁶.

Podsumowanie

Badania wiktyimizacyjne stanowią w kryminologii jedno z głównych źródeł rzetelnych danych ilościowych o rozmiarach wiktyimizacji, poziomie lęku przed przestępczością oraz społecznych ocenach organów ścigania i wymiaru sprawiedliwości. Badania te realizowane mogą być przy wykorzystaniu różnych technik badawczych, które z kolei mogą być obarczone różnego rodzaju błędami.

W artykule omówiono podstawowe rodzaje technik prowadzenia badań wiktyimizacyjnych oraz rodzaje błędów związanych z realizacją badań sondażowych. Opisano szczególne problemy, jakie może nieść za sobą wykorzystanie dwóch, alternatywnych do wywiadu bezpośredniego (PAPI lub CAPI), sposobów prowadzenia badań wiktyimizacyjnych, czyli sondaży telefonicznych (CATI) i kwestionariuszy internetowych (CAWI).

Generalnie, badania ilościowe są bardziej wrażliwe na wszelkie modyfikacje w zakresie sposobu prowadzenia badania, jego zewnętrznych warunków, czy na zmiany w treści pytań niż badania jakościowe. Bowiem z samego ich założenia, żeby można było mówić o porównywalności odpowiedzi poszczególnych respondentów, muszą być zachowane te same warunki prowadzenia badania. Dopiero wtedy odpowiedzi te można analizować łącznie i tworzyć na ich podstawie współczynniki lub wskaźniki. Bez spełnienia tego założenia możemy operować jedynie danymi jakościowymi charakteryzującymi każdą pojedynczą jednostkę i każdy przypadek przeprowadzonego badania. Każde odstępstwo od pełnej identyczności procesu badania może wywoływać szereg błędów, szczególnie istotnych przy próbie porównywania wyników uzyskanych za pomocą odmiennych technik badawczych. Podsumowując opisywane w literaturze przedmiotu różnice między rozpatrywanymi w artykule technikami prowadzenia badań wiktyimizacyjnych, warto jeszcze raz przytoczyć ich najważniejsze ustalenia.

Można spodziewać się, że przy wykorzystaniu dokładnie tych samych, dobrze skonstruowanych pytań, skierowanych do tych samych osób, ale zadawanych bądź przez telefon (CATI), bądź bezpośrednio (PAPI lub CAPI), bądź internetowo (CAWI) uzyska się praktycznie identyczne odpowiedzi. Zaobserwowano jednak istotny wpływ techniki badawczej na poziom realizacji próby i cechy zrealizowanej próby. Badania typu CATI, CAWI i wywiadu bezpośredniego zakładają odmienny sposób doboru próby i próby zrealizowane tymi technikami mogą się od siebie istotnie różnić. Te różnice mogą przekładać się na istotne rozbieżności w uzyskanych wynikach.

Ujawniły się także różnice rozpatrywanych technik odnośnie do odpowiedzi na pytania wrażliwe (np. o fakt wiktyimizacji napaścią seksualną) – w badaniu ankierskim (np. CATI) respondenci mieli większą skłonność do udzielania odpowiedzi przeczącej niż w przypadku samodzielnego wypełniania kwestionariusza. Osoby, które nie były ofiarami żadnego przestępstwa mają również większą skłonność do rezygnowania z udziału w całym badaniu wiktyimizacyjnym (zaburzając tym samym rozkład odpowiedzi).

Warte odnotowania są także zaobserwowane różnice poszczególnych technik w zakresie skłonności do udzielania odpowiedzi społecznie aprobowanych. Wydaje się, że najsilniejsza

²⁶ Tamże, s. 122.

tendencja do udzielania takich odpowiedzi może ujawniać się w przypadku wywiadów bezpośrednich i telefonicznych, najsłabsza zaś w odniesieniu do kwestionariuszy uzupełnianych samodzielnie oraz ankiety pocztowej i internetowej.

Konstruując badanie według metodologii CATI trzeba zwrócić szczególną uwagę na reprezentatywny dobór próby. Warto zastanowić się też nad innym doбором badanych niż losowy dobór numerów stacjonarnych telefonów (*random digit dialing – RDD*). Z kolei technika CAWI wydaje się być obiecującą formą prowadzenia badań panelowych, na próbie zidentyfikowanych wcześniej respondentów.