

ADAM RYBAK

UNIwersytet im. Adama Mickiewicza w Poznaniu

<http://dx.doi.org/10.18778/8142-715-9.11>

Zjawisko preferencji techniki sondażu w kontekście badań reprezentatywnych w schemacie mieszania technik¹

Preferencja techniki sondażu – przyczyny refleksji

Zmniejszające się z roku na rok wskaźniki realizacji próby sondażowej uważane były już w latach 70. XX wieku za poważny problem przy realizacji badań reprezentatywnych (Rossi, Wright i Anderson, 1983, s. 9) mimo zaledwie 30 lat historii ich szerokiego stosowania (Groves, 2011, s. 861). W Polsce po początkowym entuzjazmie względem nowej, wprowadzonej w latach 50. metody również rozpoczął się okres ciągłego spadku *response rate* – datujący się od czasu wprowadzenia stanu wojennego (Sztabiński, Żmijewska-Jędrzejczyk, 2012, s. 33). Trend ten utrzymał się, dziś coraz niższe wskaźniki realizacji badania utrudniają wyciąganie uprawnionych wniosków z sondaży – ze względu na to, że błąd wynikający z braków odpowiedzi jest najpewniej zasadniczym źródłem błędów pomiaru (Jabkowski, 2015). Choć głosy przedstawicieli środowiska badacza wahają się, od tych, dla których jest to po prostu zwyczajny element pracy związanej z sondażem, po takich, którzy uważają to za możliwy początek końca dyscypliny, panuje jednak powszechna zgoda co do tego, że wysiłki na

¹ Artykuł ten oparty jest na niepublikowanej pracy magisterskiej obronionej przez autora w 2018 roku w Instytucie Socjologii UAM.

rzecz rozsądnego zwiększania udziału wylosowanych osób w badaniu są konieczne i pożyteczne.

Podejść do rozwiązania problemu spadającego *response rate* było wiele. Za najważniejszą i najbardziej znaną należy uznać w warunkach amerykańskich stworzoną i rozwijaną przez Dona D. Dillmana koncepcję *tailored design method*, nakazującą planowanie badania jako całościowego procesu, który uwzględnia jednocześnie wiele czynników wpływających z jednej strony na przewidywanie przez jednostkę korzyści z udziału w badaniu, z drugiej – na jej przekonanie o niewielkim rozmiarze kosztów (Dillman, Smyth i Christian, 2014). Kolejną ważną koncepcją jest Roberta Grovesa teoria *leverage–saliency*, skupiająca się przede wszystkim na roli ankietera mającego podejmować dynamiczne wysiłki dopasowania podejścia do osobistych psychologicznych skłonności potencjalnego respondenta poprzez retoryczne podkreślanie pewnych aspektów badania, a umniejszanie innych, co wymaga od niego socjologicznej wrażliwości (Grzeszkiewicz-Radulska, 2009). Z kolei w warunkach europejskich największą rolę odegrał prawdopodobnie Europejski Sondaż Społeczny, który poza głównym celem badawczym stawia przed sobą także cele metodologiczne. Dane z realizacji kolejnych jego rund są znakomitymi źródłami do wykorzystania podczas prowadzenia analiz przez europejskich i polskich metodologów.

Jedną z koncepcji, która wyrosła z analiz braków odpowiedzi w sondażach, ale z czasem rozszerzyła się także na inne pola – jak np. minimalizowanie błędów pomiaru – jest hipoteza o preferowaniu przez respondentów udziału w badaniu sondażowym za pośrednictwem konkretnej techniki (ang. *mode preference*). Po raz pierwszy pojawiła się ona w badaniu Grovesa i Roberta Kahna (1979, s. 88–90) przeprowadzonym niedługo po rozwinęciu się na szerszą skalę badań CATI – dlatego autorzy zainteresowani byli poszukiwaniem źródeł tego, że niektórzy respondenci mogą nie chcieć lub nie lubić uczestniczenia w sondażach telefonicznych. Zadawali oni w przeprowadzonych równolegle badaniach PAPI i CATI pytania zarówno o preferowaną przez badanych technikę, jak i konkretne przyczyny tej preferencji. Zauważyli takie powszechnie uznawane do dziś tendencje, jak wskazywanie jako preferowanej techniki, w której aktualnie się uczestniczy – silniejsze jednak dla sondażu ankietarskiego – być może ze względu na grzeczność, chęć nieurazenia ankietera. Warto wspomnieć, że jako główną przyczynę preferowania badania bezpośredniego respondenci wskazywali przyjemność płynącą z kontaktu z ankieterem, dla badania telefonicznego – szybki, niewymagający charakter oraz brak konieczności wpuszczania obcej osoby do domu, a dla ankiety pocztowej – brak pośpiechu, możliwość przemyślenia odpowiedzi na pytania.

Teorie udziału w sondażu i preferencji techniki

Zjawisko preferencji techniki przez respondentów wymaga umocownia teoretycznego, by możliwa była jego rzeczowa analiza – samo stwierdzenie, że badani prawdopodobnie preferują którąś z technik i chętniej zgodzą się na badanie, nie jest niewystarczające. Mimo że hipoteza ta jest w nauce obecna, jak było wspomniane, już od prawie 40 lat, to wciąż brakuje jej solidnej koncepcji pojęciowej i analiz empirycznych (Olson, Smyth i Wood, 2012, s. 612). Poniżej przedstawię teorie lub koncepcje obecne w literaturze i mające zastosowanie do rozważań nad wpływem techniki na udział w badaniu.

Choć Groves był tym, który jako pierwszy zwrócił uwagę na to zjawisko, to nie odgrywa ono szczególnej roli w jego systemie. Analizując sposoby zapobiegania brakom odpowiedzi w sondażach, sugeruje on możliwość modyfikacji badania (np. stosowania różnej długości kwestionariuszy albo wybiórczych zachęt materialnych) w stosunku do reprezentantów określonych kategorii społecznych możliwych do wyróżnienia na podstawie operatu, co do których zachodzi przypuszczenie, że np. wykazują bardzo niskie zainteresowanie tematem badania (Groves, Couper, 1998, s. 45). Rzadko i zdawkowo wspomina jednak o stosowaniu różnych technik w zależności od preferencji respondenta (Groves, Singer i Corning, 2000).

Także sama analiza cech społeczno-demograficznych respondentów mających mieć wpływ na partycypację – takich jak płeć, wiek, wykształcenie, dochód i stan cywilny, a także cechy z nimi powiązane, takie jak struktura i charakterystyka gospodarstwa domowego, związek tematu z cechami respondenta (np. wiek, obecność w kwestionariuszu pytań odnośnie do sytuacji osób starszych) czy doświadczenie w sytuacjach podobnych do sondażu (np. większe przyzwyczajenie do egzaminów wśród osób z wyższym wykształceniem) – nie jest dla niego sprawą centralną. Twierdzi on (1998, s. 31–41), że są one jedynie podstawą formowania się u respondentów pewnych dyspozycji psychicznych do reakcji na prośbę udziału w sondażu. Kwestia tych predyspozycji jest kluczowa ze względu na charakter decyzji, jaką podejmuje badany – zdaniem Grovesa zdecydowana większość osób nie ma wyrobionej opinii o udziale w sondażach, nie może odwołać się więc do stałej postawy ani też ocenić ogólnych korzyści i strat przez kalkulację wykorzystującą solidne informacje. Zamiast tego decyzja taka jest procesem heurystycznym, opartym na wrażliwości na poznawcze skrypty. Autor opiera się tutaj na koncepcji Roberta Cialdiniego (1984), wyróżniającego czynniki odgrywające w takiej sytuacji główną rolę. Są to przede wszystkim zasada wzajemności, autorytetu, zaangażowania, niedostępności dóbr, pragnienia słuszności i lubienia. Groves proponuje wykorzystanie ich w procesie projektowania i przeprowadzania badania, w dużym stopniu

podkreślając wspomnianą powyżej dynamiczną rolę ankietera. W takim ujęciu nie możemy traktować preferencji techniki przez respondenta jako sprawy uświadomionej, o którą możemy go zapytać. Jej związek z łatwymi do ustalenia cechami społeczno-demograficznymi prawdopodobnie istnieje, ale jest zapośredniczony. Pytanie dotyczy wtedy głównie tego, jakie elementy związane z daną techniką mogą spowodować pozytywną lub negatywną spontaniczną reakcję respondenta.

Inaczej Don D. Dillman (Dillman, Smyth i Christian, 2014), który problematykę *mode preference* uwzględnia obszernie, a także bierze udział w jej badaniach. Choć krytykuje on inne teorie, w tym autorstwa Grovesa, które uważa za zbyt „psychologiczne”, sam umieszcza proces decyzyjny przyszłego respondenta w centralnym miejscu rozważań. Opierając się na teorii wymiany, twierdzi, że decyzję o partycypacji wywołuje przekonanie jednostki o tym, że zyski osiągnięte w wyniku podjęcia działania przeważą nad kosztami. Autor czyni zastrzeżenie, że nie należy traktować tego jako zachowania w pełni racjonalnego w wąskim rozumieniu ze względu na to, że po pierwsze, istotny wpływ mają takie elementy jak altruizm, zasada wzajemności i zaufanie, po drugie natomiast – charakter zysków i strat jest nieprecyzyjny, opiera się częściowo na wierze i może być odłożony na nieokreśloną przyszłość. Jednak nie oznacza to, że nie możemy mówić o rozsądnym namyśle i wważeniu za i przeciw – nawet jeśli decyzja musi zostać podjęta bardzo szybko, jak w wypadku sondażu telefonicznego. Autor analizuje czynniki, które mają wpływ na poszczególne elementy decyzji – zyski, koszty i zaufanie – które trzeba uwzględnić w holistycznym planowaniu badania. Wśród sposobów zmniejszania strat badanego zwraca uwagę na interesującą nas kwestię – tego, jaka technika będzie sprawiała respondentowi największy kłopot, a jaka będzie najwygodniejsza. Choć nie oznacza to, że ludzie na co dzień analizują problem tego, czy wolą spotkanie z ankieterem, czy ankiety internetowe, to jednak spytani o preferencje mogą udzielić sensownej odpowiedzi, o której można przypuszczać, że będzie podobna do ich rzeczywistego zachowania.

Choć wielu badaczy nie tworzy tak całościowych koncepcji zachowań respondentów, to jednak na potrzeby swoich badań buduje robocze zbiory czynników, które mogą być determinantami preferencji techniki, tym samym sugerując, gdzie należy poszukiwać źródła tego zjawiska. Na przykład Roger Tourangeau, Lance J. Rips i Kenneth Rasinski (2000, s. 289–312), badacze wpisujący się w nurt *cognitive aspects of survey methodology*, który kładzie szczególny nacisk na badania procesów poznawczych zachodzących w trakcie udziału w sondażu, sugerują, że wpływ na preferencję techniki może mieć struktura gospodarstwa domowego – osoby mieszkające z innymi mogą niepokoić się, że inni domownicy będą podsłuchiwać trwający nawet w osobnym pomieszczeniu wywiad, tym samym naruszając ich poczucie prywatności. Różnicować preferencję może także odczuwanie

tematu badania jako drażliwego tylko przez część respondentów – można zasadnie przypuszczać, że rozmowa o innych kwestiach stanowi większy problem dla młodych kobiet niż dla starych mężczyzn – w takiej sytuacji osoby uważające temat za drażliwy preferować będą techniki bardziej bezosobowe (autorzy czynią tu słuszną uwagę, by nie humanizować w takim przypadku interfejsu badań drogą elektroniczną, umniejszy to bowiem pozytywny efekt – a jeśli w kwestionariuszu zamieszczone zostaną zdjęcia awatara ankietera czy nagranie jego naturalnego głosu, to może wystąpić efekt ankierski znany z badań bezpośrednich). Osoby wrażliwe na kwestię powagi instytucji sponsorującej badanie będą mogły być poddane lepszemu wpływowi przez zastosowanie technik mogących bardziej podkreślić ten element – czyli wywiadu bezpośredniego i pocztowego (np. pismo z oficjalną pieczęcią). Obciążenie poznawcze wywoływane przez narzędzie może być różne w zależności od kompetencji. Wpływ może mieć także zamożność – wolniejszy komputer czy łącze internetowe uprzykrza udział w internetowej ankiecie. Także skojarzenia wywoływane przez technikę – np. widzenie wywiadu bezpośredniego jako rodzaju pogawędki, ankiety pocztowej i elektronicznej jako egzaminu, a badania CATI jako odmiany sprzedaży/marketingu – może mieć różny wpływ na różne kategorie respondentów. Wreszcie pewne grupy wiekowe mogą zachowywać się zgodnie ze swoim stereotypowym postrzeganiem – np. emeryci będą się bali wpuścić obcą osobę do domu, za to będą stosunkowo podatni na sprzedaż telefoniczną.

Kieren Diment i Sam Garrett-Jones (2007, s. 411) zwracają uwagę na kwestię wciąż istniejącej różnicy w stosunku do używania internetu ze względu na płeć – bardziej pozytywne nastawienie mężczyzn – a także na wpływ sektora zatrudnienia – ci pracujący w branżach związanych z kontaktem z urządzeniami elektronicznymi będą mieli większą łatwość w udziale w sondażu tą drogą. Z kolei Santosh Verma i in. (2014, s. 73) zwracają uwagę na kwestie przynależności do mniejszości etnicznej, wieku i wykształcenia jako mających mieć wpływ na stosunek do używania nowych technologii w sondażach. Wreszcie Jolene Smyth, Kristen Olson i Morgan M. Millar (2014, s. 135–137) stawiają hipotezę, że wpływ na *mode preference* mogą mieć takie kwestie jak przyzwyczajenie i umiejętność korzystania z danego medium, obecność czynników zakłócających, np. dzieci – jeśli obecność ankietera jest wystarczającą wymówką do odseparowania się od reszty domowników, to jednostki wybiorą tę technikę; jeśli to niemożliwe, będą wołały kwestionariusze samodzielnie wypełniane. Istotną rolę może odgrywać kwestia tego, w jakim stopniu dana technika jest wymagająca – ankieter może ułatwiać zrozumienie pytań, z kolei poczta i internet dają czas na zastanowienie – należy tu brać pod uwagę wykształcenie respondenta, jego kompetencje językowe oraz dostępne mu zasoby pamięci roboczej. Na stosunek do preferowania lub nie wpływ mogą mieć poczucie

własnej wartości i potrzeba autoprezentacji – jeśli są wysokie, to wywiad ankierski wiąże się z wyższymi subiektywnymi kosztami badanego. Istotne są również potrzeba prywatności i preferowanie towarzystwa lub samotności. Wreszcie autorki zwracają uwagę na bardzo interesującą i w sytuacji ciągłego starzenia się społeczeństw zachodnich niezwykle istotną sprawę występowania związanych z wiekiem schorzeń wzroku i słuchu, które powinny mieć istotny związek z wyborem techniki – takiej, w której dane scharczenie będzie mniej przeszkadzać.

Przedstawienie poszczególnych koncepcji umożliwia zrozumienie, jakie czynniki mogą tkwić u podstaw preferowania przez respondentów jednych technik, a niechęci do innych, przyczyn przypuszczalnej trwałości lub nie takiej postawy, jej racjonalnej lub też czysto spontanicznej natury. Jednocześnie zarysowane zostają podstawowe rozróżnienia sposobów badania *mode preference*: poprzez postawienie respondenta w sytuacji wyboru techniki, spytanie go o preferencję w jej zakresie lub też analizowanie stopnia udziału respondentów z danej kategorii społeczno-demograficznej lub podobnej w badaniu sondażowym.

Czym jest *mixed-mode survey design*?

Praktyczne zastosowanie idei preferencji techniki wymaga użycia schematu zwanego *mixed-mode survey design*. Nie jest to jednak szczególnie problem – badania realizowane w ten sposób są w wielu miejscach zachodniego świata obowiązującą normą, w Japonii natomiast właściwie nie zdarza się, by sondaż był przeprowadzany w inny sposób (de Leeuw, 2005, s. 235). W najogólniejszym rozumieniu sondaż przeprowadzony w schemacie mieszania technik to takie badanie, które w obrębie całego procesu – nawiązania kontaktu, zbierania danych, dostarczania zapowiedzi, przypomnień i ponagleń – używa więcej niż jednego medium kontaktu z respondentem. Wąskie rozumienie natomiast dotyczy wyłącznie samego zbierania danych. W dalszej części będę odwoływać się do szerokiego rozumienia, chyba że zaznaczone zostanie inaczej.

Najwcześniejszą przyczyną stosowania schematu *mixed-mode* była chęć zmniejszania błędu pomiaru (Dillman i in., 2014) – zauważono wpływ ankietera na chęć odpowiadania na pytania społecznie wrażliwe – tematy uznawane przez respondenta za wstydlive albo silnie związane z prestiżem (respondent nie chciał pokazywać się z gorszej strony) – bądź niezgodną z rzeczywistością treść odpowiedzi na te pytania. Dlatego, by korzystać z zalet techniki wywiadu kwestionariuszowego – możliwości motywowania respondenta i tłumaczenia mu niezrozumiałych pytań, nadzorowania reguł przejścia – wprowadzono kwestionariusze do samodzielnego wypełniania,

dotyczące wyłącznie wrażliwych obszarów badania. Kwestionariusz taki był wręczany respondentowi, który wypełniał go przy ankiecie, ale nie pokazując mu swoich odpowiedzi – mógł więc czuć się bardziej anonimowy. Jest to rozwiązanie szczególnie często stosowane w dziedzinie sondaży dotyczących tematyki zdrowotnej. Idea mieszania technik pojawiła się wśród szerszego grona badaczy na początku lat 60. (Couper, 2011, s. 891–892), jednak proces jej upowszechniania był powolny, a sama analiza teoretyczna konsekwencji takiego podejścia do dzisiaj daleka jest od kompletności.

Większość analiz naukowych dotyczących sondaży *mixed-mode* skupia się na problematyce zmniejszania odsetka braków odpowiedzi (de Leeuw, Hox, Dillman, 2008, s. 302). Dążenie to może realizować się poprzez odwołanie do hipotezy *mode preference* (albo do jej bardzo osłabionej formy sugerującej, że z reprezentantami niektórych kategorii społecznych łatwiej jest się skontaktować za pomocą jednych mediów niż drugich) lub do założenia, że danie respondentowi wyboru techniki wywoła u niego poprawę nastawienia do sondażu/ankietera i pozytywną odpowiedź na prośbę o udział w badaniu. Może też dotyczyć wykorzystywania kombinacji różnych technik w różnych fazach badania – np. telefonicznych kontaktów przypominających o niewypełnieniu kwestionariusza pocztowego. Wreszcie dotyczyć może tak specyficznych kwestii jak podanie respondentom nieznającym języka kraju macierzystego numerów telefonów do ankiesterów wielojęzycznych, którzy z różnych przyczyn nie mogliby efektywnie uczestniczyć w badaniu terenowym. Należy jednak zwrócić uwagę, że nie wszystkie badania dotyczące tej problematyki biorą pod uwagę skomplikowane zależności pomiędzy samą wielkością wskaźnika realizacji próby a rzeczywistym błędem braku odpowiedzi (por. Groves, 2006).

Kolejne zastosowanie *mixed-mode* może dotyczyć redukcji kosztów badania przez użycie tańszych technik badawczych przy jednoczesnej minimalizacji ułomności tych technik – np. używanie podwójnych operatorów łączących badania telefoniczne i bezpośrednie, co umożliwia dotarcie również do tych respondentów, którzy nie mają telefonów (choć dziś być może ciekawszym problemem jest łączenie jeszcze tańszych i szybszych sondaży internetowych z technikami uzupełniającymi, które umożliwią kontakt z osobami bez dostępu do sieci lub z deficytem kompetencji cyfrowych), stosowanie tańszej techniki w pierwszej fazie badania itd. Dodatkowo wpływ na koszty może mieć próba zmniejszania czasu trwania badania przez dopasowanie techniki do preferencji respondenta (por. Olson i in., 2012). Kwestie kosztów są jednak trudne do analizy, gdyż badacze niezbyt często podają je do publicznej wiadomości.

Wreszcie stosunkowo nowym zagadnieniem są badania dotyczące wpływu *mode preference* na jakość udzielanych przez respondenta odpowiedzi (por. Smyth, Olson i Kasabian, 2014a), które zostaną omówione

w dalszej części artykułu. Widoczne jest więc, że zastosowania tej techniki dotyczą redukcji wielu elementów całkowitego błędu pomiaru. Jednak ta skomplikowana konstrukcja badania rodzi również potencjalne ryzyko zwiększenia błędu. Ewentualne stosowanie łączonych operatów znacząco utrudnia analizę wielkości zmiany precyzji wnioskowania płynącej z odejścia od losowania prostego, za pomocą współczynnika DEFF (Jabkowski, 2015). Łączenie operatów może również powodować nakładanie się błędów płynących np. z ich nieaktualności. Obecnie najpopularniejszym przypadkiem stosowania tzw. *dual-frame* jest łączenie operatu (choć jest to sformułowanie nieco na wyrost, mówimy bowiem najczęściej o losowym generowaniu numerów z zastosowaniem kwot, a nie klasycznym doborze próby ze spisu) użytkowników telefonów stacjonarnych i komórkowych. Choć mamy tu do czynienia z zupełnie innymi operatami – telefony stacjonarne przypisane są do gospodarstwa domowego i wymagają drugiego stopnia losowania konkretnego respondenta, a komórkowe są zazwyczaj własnością konkretnej osoby (Jabkowski, 2015) – to w literaturze oba przypadki przyporządkowuje się najczęściej do jednej techniki – CATI, nie będą więc zajmować się tym problemem w niniejszym artykule.

Istotnym obszarem potencjalnych błędów jest również niejasność zjawiska nakładania się efektów konkretnych technik, których wpływ na pomiar stwierdzono (de Leeuw, 2008), a wreszcie kwestia „tłumaczenia” pytań pomiędzy technikami, tak by rzeczywiście w ten sam sposób mierzyły tę samą zmienną i by różnice w formie i sposobie prezentacji nie powodowały różnic w pomiarze. Tą kwestią zajmę się w dalszych częściach artykułu.

Możliwe układy technik na wszystkich etapach badania

Jak zostało wspomniane, łączenie technik sondażu dotyczy wszystkich etapów kontaktu z respondentem. Różne techniki mogą być zastosowane na różnych etapach, a także w obrębie jednego etapu może być stosowanych kilka technik – wtedy pojawia się kwestia kolejności ich użycia: równoczesnej bądź sekwencyjnej. Poniżej przedstawiono rozszerzoną koncepcję klasyfikacji układów *mixed-mode* w zależności od etapu badania autorstwa Edith de Leeuw i Joopa J. Hoxa (2008), uzupełnioną przykładami zastosowań w sytuacjach badawczych.

Pierwszą styczność z respondentem zapewnia kontakt zapowiedni – w technice, która zapewni założony efekt, zmieści się w planowanych kosztach i będzie możliwa do stosowania przy danym operacie/operatach. Kiedy technika ta różni się od techniki stosowanej np. na etapie zbierania danych, możemy już mówić o schemacie *mixed-mode*. Zastosowanie tańszej techniki jako zapowiedniej pozwala np. wykluczyć z próby nieistniejące

adresy przy technice pocztowej albo uzyskać informację zwrotną o śmierci czy trwałej wyprowadzce respondenta w sytuacji podania adresu/numeru kontaktowego – dzięki temu można już na wstępie doprowadzić do większej aktualności operatu. Zastosowanie eleganckich listów zapowiednich zaopatrzonych w emblematy poważnych instytucji badawczych pozwala zwiększyć zaufanie potencjalnego respondenta do badacza, tym samym zwiększając *response rate* (Tourangeau i in., 2000, s. 307). Odpowiednie dobranie treści listów zapowiednich do kategorii respondentów (możliwych do wyznaczenia na podstawie operatu), choć nie jest ściśle rzecz biorąc mieszaniem technik, jako elastyczne dopasowanie w obrębie jednej techniki również może przynieść dobre rezultaty. Edwards i in. (2014) w badaniu łączącym zbieranie danych techniką pocztową oraz za pomocą *web-push* (kontaktu pocztowego z linkiem internetowym – obie techniki stosowały więc jeden operat) porównywali wskaźniki odpowiedzi, w sytuacji gdy respondent z jednego z dwóch badanych stanów w USA dostawał list zapowiedni z logo uniwersytetu znajdującego się w obrębie jego stanu lub innego. Lokalne logo zawsze zwiększało *response rate*, w jednym ze stanów różnice były istotne statystycznie. Stosowanie najpierw tańszej metody nie jest jednak ani konieczne, ani nawet zawsze wskazane. Stephen R. Porter i Michael E. Whitcomb (2007) zbadali 3 tysiące absolwentów i 3 tysiące niedoszlých studentów jednej z amerykańskich uczelni. Ze względu na to, że zbierała ona dane o adresach pocztowych i elektronicznych wszystkich kontaktujących się, możliwe było zastosowanie mieszania technik na wszystkich etapach, ale samo zbieranie danych odbywało się tylko drogą internetową. W tym miejscu interesuje nas użycie go na etapie pierwszego kontaktu – zastosowanie pocztowego listu zapowiedniego zwiększyło wskaźnik odpowiedzi w każdym przypadku, analiza wariancji porównująca tę miarę w sytuacji użycia zapowiedni pocztowej i nieużycia żadnej wykazała istotną statystycznie różnicę. Możliwe jest także zastosowanie naraz kilku metod kontaktów zapowiednich, a tym samym – mieszanie technik w obrębie pierwszej fazy badania – Morgan Millar i Dillman (2011) w sondażu *mixed-mode* przeprowadzonym wśród studentów porównywali odpowiedzi w kategoriach osób, które proszono o wypełnienie ankiety internetowej, pocztowej lub dano im wybór, przy czym pierwszy kontakt odbywał się drogą pocztową. Dodatkowo wprowadzono równoliczną (700 osób) grupę czwartą, która otrzymała list zapowiedni jednocześnie drogą pocztową i elektroniczną. Autorzy uzasadniali to posunięcie tym, że z jednej strony kontakt pocztowy, który może przybrać poważną formę listu urzędowego z pieczęcią, tym samym zwiększając zaufanie, jest często trudniejszy do przeoczenia niż jeden z wielu e-maili w skrzynce i umożliwia zastosowanie bezpośrednich, bezwarunkowych zachęt gotówkowych, które są skuteczniejsze niż inne formy wynagradzania (Dillman i in., 2014). Z drugiej list elektroniczny umożliwia wygodniejsze przejście do ankiety internetowej,

bez konieczności dodatkowego uruchamiania urządzenia elektronicznego czy przepisywania adresu, nie obniża wskaźnika odpowiedzi przez konieczność zmiany medium kontaktu. Zastosowanie podwójnej zapowiedzi ankiety internetowej doprowadziło do wyższego wskaźnika odpowiedzi niż w pozostałych układach technik i były to różnice istotne statystycznie. De Leeuw wyróżnia dodatkowo etap *screeningu*, czyli pierwszego kontaktu z respondentem, który ma przede wszystkim sprawdzić, czy rzeczywiście znajduje się on w próbie oraz czy, kiedy i w jakiej formie chce uczestniczyć w badaniu. Nie będę opisywać perspektyw użycia mieszanych technik w obrębie tego etapu lub między etapami, gdyż cechy tego rozwiązania pokrywają się zasadniczo z omówionymi powyżej.

Po uzyskaniu kontaktu i ewentualnym umówieniu terminu następuje proces zbierania informacji. Możliwe są następujące konfiguracje mieszania technik:

- 1) równoczesne zbieranie danych od wszystkich członków próby za pomocą różnych technik;
- 2) sekwencyjne zbieranie danych od wszystkich członków próby za pomocą różnych technik;
- 3) zbieranie danych od poszczególnych członków próby za pomocą różnych technik (różne części kwestionariusza);
- 4) zbieranie danych od tej samej (mniej więcej) próby w różnych technikach w różnych punktach w czasie (czyli zmiana techniki panelu);
- 5) wreszcie możliwe, choć poznawczo niepewne, jest stosowanie różnych technik do różnych populacji w badaniach porównawczych.

W przypadku pierwszym należy wprowadzić dwa dodatkowe rozróżnienia. Po pierwsze, czy respondenci zostaną przydzieleni do techniki przez badacza, czy pozostawi się im wybór. Jeśli to badacz decyduje, to dodatkową kwestią jest uzasadnienie kryterium podziału. Pozostawienie wyboru wynika najczęściej z założenia albo o istnieniu w jakiejś formie preferencji techniki wśród respondentów, do której dopasowanie się miałyby poprawić jakość pomiaru w różnych aspektach, a o której nie mamy wcześniejszej informacji (np. z poprzedniej rundy sondażu albo z jakiegoś modelu teoretycznego), albo po prostu (w przypadku braku takiej refleksji) jest to forma pokazania respondentowi, że jest on ważnym, istotnym, decyzyjnym podmiotem (co jednostka w przeciwnym razie mogłaby chcieć osiągnąć np. przez odmowę udziału w badaniu) i tym samym wzbudzenie jego wdzięczności i zwiększenie motywacji.

Jakie są konsekwencje pozostawienia wyboru respondentom? Rebecca L. Medway i Jenna Fulton (2012) przeprowadziły metaanalizę 19 badań eksperymentalnych wykonanych od 2000 do 2011 roku, porównujących *response rate* dla sytuacji, gdy respondenci zostali poproszeni o wypełnienie ankiety pocztowej z sytuacją dania im wyboru (poczta) albo kwestionariusza internetowego. W 17 z 19 analizowanych badań danie respondentom

wyboru zwiększyło liczbę braków odpowiedzi – średnio o 3,8 p.p. Jednocześnie w grupach z wyborem średnia liczba osób wybierających ankietę internetową wynosiła 16,6% (mediana to odpowiednio 3,4 p.p. i 10,2%). Ważony iloraz szans dla całego badania mający za wartość odniesienia odpowiedzi w grupie „tylko poczta” wyniósł 0,87 przy poziomie istotności $p < 0,001$. Jednocześnie autorki zauważają, że wpływ innych zmiennych, takich jak istotność tematu dla respondentów czy zastosowanie zachęt materialnych, na wielkość różnicy między obiema grupami eksperymentalnymi we wszystkich badaniach był znikomy (Medway, Fulton, 2012, s. 741).

Widzimy więc, że zaoferowanie wyboru respondentowi niesie negatywne konsekwencje dla wskaźnika odpowiedzi. Należy jednak zaznaczyć, że niekoniecznie musi się to przekładać na wzrost błędu (Groves, 2006). Jednocześnie można przypuszczać, że zaoferowanie wyboru techniki będzie skutkowało idealnym dostosowaniem tej oferowanej do preferowanej. Na pozytywny wpływ takiej sytuacji na jakość pomiaru wskazują badania Smyth, Olson i Alian Kasabian (2014), natomiast Joseph W. Sakshaug, Ting Yan i Tourangeau (2010) podkreślają konieczność poszukiwania złotego środka pomiędzy *response rate*, który najczęściej osiąga najwyższe wartości w badaniach ankierskich, a wiarygodnością odpowiedzi, która często może wzrastać w przypadku oferowania kwestionariuszy do samodzielnego wypełniania. Obydwa badania zostaną szczegółowo omówione później.

Z kolei narzucenie respondentom techniki przez badaczy może znajdować zastosowanie przede wszystkim w trzech sytuacjach. Po pierwsze, gdy na podstawie wyników badań na temat wpływu zmiennych społeczno-demograficznych na *mode preference* od razu kierują oni prośby o udział w konkretnej technice do kategorii możliwych do wyróżnienia na podstawie operatu lub możliwych do częściowego wywnioskowania z operatu – np. w przypadku próby adresowej wskazanie osiedli o homogenicznym składzie pod względem wykształcenia, rasy czy pochodzenia społecznego. Jest to łatwiejsze w przypadku bardziej rozwarstwionych społeczeństw, w których poszczególne klasy separują się od siebie, czy też dużych populacji źle zintegrowanych imigrantów, zamieszkujących możliwe do wyznaczenia getta – w warunkach polskich takie podejście miałoby zapewne ograniczone zastosowanie. Jakie badania możemy brać pod uwagę w przypadku zakładania nierównego rozkładu *mode preference* wśród różnych kategorii? Millar, O’Neil i Dillman (2009) przeanalizowali wyniki sondażu Lewiston-Clarkston Quality of Life Survey, przeprowadzonego w 2007 i 2008 roku na losowej próbie adresowej 2800 mieszkańców wskazanego w nazwie obszaru, który był badaniem łączącym technikę pocztową i internetową. Poszczególne grupy eksperymentalne odpowiadały w jednej z przydzielonych technik, a dodatkowo udzielały odpowiedzi na pytanie o preferowaną technikę. Poza deklarowaną preferencją tej ostatniej zmiennymi istotnie

różnicującymi populację respondentów i nierespondentów były wiek (starsza była populacja respondentów pocztowych), poziom wykształcenia (wyższy dla respondentów internetowych), w tej zbiorowości wyższy był także średni dochód. Wśród respondentów internetowych istotnie przeważali także mężczyźni i osoby zamężne. Diment i Garrett-Jones przeprowadzili sondaż na losowej próbie 1100 badaczy zrzeszonych w australijskim Cooperative Research Centre, dając im wybór pomiędzy ankietą pocztową a internetową. Znow wiek i płeć istotnie różnicowały zbiorowości osób preferujących daną technikę – internet częściej wybierali młodszy i mężczyźni. Natomiast zależność między preferencją a pozycją zawodową była nieliniowa – internet był częściej preferowany przez osoby zajmujące średnie pozycje w instytucjach.

Za szczególnie istotne należy uznać badania nakierowane na stwierdzenie (Haan, Ongena i Aarts, 2014) wpływu techniki na możliwość przeprowadzenia wywiadu z kategoriami respondentów, którzy ogólnie są uznawani za trudnych do skłonienia do udziału w badaniu lub nawet nawiązania z nimi kontaktu. Wskazany sondaż objął losową próbę adresową mieszkańców Holandii, przy czym grupy uznane za szczególnie problematyczne – mieszkańcy nowych osiedli (młode, zapracowane osoby), obszarów biedy (główny nacisk na imigrantów) i wielkich miast – zostały celowo nadreprezentowane w próbie. Istotne wyniki osiągnięto przy zastosowaniu kontaktu bezpośredniego z ankierem (poprzedzonego listem zapowiadającym) i daniu respondentom wyboru pomiędzy sondażem CAPI a wypełnieniem ankiety internetowej. Młodzi dorośli (poniżej 35 lat) oraz pracownicy pełnoetatowi wybierali badanie internetowe częściej niż bezpośrednio w znacząco większym stopniu niż reszta populacji. Z kolei badania (Verma i in., 2014), w których proszono 475 pracowników amerykańskich restauracji o udzielanie przez 12 tygodni informacji przez system Interactive Voice Response (odpowiadanie automatowi przez telefon, a więc badania kanałem słuchowym bez udziału ankiera), kwestionariusze internetowe lub pocztowe wykazały, że częstość wybierania internetu względem IVR spadała wraz z wiekiem; respondenci hiszpańskojęzyczni (w porównaniu z angielskojęzycznymi) i czarnoskórzy niełatynoscy (w porównaniu z białymi niełatynoskimi) oraz niemający wyższego wykształcenia wybierali częściej IVR niż internet – przy czym wszystkie formy kontaktu były dostępne w języku portugalskim, hiszpańskim i angielskim. Dodatkowo respondenci korzystający z kanału internetowego nieco częściej wypełniali kwestionariusze w zadanych terminach.

Wreszcie trzeci przykład, który z jednej strony pokazuje możliwe predyktory *mode preference* w populacji, a z drugiej odwołuje się do kolejnego sposobu wykorzystania sondażu *mixed-mode* z narzucaną respondentom techniką. W przeprowadzonym w 2008 roku sondażu telefonicznym Nebraska Annual Social Indicators Survey zapytano respondentów o preferowaną

przez nich technikę, w sytuacji gdyby chcieli w przyszłości ponownie uczestniczyć w badaniach (Smyth, Olson i Millar, 2014). Następnie skonstruowano wielomianowy model logistyczny przewidujący preferencję dla badań telefonicznych (49,2% respondentów), pocztowych (24,6%) bądź internetowych (19,7%). Skłonność do badań *face-to-face* wskazało zaledwie 1,7% respondentów, więc wykluczono ich z analizy. Prawdopodobieństwo wskazania techniki internetowej (w odniesieniu do telefonicznej) istotnie wzrastało wraz ze spadkiem wieku, było także wyższe dla osób z wyższym wykształceniem i deklarujących, że korzystają z internetu, natomiast istotnie spadało dla osób deklarujących niskie umiejętności korzystania z komputera. Z kolei wzrost prawdopodobieństwa (konkretnie – ryzyka względnego) wskazania ankiety pocztowej w stosunku do telefonicznej następował dla respondentek płci żeńskiej i osób deklarujących, że boją się kradzieży informacji na ich temat.

Warto zwrócić w tym miejscu uwagę czytelnika na powracający problem: czy efekt wieku jest efektem kohortowym, wynikającym z wychowywania się w środowisku sprzed rewolucji cyfrowej (ale w takim razie dlaczego działa on też na niekorzyść ankiety pocztowej?), czy też jest związany z jakimiś procesami następującymi wraz ze starzeniem się respondentów, np. spadkiem efektywności pamięci roboczej (por. Constantinidou, Baker, 2002; Kumar, Priyadarshi, 2013; Rybak, 2019) – albo też z nierównomiernym tempem pogarszania się działania zmysłów wzroku i słuchu (por. Piekarzewska, Wieczorkowski i Zajenkowska-Kozłowska, 2016 – w Polsce jednak zauważa się odwrotną tendencję: znacząco więcej respondentów w podeszłym wieku wskazuje na problemy ze wzrokiem niż ze słuchem)?

Jak zostało wskazane powyżej, drugim z głównych uzasadnień dla narzucenia respondentom techniki w schemacie równoczesnym jest uzyskanie od nich informacji na temat preferencji w poprzednim badaniu. Przy założeniu wiarygodności takiej odpowiedzi i stabilności preferencji może to przynieść korzyści dla wskaźnika odpowiedzi i jakości pomiaru przy jednoczesnym uniknięciu omówionych problemów płynących z dania respondentom bezpośredniego wyboru techniki. Kontynuacją omówionego powyżej badania NASIS był eksperyment (Olson i in., 2012) przeprowadzony w 2009 roku na respondentach, którzy wskazali dalszą chęć uczestnictwa w badaniach. Kolejna runda została przez badaczki potraktowana jako grupa eksperymentalna z techniką CATI (która to technika była przez nie traktowana jako zgodna z preferencją wszystkich osób, które wyraziły chęć udziału w badaniu za pośrednictwem ankietera ze względu na bardzo niskie preferencje dla innych technik ankierskich). Dodatkowo wśród chętnych z badania z 2008 roku wylosowano cztery grupy – do ankiety pocztowej i internetowej oraz dwie do tych technik zastosowanych w sekwencji. Istotny wpływ na poziom wskaźnika odpowiedzi miała zgodność techniki, której używał respondent do udzielenia odpowiedzi, z poprzednio

zgłoszoną preferencją. Dla całego badania iloraz szans udziału w badaniu preferowaną drogą wynosił 1,38 w stosunku do niepreferowanej i był to wynik istotny statystycznie na poziomie $p < 0,001$. W rozbiciu na poszczególne grupy istotny statystycznie wynik uzyskano dla techniki telefonicznej i internetowej, przy czym ankieta internetowa stosowana samodzielnie na osobach ją preferujących miała niższy *response rate* niż wszystkie inne techniki dla grup ich niepreferujących.

Trzecim uzasadnieniem dla narzucenia respondentom techniki badania jest sytuacja łączenia operatorów, która może wynikać z chęci zmniejszenia błędu pokrycia przy korzystaniu z operatorów niekompletnych, lub chęć ograniczenia kosztu badania przez dodanie do droższego w wykorzystaniu operatoru kompletnego drugiego operatora niekompletnego. W przeszłości, szczególnie w warunkach amerykańskich (Groves, Lepkowski, 1985), przeprowadzano badania telefoniczne na próbie wylosowanej ze spisów abonenckich bądź przy użyciu metody *random digit dialing* (obie drogi mają swoje zalety i wady) przy uzupełnieniu jej o respondentów niemających dostępu do łącza telefonicznego dobieranych z próby adresowej. Taka praktyka wymaga dodatkowo uwzględniania nierównego prawdopodobieństwa wylosowania jednostek ze względu na ich potencjalną obecność w jednym bądź obu operatorach (Groves i in., 2004). Rozwiązanie takie stało się bardziej problematyczne wraz z rozwojem sieci komórkowej (dla której nie istnieją spójne operaty ani sposoby geograficznej identyfikacji abonenta przez numer kierunkowy w przypadku doboru metodą RDD), który spowodował znaczny spadek penetracji telefonii stacjonarnej (Brick, Cervantes, Lee, Norman, 2011). Także w Polsce nastąpił znaczący spadek łącz stacjonarnych kosztem komórkowych (Czapiński, Panek, 2011, s. 301). Nie omawiam jednak tej problematyki ze względu na jej specyficzny charakter. W przyszłości dynamiczny rozwój łączenia operatorów może mieć podstawy, jeśli pojawią się kompletne spisy użytkowników internetu – mogłoby to nastąpić np. w sytuacji zwiększania się presji na kontrolę przez państwo wykorzystania sieci przez obywateli i zmniejszenie niektórych społecznie uciążliwych występujących w niej zjawisk przez minimalizację bądź likwidację anonimowości użytkowników. Także wzrost udziału urządzeń mobilnych, które są zazwyczaj traktowane jako przedmioty bardzo osobiste i rzadko udostępniane innym użytkownikom, mógłby być podstawą do tworzenia spisu użytkowników ze względu na potencjalny obowiązek rejestracji i możliwość identyfikacji aktywności poszczególnych urządzeń.

Głównym uzasadnieniem dla stosowania drugiego, sekwencyjnego układu technik jest zazwyczaj chęć zmniejszenia kosztów badania poprzez oferowanie respondentom na początek technik tańszych, przy zachowaniu porównywalnego z badaniem *single-mode* poziomu realizacji. Koncepcja pochodzi od Josepha Hochstima (1967), który porównał badania sekwencyjne rozpoczynające się najdroższą techniką ankierską,

najtańszą telefoniczną i umiarkowanie drogą pocztową. Jego wyniki wskazywały na całkowitą porównywalność wszystkich układów pod względem poziomu realizacji i wyników oraz istotne różnice po stronie kosztów. Jako że w podstawowej formie wszystkim respondentom oferuje się jedną technikę (choć oczywiście istnieją bardziej złożone formy, łączące schemat równoczesny z sekwencyjnym i docierające do różnych kategorii badanych za pomocą różnych sekwencji technik), rozwiązanie to w mniejszym stopniu dotyczy problematyki *mode preference*. Odniesienie występuje w sytuacji założenia, że populacje wczesnych i późnych respondentów mogą różnić się od siebie pod względem cech społeczno-demograficznych, co skutkuje u nich preferencją różnych technik i z kolei owocuje ich udziałem w pierwszej bądź drugiej fazie badania. W tym rozumieniu celem badania sekwencyjnego jest dodatkowo zmniejszenie błędu braku odpowiedzi wynikającego ze skrzywienia populacji odpowiadających w stronę preferujących jedną technikę (tą użytą w badaniu *single mode*) – co w sytuacji związku cech i *mode preference* respondentów jest realnym zagrożeniem. Dillman wraz ze współpracownikami z Instytutu Gallupa (2009) przeprowadził badanie sekwencyjne na próbie 8999 amerykańskich respondentów, których dane teleadresowe zostały zakupione u prywatnego dostawcy. Podzielił on ich na pięć grup eksperymentalnych stosujących różne układy techniki pocztowej, internetowej, CATI i IVR. Dodatkowo kontrolowano efekt *recency* – preferowania przez respondentów ostatnich usłysanych odpowiedzi w badaniach technikami auralnymi – poprzez stosowanie różnej kolejności kafeterii. Przeciwnieństwem tego zjawiska jest *primacy effect*, czyli preferowanie pierwszych odpowiedzi w wynikach wizualnych. Wprowadzenie drugiej fazy badania spowodowało wzrost wskaźnika odpowiedzi od 22 do 37 p.p. – za wyjątkiem schematu poczta → CATI, który już w pierwszej fazie uzyskał 75-procentowy *response rate* i badanie telefoniczne podniosło go tylko o niecałe 8 p.p. Różnice w charakterystykach respondentów późnych i wczesnych na podstawie ich deklaracji były minimalne, za wyjątkiem zgodnej z wcześniej omówionymi badaniami nadreprezentacji mężczyzn w badaniach internetowych. Natomiast grupy respondentów i nierespondentów, których dane były dostępne na podstawie zakupionego operatu komercyjnego, istotnie różniły się od siebie. Jednakże nie można w pełni polegać na tej obserwacji, gdyż osoba rejestrująca się w komercyjnej bazie danych nie musiała być tą samą, która odpowiadała na pytania sondażowe. Można jednak przypuszczać, że za tym, który z członków gospodarstwa podaje swoje dane do rejestru, stoją jakieś zmienne ukryte determinujące jego charakterystykę również w innych wymiarach. Jedyna zmienna w operacie, która dotyczyła całego gospodarstwa – dochód na głowę – wykazała istotną różnicę między respondentami i nierespondentami jedynie

w schemacie IVR → CATI, czyli takim, który dotyczy podobnych technik, i to takich, w których respondent słyszy głos innej osoby

Układ trzeci dotyczy sytuacji stosowania kilku technik wobec jednego respondenta w trakcie jednego badania. Jego głównym celem jest zmniejszanie błędów wynikających z konieczności udzielenia odpowiedzi na pytania społecznie wrażliwe. Podstawowym rozwiązaniem tego problemu jest wspomniane wcześniej wręczenie przez ankietera badanemu kwestionariusza do samodzielnego wypełnienia dotyczącego części wrażliwej badania. Takie działanie wpływa pozytywnie na proces zbierania danych (de Leeuw i in., 2008, s. 306), możliwe są jednak metody bardziej wysublimowane, zapewniające respondentowi większe poczucie prywatności. Trzy główne z nich to eksperyment listy, eksperyment poparcia i metoda losowanych odpowiedzi (ang. *list experiment*, *endorsement experiment*, *randomized response*) (Rosenfeld i in., 2016). W pierwszej metodzie przedstawiamy dwóm losowo dobranym grupom respondentów listę wydarzeń czy opinii. Lista dla grupy eksperymentalnej zawiera element społecznie wrażliwy, a dla kontrolnej jest go pozbawiona. Prosimy respondenta o wskazanie jedynie określonej liczby opinii z listy, które podziela, albo wydarzeń, których doświadczył. Druga metoda stosuje losowe przydzielanie pytania o opinię na temat jakiegoś neutralnego działania z fragmentem mówiącym o poparciu go przez grupę, do której stosunek chcemy zbadać, bądź bez tego fragmentu. Trzecia polega na poproszeniu respondenta o wylosowanie (najczęściej za pomocą monety), na które z dwóch pytań odpowie lub też czy odpowie na pytanie wrażliwe, czy udzieli zdefiniowanej odpowiedzi – przy czym ankieter nie jest informowany o wyniku losowania. Analiza takich odpowiedzi możliwa jest za pomocą porównania rozkładów zmiennych dla poszczególnych grup przy użyciu modeli statystycznych. Zgodnie z wynikami przywołanego wyżej badania eksperymenty listy i poparcia najskuteczniej minimalizują względem sytuacji zadania bezpośrednio pytania wrażliwego wskaźnik braków odpowiedzi, natomiast metoda losowanych odpowiedzi najskuteczniej zmniejsza błąd pomiaru – jest jednak poznawczo najtrudniejsza dla respondentów.

Zmiana techniki w obrębie panelu jest jedynym zastosowaniem dla układu czwartego. W celu wzbudzenia początkowego zaufania respondentów pierwsza runda panelu może być przeprowadzona za pomocą techniki ankierskiej, podczas której zebrane mogą zostać informacje na temat możliwości kontaktu inną drogą (telefony, adresy e-mail). Przy zastosowaniu skutecznych metod motywowania respondentów do pozostania w panelu zmiana techniki całego badania może wiązać się w pierwszej kolejności ze znaczną oszczędnością. Istnieje jednak ryzyko, że błędy powodowane przez poszczególne techniki mogą się różnić i tym samym uniemożliwiać rozpoznanie, czy mamy do czynienia ze zmianą wartości badanej zmiennej w czasie, czy ze zmianą efektu techniki. Dodatkowo mamy

do czynienia z wynikającą z preferencji techniki możliwą niechęcią części członków panelu do jej zmiany i tym samym spadkiem ich liczby (por. Jäckle, Lynn i Burton, 2015). Nie będę omawiać układu (5) ze względu na jego wysoką problematyczność, a wręcz wątpliwości co do możliwości jego praktycznego zastosowania.

Ostatnim etapem badania, w którym możliwe jest mieszanie technik, są różne formy kontaktów ponagających i przypominających o udziale w nim. Ich celem jest skłonienie osób dotąd niechętnych do udzielenia odpowiedzi bądź też przypomnienie takiej możliwości tym, którzy o niej zapomnieli. Kontakty takie mogą również służyć zbieraniu pomocnych w szacowaniu rzeczywistego błędu braku odpowiedzi *paradata*, przykładowo przez przedstawienie jedynie metryczki do wypełnienia jednostkom niezgadzającym się na udział w całym badaniu czy dokonanie opisu fizycznego stanu miejsca zamieszkania nierespondenta przez ankietera. Wreszcie kontakty takie mogą pośredniczyć w przekazaniu zachęt materialnych do zmiany decyzji. Według cytowanych wyżej badań Portera i Whitcomba (2007), którzy zbierali odpowiedzi internetowo, testując skuteczność różnych technik kontaktów i ponagleń, nie stwierdzono różnicy między stosowaniem kontaktów przypominających drogą internetową i pocztową – przy czym ta druga technika istotnie zwiększała koszty badania, wskazane było więc użycie wyłącznie pierwszej, jeśli była taka możliwość techniczna.

Koncepcje przekładu pytań kwestionariusza między technikami

W sytuacji stosowania kilku technik przede wszystkim na etapie zbierania danych istnieje konieczność wzajemnego dostosowania konstrukcji kwestionariuszy/scenariuszy wywiadu telefonicznego pomiędzy technikami. Jak zauważają Franciszek Sztabiński i Żmijewska-Jędrzejczyk (2012), jeśli mamy do czynienia ze schematem sekwencyjnym, w którym pierwsza technika jest uznawana za główną, a kolejna/kolejne za pomocnicze, to właściwym będzie dostosowanie konstrukcji do techniki podstawowej. Przykładowo jeśli główną techniką jest CATI, a pomocniczą wywiad kwestionariuszowy, wskazane być może użycie krótkiej kafeterii, odpowiedniej do charakteru wywiadów telefonicznych, niewykorzystującej jednak w pełni potencjału techniki ankierskiej (de Leeuw i in., 2008, s. 311).

W sytuacji równorzędnego potraktowania wszystkich używanych technik możliwe są zasadniczo trzy podejścia. Po pierwsze, optymalizowanie kwestionariusza dla każdej techniki z osobna, tzw. *mode-specific design*. Przykładem będzie użycie pełnej kafeterii Likerta w technice ankierskiej i dwustopniowego zadawania pytań w CATI. Polega ono na zapytaniu

respondenta najpierw o to, czy zgadza się z jakimś twierdzeniem, czy nie, a następnie o to, w jakim stopniu. Takie podejście zwiększa ryzyko wystąpienia efektu techniki i utrudnia możliwość formułowania wniosków na temat różnic między podgrupami respondentów, jeśli ich dystrybucja pomiędzy poszczególne techniki nie była jednakowa.

Drugim podejściem jest zaproponowany przez Dillmana *unimode design*. Polega on na

(a) zachowaniu jednolitej struktury pytań dla poszczególnych technik (zmiana struktury może zmienić bodziec); (b) redukowaniu liczby możliwych odpowiedzi (winny być takie same i włączone w treść pytania, także odpowiedź „Trudno powiedzieć”); (c) ustaleniu ekwiwalentnych instrukcji i reguł przejścia (jeśli byłyby bardzo skomplikowane, należy rozważyć restrukturyzację kwestionariusza w celu ich uproszczenia)” (Sztabiński i in., 2012, s. 45–46).

Przykładem jest (skądinąd właściwe również z punktu widzenia minimalizowania tzw. satisficingu) zastosowanie zestawu pytań „tak/nie” zamiast „zaznacz wszystkie odpowiedzi” w pytaniach dotyczących udziału respondenta w wymienianych aktywnościach czy podzielania wymienionych opinii – formuła „zaznacz wszystkie” bowiem jest niemożliwa do skutecznego zastosowania w technice CATI, poleganie na zapamiętaniu przez badanego wszystkich wymienionych odpowiedzi znacząco godziłoby w jakość zbieranych danych (Dillman, 2007, s. 234).

Wreszcie ostatnim podejściem jest to dążące do ujednolicenia nie jakichś form konstrukcji narzędzia, ale ich kognitywnego obrazu. De Leeuw (2005, 2008), opierając się na wynikach analiz badaczy z nurtu CASM, głównie Norberta Schwarza, dotyczących wpływu modalności komunikacji z respondentem na interpretację pytań, wskazuje na konieczność rozwinięcia badań dotyczących tego obszaru. Najprostszym przykładem *general mode design*, jak nazwała go autorka, jest kwestia uzyskania od respondentów informacji o dacie jakiegoś wydarzenia w standaryzowanym formacie. Zastosowanie znacząco odmiennego sposobu konstrukcji tego pytania w technice CATI i internetowej dało lepszy rezultat niż ich ujednolicenie formalne, nawet szczegółowo zaznaczające, jak uzyskana informacja powinna wyglądać.

Przewidywanie preferencji techniki respondenta

Najciekawszym potencjalnym obszarem rozwoju badań dotyczących metodologii sondaży *mixed-mode* wydaje się problematyka przewidywania preferencji respondenta tak, by możliwe było spożytkowanie pozytywnego wpływu odpowiadania w technikach preferowanych na wskaźniki realizacji badania i jakość pomiaru. Ze względu na to, że jak ukazują cytowane powyżej badania, zaoferowanie respondentom wyboru okazuje się mieć skutek odwrotny do zamierzonego – spadek częstości zgód na udział w sondażu

– konieczne wydaje się przeprowadzenie owych przewidywań przed właściwą fazą badania. Jako że wiedza o konkretnych respondentach posiadana przez badaczy przed kontaktem z nimi ogranicza się zazwyczaj do informacji zawartych w operacie, pożyteczne wydaje się oparcie przewidywań na zmiennych, które można wyprowadzić z niego bezpośrednio lub pośrednio.

Do pierwszej grupy zaliczę podstawowe informacje dotyczące płci, wieku i miejsca zamieszkania, druga zaś, potencjalnie szersza, może zawierać zmienne środowiskowe powiązane z charakterystyką miejsca zamieszkania respondenta. Są to przykładowo dane dotyczące przestępczości, zamożności czy gęstości zaludnienia, o których na podstawie badań (Groves, Couper, 1998, s. 177; Sztabiński, Sztabiński i Przybysz, 2007, s. 29–31) możemy wnosić, że mają istotny wpływ na skłonność do udziału w sondażu, w tym za pośrednictwem konkretnej techniki. Drugą grupę zmiennych możemy uzyskać np. poprzez połączenie danych z operatu z danymi urzędowymi, bazując na kodzie pocztowym respondenta. Przewidywanie na podstawie powyższych zmiennych możliwe jest m.in. poprzez zastosowanie modeli regresji logistycznej, które opierać mogą się w tym wypadku przede wszystkim na danych behawioralnych i deklaracyjnych, a więc informacjach o tym, które kategorie respondentów wybierały udział w konkretnej technice w przeszłych sondażach *mixed-mode* w schemacie równoległym oraz jak poszczególne kategorie respondentów odpowiadały na pytania o preferowaną technikę w sondażach wszelkiego typu. Należy zaznaczyć, że powyższe analizy należałoby wykonywać osobno dla obszarów (krajów) różniących się od siebie zarówno pod względem kulturowym, jak i jakości infrastruktury różnego typu. Należy bowiem przypuszczać, że na preferencję techniki wpływają np. takie czynniki jak gościnność czy jakość funkcjonowania usług pocztowych.

Jak skorzystać z wyników proponowanych przewidywań? Strategią niosącą największe korzyści zdaje się wyróżnienie na podstawie wartości przewidywanych przez modele prawdopodobieństw preferencji konkretnych technik tych grup respondentów, dla których przyjmują one wartości największe lub przekraczające arbitralnie ustalone minimum. Optymalna wartość owego minimum musi być ustalona metodą prób i błędów. Jako że technika wywiadu kwestionariuszowego jest *ceteris paribus* najskuteczniejsza w skłanianiu potencjalnych respondentów do współpracy (ale jednocześnie generuje największe koszty realizacji oraz pogarsza jakość pomiaru przez efekt ankierski), respondenci o niejasnych preferencjach powinni być przyporządkowani do tej właśnie techniki. W dalszej kolejności kierowano by do poszczególnych kategorii osób prośby o udział w badaniu od razu w technice, której preferencja byłaby najbardziej prawdopodobna. Jako podstawowego efektu przyjęcia powyższej strategii należałoby spodziewać się zmniejszenia kosztów realizacji badania przy zachowaniu wskaźników realizacji zbliżonych do sondaży *single-mode* wykonywanych przy użyciu techniki bezpośredniego wywiadu standaryzowanego.

Bibliografia

- Brick, J.M., Cervantes, I.F., Lee, S., Norman, G. (2011). Nonsampling Errors in Dual Frame Telephone Surveys. *Survey Methodology*, 37(1), 1–12.
- Cialdini, R. (1984). *Influence: The New Psychology of Modern Persuasion*. New York: Quill.
- Constantinidou, F., Baker, S. (2002). Stimulus Modality and Verbal Learning Performance in Normal Aging. *Brain and Language*, 82(3), 296–311.
- Couper, M.P. (2011). The Future of Modes of Data Collection. *Public Opinion Quarterly*, 75(5), 889–908.
- Czapiński, J., Panek, T. (2011). Social Diagnosis 2011: Objective and Subjective Quality of Life in Poland. *Contemporary Economics*, 5(3), 1.
- de Leeuw, E.D. (2005). To Mix or Not to Mix Data Collection Modes in Surveys. *Journal of Official Statistics*, 21(2), 233–255.
- de Leeuw, E.D. (2008). Choosing the Method of Data Collection (s. 113–135). W: E.D. de Leeuw, J.J. Hox, D.A. Dillman (eds.), *International Handbook of Survey Methodology*. New York: L. Erlbaum Associates.
- de Leeuw, E.D., Hox, J.J., Dillman, D.A. (2008). Mixed-mode Surveys: When and Why (s. 299–316). W: E.D. de Leeuw, J.J. Hox, D.A. Dillman (eds.), *International Handbook of Survey Methodology*. New York: L. Erlbaum Associates.
- Dillman, D.A. (2007). *Mail and Internet Surveys: The Tailored Design Method*. Hoboken: Wiley.
- Dillman, D.A., Smyth, J.D., Christian, L.M. (2014). *Internet, Phone, Mail, and Mixed Mode Surveys: The Tailored Design Method (4th ed.)*. Hoboken: John Wiley Sons Inc.
- Dillman, D.A., Phelps, G., Tortora, R., Swift, K., Kohrell, J., Berck, J., Messer, B.L. (2009). Response Rate and Measurement Differences in Mixed-mode Surveys Using Mail, Telephone, Interactive Voice Response (IVR) and the Internet. *Social Science Research*, 38(1), 1–18.
- Diment, K., Garrett-Jones, S. (2007). How Demographic Characteristics Affect Mode Preference in a Postal/Web Mixed-Mode Survey of Australian Researchers. *Social Science Computer Review*, 25(3), 410–417.
- Edwards, M.L., Dillman, D.A., Smyth, J.D. (2014). An Experimental Test of the Effects of Survey Sponsorship on Internet and Mail Survey Response. *Public Opinion Quarterly*, 78(3), 734–750.
- Groves, R. (2006). Nonresponse Rates and Nonresponse Bias in Household Surveys: What Do We Know About the Linkage Between Nonresponse Rates and Nonresponse Bias? *Public Opinion Quarterly*, 70(5), 646–675.
- Groves, R.M. (2011). Three Eras of Survey Research. *Public Opinion Quarterly*, 75(5), 861–871.
- Groves, R.M., Couper, M.P. (1998). *Nonresponse in Household Interview Surveys*. New York: John Wiley Sons.
- Groves, R.M., Kahn, R.L. (1979). *Surveys by Telephone: A National Comparison with Personal Interviews*. New York: Academic Press.
- Groves, R.M., Lepkowski, J.M. (1985). Dual Frame, Mixed Mode Survey Designs. *Journal of Official Statistics*, 1(3), 263–286.
- Groves, R.M., Singer, E., Corning, A. (2000). Leverage-Saliency Theory of Survey Participation. *Public Opinion Quarterly*, 64(3), 299–308.
- Groves, R.M., Fowler Jr., F.J., Couper, M.P., Lepkowski, J.M., Singer, E., Tourangeau, R. (2004). *Survey Methodology (t. 2)*. Hoboken: Wiley.
- Grzeszkiewicz-Radulska, K. (2009). *Respondenci niedostępni w badaniach sondażowych*. Łódź: Wydawnictwo UŁ.
- Haan, M., Ongena, Y.P., Aarts, K. (2014). Reaching Hard-to-survey Populations: Mode Choice and Mode Preference. *Journal of Official Statistics*, 30(2), 355–379.
- Hochstim, J.R. (1967). A Critical Comparison of Three Strategies of Collecting Data from Households. *Journal of the American Statistical Association*, 62(319), 976–989.

- Jabkowski, P. (2015). *Reprezentatywność badań reprezentatywnych: analiza wybranych problemów metodologicznych oraz praktycznych w paradygmacie całkowitego błędu pomiaru*. Poznań: Wydawnictwo Naukowe UAM.
- Jäckle, A., Lynn, P., Burton, J. (2015). Going Online with a Face-to-Face Household Panel: Effects of a Mixed Mode Design on Item and Unit Non-response. *Survey Research Methods*, 9(1), 57–70.
- Kumar, N., Priyadarshi, B. (2013). Differential Effect of Aging on Verbal and Visuo-spatial Working Memory. *Aging and Disease*, 4(4), 170–177.
- Medway, R.L., Fulton, J. (2012). When More Gets You Less: A Meta-analysis of the Effect of Concurrent Web Options on Mail Survey Response Rates. *Public Opinion Quarterly*, 76(4), 733–746.
- Millar, M.M., Dillman, D.A. (2011). Improving Response to Web and Mixed-mode Surveys. *Public Opinion Quarterly*, 75(2), 249–269.
- Millar, M.M., O'Neill, A.C., Dillman, D.A. (2009). *Are Mode Preferences Real?*. Pullman, WA: Social and Economic Sciences Research Center at Washington State University.
- Olson, K., Smyth, J.D., Wood, H.M. (2012). Does Giving People Their Preferred Survey Mode Actually Increase Survey Participation Rates? An Experimental Examination. *Public Opinion Quarterly*, 76(4), 611–635.
- Piekarzewska, M., Wieczorkowski, R., Zajenowska-Kozłowska, A. (2016). *Stan zdrowia ludności Polski w 2014 r.*, Warszawa: GUS.
- Porter, S.R., Whitcomb, M.E. (2007). Mixed-mode Contacts in Web Surveys. *Public Opinion Quarterly*, 71(4), 635–648.
- Rosenfeld, B., Imai, K., Shapiro, J.N. (2016). An Empirical Validation Study of Popular Survey Methodologies for Sensitive Questions. *American Journal of Political Science*, 60(3), 783–802.
- Rossi, P.H., Wright, J.D., Anderson, A.B. (1983). *Handbook of Survey Research*. New York, N.Y.–London: Academic Press.
- Rybak, A. (2019). Komunikacja oparta na kanale wizualnym i audialnym a jakość pomiaru sondażowego. W stronę nowego ujęcia hipotezy preferencji techniki przy wykorzystaniu koncepcji pamięci roboczej. *Przegląd Socjologiczny*, 68(3), 207–227.
- Sakshaug, J.W., Yan, T., Tourangeau, R. (2010). Nonresponse Error, Measurement Error, and Mode of Data Collection: Tradeoffs in a Multi-mode Survey of Sensitive and Non-sensitive Items. *Public Opinion Quarterly*, 74(5), 907–933.
- Smyth, J.D., Olson, K., Kasabian, A. (2014). The Effect of Answering in a Preferred Versus a Non-Preferred Survey Mode on Measurement. *Survey Research Methods*, 8(3), 137–152.
- Smyth, J.D., Olson, K., Millar, M.M. (2014). Identifying Predictors of Survey Mode Preference. *Social Science Research*, 48, 135–144.
- Sztabiński, F., Żmijewska-Jędrzejczyk, T. (2012). Mixed Mode Survey Design: problem efektu techniki. *Przegląd Socjologiczny*, 61(1), 31–63.
- Sztabiński, F., Sztabiński, P.B., Przybysz, D. (2007). Are Non-Respondents Similar to Respondents? Findings from the ESS-2004 in Poland. *ASK. Społeczeństwo, Badania, Metody*, 16(16), 25–54.
- Tourangeau, R., Couper, M.P., Steiger, D.M. (2003). Humanizing Self-administered Surveys: Experiments on Social Presence in Web and IVR Surveys. *Computers in Human Behavior*, 19(1), 1–24.
- Tourangeau, R., Rips, L.J., Rasinski, K.A. (2000). *The Psychology of Survey Response*. Cambridge: Cambridge University Press.
- Verma, S.K., Courtney, T.K., Lombardi, D.A., Chang, W.R., Huang, Y.H., Brennan, M.J., Perry, M.J. (2014). Internet and Telephonic IVR Mixed-mode Survey for Longitudinal Studies: Choice, Retention, and Data Equivalency. *Annals of Epidemiology*, 24(1), 72–74.

Mode preference hypothesis in the context of mixed-mode survey design

SUMMARY

The aim of this chapter is to show the state of knowledge in the field of mixed-mode survey design methodology, the known theoretical and practical issues in its use, and to highlight knowledge gaps on this subject. Inclusion of survey mode preference hypothesis in this context is particularly interesting because of possible positive effects of its proper implementation on increasing response rates and measurement quality and lowering survey costs. However, this phenomenon is still not deeply understood and lacks a well-developed theoretical context. This chapter connects it with a wider range of theories of survey participation and gathers the elements of possible future theory. In conclusion, it presents a possible strategy for the utilization of knowledge about participants' preferences concerning a mode in survey research.

Keywords: mixed-mode survey, representative research, survey mode preference, survey research methodology