

ARTYKUŁY

Jean-Paul Grémy*

WPŁYW SPOSOBU FORMUŁOWANIA PYTAŃ KWESTIONARIUSZOWYCH NA UZYSKIWANE ODPOWIEDZI¹

Począwszy od 1987 roku przeprowadzono we Francji serię eksperymentów split-ballot dotyczących wpływu formułowania pytań na wyniki badań surveyowych. Główne wnioski płynące z tych eksperymentów dotyczą takich problemów, jak umieszczenie kategorii odpowiedzi będącej w środku skali, włączenie odpowiedzi niezdecydowanych („trudno powiedzieć”) do kategorii odpowiedzi dostępnych respondentowi, intonacji przy zadawaniu pytań oraz kolejności pytań w kwestionariuszu wywiadu. Analizy pokazują, w jakim stopniu wymienione czynniki różnicują odpowiedzi udzielane przez respondentów w trakcie wywiadu.

Problem znaczenia odpowiedzi uzyskiwanych w sondażach jest przedmiotem empirycznych badań od ponad pięćdziesięciu lat; w Stanach Zjednoczonych Amerykański Instytut Badań Opinii Publicznej (Instytut Gallupa) podjął pierwsze analizy dotyczące sformułowań pytań już w roku 1936 (Cantril i Strunk 1951: 597, nr 3 i 598, nr 8); we Francji podobne analizy zostały przeprowadzone po II wojnie przez Jeana Stoetzela we Francuskim Instytucie Opinii Publicznej (IFOP) i przez Alaina Girarda w Narodowym Instytucie Studiów Demograficznych (INED) (Grémy 1987).

Zainteresowanie takimi badaniami wydaje się obecnie wzrastać pod wpływem dwóch czynników. Z jednej strony, kumulacja wiedzy w dziedzinie zachowań, wyobrażeń, przekonań i postaw prowadzi do zapotrzebowania na informacje specyficzne, a więc bardziej szczegółowe i bardziej precyzyjne. Wynika z tego konieczność zwiększenia subtelności technik zbierania materiałów. Z drugiej strony, postępy w zakresie analizy danych wielowymiarowych pociągają za sobą wzrost wymagań co do jakości analizowanych informacji zgodnie z przysłowiem GIGO (*Garbage in, Garbage out*). Wynika stąd konieczność wzrostu rzetelności odpowiedzi poprzez lepsze opanowanie technik ankietowych. Można zaobserwować zresztą podejmowanie tego typu badań we Francji od dwóch-trzech lat, w szczegól-

* Uwagi do autora lub prośby o nadbitki prosimy kierować do: Jean-Paul Grémy, Institut des Hautes Études de la Sécurité Intérieure, 19, rue Pecllet, 75015 Paris, France.

¹ Rozdział książki L. Lebert (red.), *La qualité de l'information dans les enquêtes*, Dunod: Paris 1992. Jean-Paul Grémy jest profesorem socjologii na Uniwersytecie René Descartes, Paryż V.

ności przez Instytut sondażowy Lavialle (Roussel i Mouthon 1988; Grémy 1989; Vincent 1990).

Badania metodologiczne nad wywiadem kwestionariuszowym można podzielić na dwie zasadnicze kategorie w zależności od tego, czy dotyczą one wpływu warunków jego przeprowadzania (kontekst fizyczny i psychologiczny wywiadu, charakterystyki ankietera i respondenta, specyficzne procedury każdego z instytutów sondażowych), czy też efektów wprowadzonych przez kwestionariusz (sposób jego zastosowania, miejsce i sformułowania pytań). W książce Seymoura Sudmana i Normana M. Bradburna (1974) znajduje się syntetyczna prezentacja najważniejszych badań na ten temat zrealizowanych przed rokiem 1974.

Wszystkie badania tego typu oparte są na hipotezie, że ta sama osoba, pytana na ten sam temat w tym samym momencie, może udzielać różnych odpowiedzi w zależności od kontekstu ankiety i tekstu kwestionariusza. Nie jest oczywiście możliwe bezpośrednie sprawdzenie tej hipotezy. Badania empiryczne dotyczące znaczenia pytań zadawanych w kwestionariuszach odwołują się generalnie do jednej z dwóch następujących metod: 1. Podejście nastawione na rozumienie, w którym pyta się respondenta o rozumienie pytania zamkniętego i powody, dla których udzielił określonej odpowiedzi (bądź przy okazji wywiadu zasadniczego (Cantril 1944), bądź też podczas wywiadu pogłębionego realizowanego nazajutrz po wywiadzie zasadniczym (Belson 1981); 2. Podejście eksperymentalne, które polega na testowaniu, podczas tego samego wywiadu, kilku wersji pytań lub kwestionariusza na kilku identycznych pod próbkach (metoda „dzielonych prób” lub *split ballot*).

Wnioski, które można sformułować na podstawie eksperymentów w zakresie formułowania pytań ankietowych są w znacznym stopniu związane z używanym językiem i kontekstem kulturowym, który mu odpowiada; tak więc prosta transpozycja wniosków z eksperymentów prowadzonych w innych krajach (w szczególności w Stanach Zjednoczonych) nie zawsze jest możliwa. Poniżej przedstawiamy kilka rezultatów eksperymentów zrealizowanych ostatnio we Francji, według metody „dzielonych prób”. Badania te dotyczą odpowiedzi „środkowej”, odpowiedzi „nie mam zdania”, porządku prezentacji możliwych odpowiedzi, wyboru używanych słów i miejsca pytań w kwestionariuszu.

PROBLEM ODPOWIEDZI „ŚRODKOWEJ”

Specjaliści w zakresie przygotowania ankiet są podzieleni jeżeli chodzi o celowość przedstawiania respondentom odpowiedzi „środkowej” (*logical, middle alternative*), takich jak na przykład „w takim samym stopniu jeden, jak i drugi”, „między jednym a drugim”, „bez różnicy”. Stanley Payne, autor pierwszej pracy dotyczącej przygotowania kwestionariusza ankiety, odradzał taką praktykę oceniając że prowadzi ona do utraty informacji (Payne 1951). Converse i Presser są tego samego zdania, ale proponują dodanie pytania mierzącego intensywność postawy lub wyrażanej opinii (Converse i Presser 1986). Natomiast Sudman i Bradburn zalecają włączenie takiej odpowiedzi w przypadku alternatyw, a uznają to za niezbędne w przypadku skal zawierających oceny (Sudman i Bradburn 1983). Jak zauważył Hubert Jacquart, obecność lub brak odpowiedzi „środkowej” powoduje powstanie dwóch pytań nie ekwiwalentnych, z których każde ma swoje zalety (Jacquart 1988).

Instytut sondażowy Lavalie (ISL) zrealizował, od lipca 1988 do lutego 1990, pierwszą serię eksperymentów dotyczących łącznie odpowiedzi „średkowej” i odpowiedzi „nie mam zdania”. Jeżeli chodzi o odpowiedź „średkową”, zostały zadane równolegle dwa pytania (w lipcu 1988) na dwóch podpróbach, z których każda liczyła 1000 osób. Pierwszej zaproponowano prostą alternatywę: „Czy Pana(i) zdaniem rząd powinien zwracać szkołom niepaństwowym więcej pieniędzy niż obecnie czy mniej pieniędzy niż obecnie?”; drugiej przedstawiono dodatkowo możliwość odpowiedzi pośredniej („tyle samo pieniędzy co obecnie) zamieszczonej między dwiema odpowiedziami ekstremalnymi. Ankieterzy mieli obowiązek zanotować brak zdania, jeśli taka odpowiedź została udzielona spontanicznie przez osobę badaną (należącą do pierwszej próby).

Tabela 1 przedstawia rezultaty tego eksperymentu. (Jak w większości zamieszczonych tutaj tabel, efekt różnicujący przedstawianych pytań został określony jako różnica procentów w drugiej kolumnie i w pierwszej).

Tabela 1.
Odpowiedź „średkowa” (zestawienie wyników)

Odpowiedź	Bez odpowiedzi „średkowej”	Z odpowiedzią „średkową”	Różnica
Więcej	47,9%	33,5%	-14,4%
Tyle samo	0,0%	32,3%	+32,3%
Mniej	18,7%	13,5%	-5,2%
(Brak zdania)	33,4%	20,7%	-12,7%

Z tabeli tej wynikają dwa wnioski: 1. Obecność odpowiedzi „średkowej” zmniejsza liczbę „braków opinii”. Chodzi o zjawisko najczęściej obserwowane w podobnych eksperymentach prowadzonych w krajach anglo-saksońskich, nawet jeśli występują pewne wyjątki od tej reguły (Schuman i Presser 1981). 2. Dodanie odpowiedzi „tyle samo pieniędzy” nie zmienia rozkładu odpowiedzi „znaczących” (*substantive answers*). Innymi słowy, jeżeli bierze się pod uwagę tylko odpowiedzi „więcej” i „mniej” otrzymuje się podobne odsetki niezależnie od sformułowania (Tabela 2). Taki rezultat obserwuje się praktycznie zawsze (Schuman i Presser 1981).

Tabela 2.
Odpowiedź „średkowa” (tylko odpowiedzi „znaczące”)

Odpowiedź	Bez odpowiedzi „średkowej”	Z odpowiedzią „średkową”	Różnica
Więcej	479=71,9%	335=71,3%	-0,6%
Mniej	187=28,1%	135=28,7%	+0,6%

Dodanie odpowiedzi „średkowej” powoduje więc zmniejszenie „braków zdania”, jednak bez modyfikowania stosunku między opiniami ekstremalnymi. Schuman i Presser przedstawiają hipotezę, że respondenci, którzy wybierają odpowiedź „średkową”, jeśli jest ona przedstawiona, mają mniej ugruntowane przekonania lub uważają, że problematyka pytania w mniejszym stopniu ich dotyczy. Wynika stąd wniosek, żeby nie przedstawiać odpowiedzi „średkowej”, lecz raczej dodać pytania dodatkowe o siłę przekonania respondenta (Converse i Presser 1986).

ODPOWIEDŹ „NIE MAM ZDANIA”

Wśród anglosaskich specjalistów od budowy kwestionariusza istnieje niemal jednomyślność w ocenie, że odpowiedź „nie mam zdania” („brak opinii”, „nie wiem”) powinna być *explicit*e przedstawiana, z wyjątkiem niektórych specyficznych przypadków (np. zastosowania karty respondenta, zawierającej listę możliwych odpowiedzi) (Payne 1951). Wyraźne przedstawienie takiej możliwości odpowiedzi zmniejsza, jak to się przyjmuje, niepokój osoby badanej (zwłaszcza w pytaniach o wiedzę) (Sudman i Bradburn 1983). Dlatego też jest to zalecane zwłaszcza na początku wywiadu (Cantril i Strunk 1951). Stanley Payne już w roku 1951 snuł rozważania, że „jeśli osoby odpowiadające nie mają solidnie ugruntowanych opinii, ale muszą dla udzielenia odpowiedzi dokonywać operacji w rodzaju gry w orła i reszkę, lepiej jest w ogóle nie zbierać takich odpowiedzi” (Payne 1951, s. 23), dodając, w związku z pytaniami sformułowanymi w sposób niejasny, że „odpowiedź »tak«, która znaczy »nie« jest gorsza, niż »nie wiem«” (s. 234).

Podczas serii wspomnianych już eksperymentów ISL porównywano trzy sformułowania: 1. Odpowiedź „nie mam zdania”, której nie proponowano *explicit*e, ale która była odnotowywana, jeśli pojawiła się spontanicznie; 2. Odpowiedź „nie mam zdania” przedstawiana na końcu możliwych odpowiedzi; 3. Odpowiedź „nie mam zdania” przedstawiana w postaci pytania filtrującego. Ponieważ w przypadku każdej ankiety dysponowano jedynie dwiema tysiąc osobowymi podpróbami, nie było możliwe przeprowadzenie jednocześnie trzech porównań; zdecydowano się więc uznać pierwszą z form za podstawową i porównywać z nią, podczas przeprowadzania dwóch następnych ankiet, każdą z dwóch pozostałych form (tę samą regułę przyjęto również w pierwszej serii eksperymentów ISL, do których odwoływaaliśmy się wcześniej).

W konsekwencji sformułowanie podstawowe było więc następujące: „Czy Pana(i) zdaniem rząd powinien wypłacać szkołom niepublicznym więcej pieniędzy niż obecnie, tyle samo co obecnie czy też mniej pieniędzy niż obecnie?” W październiku 1988 sprawdzono efekty dodania do tego pytania odpowiedzi „czy też nie ma Pan(i) zdania w tej sprawie?” (tabela 3).

Tabela 3.
„Nie mam zdania” na końcu pytania (zestawienie rezultatów)

Odpowiedź	Bez „nie mam zdania”	„Nie mam zdania” na końcu	Różnica
Więcej	35,4%	29,3%	-6,1%
Tyle samo	25,5%	23,0%	-2,5%
Mniej	13,5%	14,4%	+0,9%
Brak zdania	25,5%	33,3%	+7,8%

Jak można było oczekiwać, sformułowanie takie spowodowało zwiększenie liczby „braków zdania” (Schuman i Presser 1981). Jednakże wzrost ten jest relatywnie niski i występuje głównie kosztem odpowiedzi „więcej pieniędzy”.

Jeśli możliwość „nie mam zdania” przedstawia się w postaci pytania filtrującego, jej efekt jest naturalnie znacznie bardziej wyraźny. W maju 1989 sprawdzono następujące sformułowanie: „Czy w sprawie pieniędzy, które rząd wypłaca szko-

tom niepublicznym, ma Pan(i) wyrobione zdanie?”. Jeśli tak: „Czy Pana(i) zdaniem rząd powinien wypłacać więcej pieniędzy...?”. To nowe sformułowanie spowodowało, że liczba „braków zdania” wzrosła ponad dwukrotnie w stosunku do pytania podstawowego (por. tabela 4).

Powyższy wzrost udziału „braków zdania” jest w pełni porównywalny ze stwierdzonym przez Schumana i Pressera podczas podobnych eksperymentów (Schuman i Presser 1981). Jednakże autorzy ci nie stwierdzili w większości swoich eksperymentów wpływu pytania-filtru na odpowiedzi „znaczące”. Tabela 5 wskazuje na występowanie takiego wpływu, gdy włącza się lub nie odpowiedź „średniową” do odpowiedzi „znaczących”, a dzieje się to każdorazowo kosztem odpowiedzi „więcej pieniędzy”.

Tabela 4.
„Nie mam zdania” jako pytanie filtrujące (zestawienie rezultatów)

Odpowiedź	Bez „nie mam zdania”	„Nie mam zdania” jako filtr	Różnica
Więcej	33,7%	18,1%	-15,6%
Tyle samo	31,5%	20,6%	-10,9%
Mniej	12,9%	10,2%	- 2,7%
Brak zdania	21,9%	51,1%	+29,2%

Tabela 5.
Odpowiedź „nie mam zdania”.
Różnice między odpowiedziami na pytania eksperymentalne i pytanie podstawowe
(tylko odpowiedzi „znaczące”)

Odpowiedź	„Nie mam zdania” na końcu	„Nie mam zdania” jako filtr
Więcej	43,9–47,6=-3,7%	37,0–43,1=-6,1%
Tyle samo	34,5–34,3=+0,2%	42,1–40,3=+1,8%
Mniej	21,6–18,2=+3,4%	20,8–16,5=+4,3%
Więcej	67,1–72,4=-5,3%	64,0–72,3=-8,3%
Mniej	32,9–27,6=+5,3%	36,0–27,7=+8,3%

Tak więc, jak to stwierdzili Schuman i Presser, pytanie filtrujące zwiększa liczbę „braków zdania”, przy czym nie wydaje się to być związane ani z treścią pytania, ani też z udziałem spontanicznych „braków zdania” odpowiadających niefiltrującej formie pytania. Autorzy ci podjęli próbę analizy przyczyn procesu „wahań” (*floating process*) w odpowiedziach. Definiują oni respondentów wahających się (*floaters*) jako osoby, które udzielają odpowiedzi „znaczącej”, jeśli odpowiedź „nie mam zdania” nie została przedstawiona *explicite*, ale wybierają tę ostatnią, jeśli jest ona przedstawiona w postaci filtru. Wysuwają oni hipotezę, zgodnie z którą „respondenci wahający się” nie różnią się od innych specjalnymi cechami psychologicznymi, ale raczej charakterystykami socjologicznymi, związanymi z niskim poziomem wykształcenia: brakiem znajomości pojęć z zakresu polityki, brakiem wyrazistych opinii *etc.* Respondenci ci dysponują jednakże wyobrażeniami politycznymi

mi lub dotyczącymi generalnie świata, co pozwala im ewentualnie wypowiadać się na tematy przez nich ignorowane na co dzień (Schuman i Presser 1981). Konkluzje te należałoby skonfrontować z uzyskanymi w wyniku zastosowania innej metody; chodzi o badania Guya Michelata i Michela Simona na temat „braków zdania” w pytaniach politycznych (Michelat i Simon 1985).

Należałoby jeszcze przeanalizować przesunięcia między odpowiedzią „nie mam zdania” a odpowiedziami „znaczącymi”. Erozja odpowiedzi „więcej pieniędzy” jest silniejsza, im wyższy jest udział „braków zdania”. Można wysunąć hipotezę, że osoby, dla których problematyka ta jest mało ważna, chcą być możliwie najbardziej akceptowane społecznie i dlatego też wybierają odpowiedź „więcej pieniędzy”, to jest tę, która wydaje się im być najbardziej społecznie pożądana (*social desirability bias*). Z drugiej strony, odpowiedź „średkowa” podlega erozji w tym samym sensie, choć w mniejszym zakresie. Schuman i Presser również stwierdzili, że przedstawienie odpowiedzi „średkowej” powoduje najczęściej pewne zmniejszenie udziału „braków zdania”. Dodają jednakże, iż „właściwa interpretacja różnic udziałów „braków zdania” w zależności od formy pytania nie musi ograniczać się do stwierdzenia, że przedstawiona *explicite* odpowiedź „średkowa” „przyciąga” osoby, które nie mają opinii. Jest również możliwe, że pominięcie odpowiedzi „średkowej” zwiększa udział „braków zdania” zmuszając osoby, które rzeczywiście mają taką opinię, do stwierdzenia, iż nie mają zdania, gdyż nie mogą wybrać żadnej z proponowanych odpowiedzi” (Schuman i Presser 1981: 170).

KOLEJNOŚĆ PREZENTACJI ODPOWIEDZI

Porządek, w jakim przedstawiane są respondentom możliwe odpowiedzi na dane pytanie, może nie wywierać żadnego wpływu na ich rozkład. Podczas pierwszej serii eksperymentów realizowanych przez ISL sprawdzaliśmy (w kwietniu 1988) skutki umieszczenia odpowiedzi „tyle samo pieniędzy” na końcu pytania. Oczekiwano, że skutkiem będzie zwiększenie udziału tej właśnie odpowiedzi „średkowej”; efekt taki jednak nie wystąpił (por. tabela 6).

Tabela 6.
Miejsce odpowiedzi „średkowej”

Odpowiedź	Odpowiedź „średkowa” w środku	Odpowiedź „średkowa” na końcu	Różnica
Więcej	37,6%	34,9%	-2,7%
Tyle samo	32,0%	29,2%	-2,8%
Mniej	16,3%	19,3%	+3,0%
Brak zdania	16,9%	19,3%	+2,4%
Więcej	45,2%	43,3%	-1,9%
Tyle samo	38,5%	36,2%	-2,3%
Mniej	16,4%	19,6%	+3,2%

W przypadku zadania tego pytania ustnie, bez „wsparcia” pisemnego w postaci karty respondentanta, należałoby przyjąć wystąpienie tzw. „odpowiedzi echo” (*recency effect*) polegającej na tym, że respondenci wybierają przede wszystkim ostatnią z przedstawionych odpowiedzi (Payne 1951; Schuman i Presser 1981).

Efekt przeciwny, *primacy effect*, polega na tendencji respondentów do wybierania przede wszystkim pierwszej lub pierwszych spośród przedstawionych odpowiedzi. Pośredniej ilustracji tej tendencji dostarczyło porównanie dwóch eksperymentów zrealizowanych przez ISL: pierwszy z nich (październik 1988) został przedstawiony powyżej (tabela 3: „Brak zdania” na końcu pytania); w drugim (luty 1989) z formą podstawową pytania porównano sformułowania, w którym odpowiedź „nie mam zdania” była przedstawiona na początku pytania („Jeżeli chodzi o pieniądze, które rząd wypłaca szkołom niepublicznym, to czy nie ma Pan(i) opinii w tej sprawie, czy też uważa Pan(i), że rząd (...)?”).

Tabela 7.
Miejsce odpowiedzi „nie mam zdania”. Różnice między odpowiedziami na pytania eksperymentalne i na pytanie podstawowe

Odpowiedź	„Nie mam zdania” na końcu	„Nie mam zdania” na początku
Więcej	29,3–35,4=–6,1%	29,3–33,0=–3,7%
Tyle samo	23,0–25,5=–2,5%	25,8–29,3=–3,5%
Mniej	14,4–13,5=+0,9%	8,9–14,1=–5,2%
Brak zdania	33,3–25,5=+7,8%	36,0–23,5=+12,5%

Porównanie wyników tych eksperymentów (tabela 7) wskazuje, że odpowiedź „nie mam zdania” była częściej wybierana, jeśli została przedstawiona na początku pytania. Podobny rezultat zaobserwowano przy okazji nieco innego eksperymentu. Podczas wystawy „Ryzyko różnych zawodów (8–12 kwietnia 1987) na stoisku CNRS wręczano młodym zwiedzającym krótką ankietę do samodzielnego wypełnienia, opracowaną przez Laboratorium Analiz Wtórnych i Metod Stosowanych w Socjologii (LASMAS). Ankieta ta została przygotowana w dwóch wersjach, różniących się porządkiem, w jakim przedstawiono możliwe odpowiedzi na następujące pytanie: „Jakie są, według Pana(i), dwa obecnie najważniejsze problemy?: pierwszy pod względem ważności, drugi pod względem ważności”. W pierwszej wersji odpowiedzi przedstawiono w takim porządku, jak to widnieje w tabeli 8, w drugiej zaś, w porządku odwrotnym. Na pytanie to odpowiedziało 2180 dorastającej młodzieży obydwu płci (1186 na pierwszą jego wersję, 994 – na drugą).

Tabela 8.
Efekt porządku prezentacji odpowiedzi (na pytanie o najważniejszy obecnie problem)

Odpowiedź	Porządek	Porządek	Różnica
Bezrobocie	34,2%	19,2%	–15,0%
Terroryzm	8,5%	7,0%	–1,5%

c.d. tabeli 8.

Głód na świecie	21,3%	17,7%	-3,6%
Wojna	15,3%	18,3%	+3,0%
Przeludnienie	1,5%	1,7%	+0,2%
Rasizm	5,0%	8,5%	+3,5%
Nierespektowanie praw człowieka	8,0%	10,4%	+2,4%
Niedostateczne przygotowanie zawodowe młodzieży	1,3%	6,2%	+4,9%
Przestępczość	1,2%	8,2%	+7,0%

Tabela 8 zawiera nieważone odsetki odpowiedzi dotyczące problemu uważanego za najważniejszy. (Dystrybucja dwóch wersji ankiety nie odbywała się w sposób przypadkowy i wypróbowywano różne metody ważenia; nie modyfikują one jednak ogólnego rozkładu rezultatów).

Daje się zauważyć występowanie dość wyraźnego efektu porządku prezentacji: odpowiedzi wybierano częściej wówczas, gdy przedstawione były na początku listy, niż gdy zamieszczono je na końcu. Należy jednak zwrócić uwagę, że w tym przypadku trzy problemy uznane za najważniejsze przez większość respondentów są takie same niezależnie od porządku prezentacji.

Schuman i Presser zrealizowali dość podobny eksperyment (z listy problemów, przed którymi stoją Stany Zjednoczone, należało wybrać problem najważniejszy), w którym przedstawiano odpowiedzi w pięciu różnych porządkach. Nie zaobserwowano żadnego systematycznego efektu. Jeśli zrobi się bilans innych eksperymentów dotyczących tej problematyki wydaje się, że w przypadku wystąpienia efektu porządku faworyzowane są ostatnie odpowiedzi, jeśli pytanie jest krótkie (*recency effect*), a pierwsze, jeśli pytanie jest długie (*primacy effect*) (Schuman i Presser 1981). Nico J. Molenaar podsumowując najważniejsze rezultaty badań prowadzonych na ten temat stwierdza, że w przypadku pytań krótkich (alternatyw) efekt porządku odpowiedzi ujawnia się jedynie dla „najbardziej abstrakcyjnych pytań dychotomicznych”: obserwuje się wówczas *recency effect*; w przypadku pytań długich im lista odpowiedzi jest dłuższa, tym większa jest szansa wystąpienia *primacy effect* (Molenaar 1982, s. 67).

WYBÓR SŁÓW I PROBLEM FAŁSZYWEJ SYNONIMICZNOŚCI

Problem „niuansów w sformułowaniach pytań” (*tone of wording*) ma wiele wymiarów i jest dobrze znany profesjonalistom od budowy ankiet, choć bardzo często ignorowany przez osoby korzystające z sondaży opinii. Najbardziej znanym w historii ankiet przykładem fałszywej synonimiczności jest różnica między „zabraniać” i „nie zezwalać”. W roku 1940, przy zastosowaniu procedury *split-ballot*, 54% pytaných Amerykanów zadeklarowało się jako zwolennicy zakazu (*forbid*) dyskusji skierowanych przeciw demokracji, a 75% – jako zwolennicy sądu, że nie wolno pozwalać (*not allow*) na takie dyskusje. Eksperyment ten, przeprowadzony przez Elmo Rope-

ra, został powtórzony 35 lat później przez Schumana i Pressera, przynosząc analogiczne rezultaty (Schuman i Presser 1981).

ISL podjął w marcu 1990 we współpracy z Marc Vincentem nową serię eksperymentów według procedury „dzielonych prób”, spośród których część dotyczyła użycia quasi-synonimów (Vincent 1990). W pierwszym z nich porównano dwa sformułowania pytania dotyczące prawa głosowania dla cudzoziemców („Czy uważa Pan(i) za normalne czy też nienormalne, że obcokrajowcy od kilku lat mieszkający i pracujący we Francji głosują w wyborach municypalnych?” z zastąpieniem „głosują” przez „mogą uczestniczyć w życiu lokalnym głosując”) oraz budowy meczetów („Czy uważa Pan(i) za normalne czy też nienormalne, żeby we Francji były meczety?” i („Czy uważa Pan(i) za normalne czy też nienormalne, żeby muzułmanie mieszkający we Francji mieli meczety dla praktykowania swojej religii?”). Różnice w sformułowaniach nie miały istotnego wpływu na uzyskane odpowiedzi (największa różnica wyniosła 3%).

W drugim eksperymencie przeciwstawiono formę „obiektywną” pytania (opinię) formie „subiektywnej” (postawie). Tabela 9 zawiera rezultaty uzyskane za pomocą pytań: „Czy uważa Pan(i) za normalne czy też nienormalne, że imigranci od kilku lat mieszkający i pracujący we Francji głosują w wyborach municypalnych?” i „Czy jest Pan(i) zwolennikiem czy też przeciwnikiem przyznania prawa głosu w wyborach municypalnych imigrantom mieszkającym i pracującym we Francji od kilku lat?”. Przejście od sformułowania „obiektywnego” do sformułowania „subiektywnego”, implikujące pewne zaangażowanie, spowodowało niewielkie zmniejszenie udziału odpowiedzi negatywnych na korzyść odpowiedzi niezaangażowanych.

Tabela 9.
Stopień implikowanego zaangażowania w odpowiedzi (głosowanie imigrantów w wyborach municypalnych)

Opinia ogólna		Skłonność osobista		Różnica
Normalne	38%	zwolennik	37%	-1%
Nienormalne	51%	przeciwnik	46%	-5%
Brak opinii	11%	indyferentny	17%	+6%

Trzeci eksperyment przyniósł rezultaty podobne, aczkolwiek ich kierunek jest przeciwny. Zadano następujące pytania: „Czy uważa Pan(i) za normalne czy też nienormalne budowanie meczetów we Francji?” i „Czy jest Pan(i) zwolennikiem czy też przeciwnikiem budowy meczetów we Francji?”.

Tabela 10.
Stopień implikowanego zaangażowania w odpowiedzi (budowa meczetów we Francji)

Opinia ogólna		Skłonność osobista		Różnica
Normalne	40%	zwolennik	30%	-10%
Nienormalne	45%	przeciwnik	44%	-1%
Brak opinii	15%	indyferentny	26%	+11%

Stwierdzono (tabela 10), że jeśli 40% respondentów uznaje za coś normalnego budowanie meczetów we Francji, to tylko 30% to zwolennicy. Sformułowanie „subiektywne” spowodowało tutaj zmniejszenie udziału odpowiedzi pozytywnych na korzyść odpowiedzi niez zaangażowanych. W obydwóch przypadkach za pomocą pytania o postawę, wymagającego bardziej osobistego zaangażowania osoby badanej, uzyskiwano wyraźnie wyższy udział odpowiedzi „neutralnych” niż za pomocą pytania o opinię.

Paralelnie do przedstawionych tutaj analiz na „dzielonych próbach”, Marc Vincent porównywał rezultaty ankiet odrębnych, ale realizowanych w odstępie zaledwie kilkudniowym. Stwierdził on, że zastąpienie słowa „imigranci” słowem „obcokrajowcy” nie miało znaczącego wpływu na odpowiedzi na pytanie dotyczące głosowania w wyborach municypalnych. Kiedy jednak zastąpiono „Czy uważa Pan(i) za normalne [...] żeby były meczety we Francji” przez „budowanie meczetów”, odpowiedzi „normalne” obniżyły się z 52% do 40%, a odpowiedzi „nienormalne” wzrosły z 35% do 45% (tamże).

Problem wyboru słów w pytaniu jest sprawą delikatną dla praktyków układających ankietę zwłaszcza, że jak się wydaje, nawet subtelne zmiany mogą mieć istotne skutki dla uzyskiwanych odpowiedzi (Schuman i Presser 1981). Trudność badań w tej dziedzinie polega na heterogeniczności zagadnień, które się na nią składają. Eksperymenty dotyczące pozornej synonimiczności opisywane w literaturze naukowej dotyczą tak różnych problemów, jak stopień zaangażowania osobistego badanych (opinie, postawy, zachowania przyszłe lub hipotetyczne), skojarzeń, które pobudzają pewne słowa (konotacje pozytywne lub negatywne, *loaded items*), czy też zawartość informacyjna lub przedstawiane argumenty. Wydaje się, że w zakresie tej właśnie problematyki istnieje pilna potrzeba wypracowania koherentnej ramy teoretycznej.

PORZĄDEK PYTAŃ I EFEKT KONTEKSTU

Miejsce pytania w kwestionariuszu nie jest sprawą bez znaczenia; w szczególności refleksje pobudzane w toku wywiadu, informacje, których *explicite* lub *implicite* dostarczają pytania poprzedzające, czy też ewolucja relacji między ankieterem, a respondentem mogą wpływać na udzielane odpowiedzi. W związku z ankietą dotyczącą opinii Francuzów na temat zagadnień demograficznych (zrealizowaną na przełomie lat 1959–1960) Alain Girard i Henri Bastide wykazali, że udział osób oceniających liczbę urodzin we Francji jako zbyt wysoką był w znaczący sposób niższy, jeśli pytanie zostało zadane na końcu wywiadu niż wówczas, kiedy zadawano je na początku (cyt. za Grémy 1987).

Podczas drugiej serii eksperymentów ISL (marzec 1990) Marc Vincent ocenił efekty permutacji pytania dotyczącego głosowania obcokrajowców w wyborach municypalnych i pytania na temat istnienia meczetów we Francji. Zrealizowano dwa analogiczne eksperymenty, w których różnice polegały na użytych sformułowaniach. Rezultaty ich zastosowania są identyczne: podczas gdy udział odpowiedzi na pytanie dotyczące meczetów nie zmienił się, odpowiedzi dotyczące głosowania są mniej pozytywne, kiedy pytanie dotyczące meczetów zostało zadane wcześniej (tabela 11).

Tabela 11.

Efekt kontekstu. (Głosowanie obcokrajowców w wyborach municypalnych)

Odpowiedź	Po „meczetach” – Przed „meczetaami” = różnica	
	„uczestniczyć w życiu lokalnym	„głosować”
Normalne	34%–42%=–8%	31%–40%=–9%
Nienormalne	52%–46%=+6%	56%–49%=+7%
Brak opinii	14%–12%=+2%	13%–11%=+2%

Autor stwierdza, że „porządek tych dwóch pytań [...] ma tylko bardzo słaby wpływ na odpowiedzi na pytanie dotyczące meczetów, ale występuje silny wpływ pytania dotyczącego meczetów na odpowiedzi dotyczące głosowania [...]. Oczywiście słowo »obcokrajowiec« nie powoduje tych samych ewokacji, kiedy jest użyte bezpośrednio po pytaniu dotyczącym muzułmanów i wówczas, kiedy jest użyte w pierwszym pytaniu” (Vincent 1990: 6).

Problem kontekstu dyskursywnego pytania jest dobrze znany praktykom zajmującym się przygotowaniem ankiet. Sygnalizowany już przez Hadleya Cantrila w związku z permutacją dwóch podobnych (Cantril 1944), stanowił następnie przedmiot refleksji teoretycznej Schumana i Pressera (1981) i Nico J. Molenaar (1982). Według tych autorów, efekt kontekstu występuje jedynie wówczas, gdy pytania dotyczą tego samego tematu lub tematów pokrewnych; pytania podsumowujące („Aby podsumować to, co Pan(i) mi powiedział(a), czy [...]?”) są bardziej wrażliwe na ten rodzaj efektu niż pytania szczegółowe; i wreszcie można wyróżnić co najmniej pięć różnych efektu kontekstu (tamże).

W praktyce efekt ten nie jest łatwy do kontroli, zwłaszcza, gdy kwestionariusz zawiera jedno lub więcej pytań filtrujących, kierujących sekwencją pytań. Możliwe rozwiązanie polega na równoległym stosowaniu w tej samej ankiecie kilku różnych porządków pytań (Jacquart 1988).

WNIOSKI

Przedstawione tutaj rezultaty badań kładą akcent na sytuacje, w których zróżnicowanie sformułowań pytań powoduje skutki dla dystrybucji uzyskanych odpowiedzi. Nie należy jednak z tego wyciągać wniosku, że każda różnica w sformułowaniu pytania ma wpływ na uzyskiwane rezultaty; badania empiryczne, co jest naturalne, kładą nacisk na te punkty, o których praktycy z góry wiedzą, że są podatne na efekt sformułowania. Przegląd eksperymentów zrealizowanych między 1936 a 1940 rokiem według metody „dzielonych prób” wskazuje, że gdy chodzi o pierwsze badania, znaczący efekt wystąpił jedynie w mniej niż połowie przypadków.

Przedstawiony tutaj przegląd eksperymentów prowadzonych ostatnio we Francji dotyczył jedynie wyników brutto („marginesów”). Pełna analiza rezultatów nie została jeszcze zakończona. Na obecnym etapie ma ona dwa cele: 1. Weryfikacja hipotezy, zgodnie z którą osoby, których problematyka pytania nie dotyczy bezpośrednio, są bardziej podatne na efekt sformułowania niż te, które są silnie w nią

zaangażowane; 2. Określenie, jakie czynniki zabezpieczają (przy kontroli pozostałych) przed wystąpieniem tego efektu. Jeżeli chodzi o tę ostatnią sprawę, to podstawowa teza Schumana i Pressera, zgodnie z którą osoby o wyższym poziomie wykształcenia są mniej podatne na efekty sformułowania, nie znajduje potwierdzenia w pierwszych wynikach analiz. Jednakże na obecnym ich etapie problem pozostaje otwarty.

Konkludując, przyszłe badania nad problematyką sformułowań pytań ankietowych powinny być prowadzone zgodnie ze ścisłymi wymogami metodologicznymi. Przede wszystkim należy zachować ścisłe warunki eksperymentalne; przedmiot eksperymentu powinien być ograniczony do jednego zagadnienia (np. różnice w formułowaniu jednego tylko pytania) i wszystkie pozostałe pytania kwestionariusza powinny być identyczne dla wszystkich prób, aby uniknąć ewentualnego efektu kontekstu. Ponadto, co podkreśla Nico J. Molenaar, konieczne jest opracowanie koherentnej ramy teoretycznej, zawierającej jasne hipotezy i opisującej sposoby ich rozstrzygnięcia na drodze empirycznej. Przy opracowaniu takiej ramy teoretycznej powinni współpracować specjaliści tak różnych dziedzin, jak logika stosowana (logika erotetyczna, teoria dialogu), lingwistyka (semantyka, syntaksa pytań), psychologia (technika wywiadu, interakcja między ankierem a respondentem), nie zapominając oczywiście o praktykach w zakresie przygotowania ankiet.

Przekład Paweł B. Sztabiński

LITERATURA

- Belson, W. A. 1981, *The Design and Understanding of Survey Questions*, Aldershot: Hower.
- Cantril, H. 1944 (red.) *Gauging Public Opinion*, Princeton: Princeton University Press.
- Cantril, H., Strunk M. 1951 *Public Opinion 1935–1946*, Princeton: Princeton University Press.
- Converse, J. M., Presser S. 1986 *Survey Questions. Handcrafting the Standardized Questionnaire*, Beverly Hills: Sage.
- Grémy, J.-P. 1987 *Les expériences françaises sur la formulation des questions d'enquête. Resultat d'un premier inventaire*, „Revue française de sociologie”, 4, (XXVIII).
- Grémy, J.-P. 1989 *La question en question*, „ISL Information”, (6–7) 37.
- Jacquart, H. 1988 *Qui? Quoi? Comment? ou La pratique des sondages?*, Paris: Eyrolles.
- Michelat G., Simon M. 1985 *Les „sans reponse” aux questions politiques*, „Pouvoirs”, (41–56) 33.
- Molenaar, N. J. 1982 *Response – effects of „Formal” Characteristics of Questions* [w:] Dijkstra, W. J. van der Zouwen, *Response Behaviour in the Survey. Interview*, London: Academic Press.
- Payne, S. 1951 *The Art of Asking Questions*, Princeton: Princeton University Press.
- Roussel, J. D. Mouthon 1988 *La question en question...*, „ISL Information”, 6 (35).
- Schuman, H. S. Presser 1981 *Questions and Answers in Attitude Surveys*, New York: Academic Press.
- Sudman, S. Bradburn N. M. 1974 *Response Effects in Surveys*, Chicago: Aldine.
- Sudman, S. Bradburn N. M. 1983 *Asking Questions. A Practical Guide to Questionnaire Design*, San Francisco: Jossey – Bass.
- Vincent, M. 1990 *Immigres: voile, mosquées, vote ... pour ou contre? ... Poids des mots et realites*, „ISL Informations”, 2–7 (39).

THE ADMINISTRATION OF QUESTIONS IN A QUESTIONNAIRE SCHEDULE

A new series of *split-ballot* experiments into the effects of the formulation of questions upon the results of surveys have been carried out in France since 1987. The main findings of these experiments deal with such problems as middle alternative, *dont knows* answers, response order, tone of wording, and questions' order. The analyses demonstrate, how far these elements affect answers given by respondent during an interview.