

DANIEL DARIUSZ MIDER¹

Sztuka i nauka formułowania pytań w badaniach ilościowych. Nowa rzeczywistość, nowe wyzwania

Art and science of creating a questionnaires in quantitative research. New reality and new challenges

Streszczenie

Tekst wzbogaca refleksję teoretyczną nad standaryzacją narzędzi pomiarowych w metodologii nauk społecznych. Potrzeba powstania niniejszego tekstu jest również dydaktyczna oraz porządkująca dotychczasową refleksję. Przedstawiono autorską koncepcję analizy pytań kwestionariuszowych na trzech płaszczyznach: merytorycznej, logicznej i psychologicznej. Płaszczyzna merytoryczna odnosi się do podziału ze względu na sposób konstrukcji – strukturę, format – pytania i / lub odpowiedzi. Z kolei płaszczyzna logiczna to podział ze względu na funkcje, jakie dane pytania, względnie grupy pytań, spełniają w szerszym kontekście, a zatem w odniesieniu do całości kwestionariusza. Płaszczyzna psychologiczna została wyróżniona ze względu na zamierzone lub niezamierzone, lecz nieuniknione reakcje emocjonalne wywoływane u respondenta. Ostatnia część zawiera refleksję nad najczęstszymi błędami w narzędziach badawczych oraz zalecenia ich unikania (w szczególności jest to postulat opracowywania złożonych, opartych na wcześniejszych obserwacjach wskaźników, próba wstępnej typologii błędów w formułowaniu pytań oraz kafeterii, czasu trwania wywiadu oraz zapewnieniu badanemu komfortu zarówno w trakcie, jak również po zakończeniu badania (debriefing).

¹ Daniel Dariusz Mider, Wydział Nauk Politycznych i Studiów Międzynarodowych, Uniwersytet Warszawski, Polska, e-mail: d.mider@uw.edu.pl, ORCID ID: <https://orcid.org/0000-0003-2223-5997>.

Słowa kluczowe:

metodologia, badania ilościowe, kwestionariusz, standaryzacja narzędzi badawczych, badania sondażowe

Abstract

The text enriches the theoretical reflection on the standardization of measurement tools in the methodology of social sciences. The need for this text is also, to some extent, didactic and organizing the reflection so far. This article presents the original concept of analyzing questionnaire questions on three following levels: content-related, logical, and psychological. The substantive / content-related level refers to the division according to the way of construction - structure, format - questions and/or answers. On the other hand, the logical plane is a division according to the functions given questions or groups of questions fulfilled in a broader context, i.e. concerning the whole questionnaire. The psychological plane was distinguished due to the intended or unintended but unavoidable emotional reactions evoked by the respondent). In the last part of the text, a reflection on the most common errors in research tools is presented and recommendations for avoiding them are formulated.

Keywords:

methodology, quantitative research, questionnaire, standardization of measurement tools, survey research

Pogłębiona refleksja nad strukturą, funkcjami, sformalizowanymi i niesformalizowanymi zasadami konstruowania narzędzi badawczych w badaniach ilościowych stanowi przedmiot poniższego tekstu. Złożoność świata społecznego i politycznego, niedoskonałość oraz niepewność wykorzystywanych powszechnie klasycznych metod pomiaru wymusza na badaczu poszukiwanie wszelkich dostępnych możliwości uczynienia pomiaru bardziej precyzyjnym i adekwatnym. Waga ilościowych badań empirycznych jest nie do przecenienia – George Gallup i Saul Rae (1940) w wydanej w latach czterdziestych XX wieku wpływowej książce *The Pulse of Democracy* mianem tytułowego pulsu demokracji obdarzają sondaże polityczne, co stanowi jasną sugestię, że są one jednym z kluczowych mechanizmów demokratycznej gry władzy. Literatura przedmiotu rzadko odnotowuje akademicką i jednocześnie praktyczną, kompleksową refleksję nad konstruowaniem narzędzi badawczych (Babiński, 1984; Gruszczyński, 2003; Krok, 2015; Lutyńska, 1984; Sawiński, P.B. Sztabiński i F. Sztabiński, 2000; Zagańczyk, 2014); brak również nacisku, by sztukę ich tworzenia jako niezbędnych do osiągnięcia poznawczych

celów naukowych traktować w kategoriach nieodzownego komponentu dydaktyki nauk społecznych. Potrzeba powstania niniejszego tekstu jest więc w głównej mierze dydaktyczna, a także systematyzująca, wzbogacająca refleksję teoretyczną nad standaryzacją narzędzi pomiarowych w metodologii nauk społecznych.

Wysiłki ogniskowano na poszczególnych częściach znaczących kwestionariusza wywiadu – pytaniach oraz blokach pytań. Te elementy każdego narzędzia badawczego posiadają złożoną specyfikę oraz funkcje. Z punktu widzenia praktyki badań wyróżnić dają się trzy płaszczyzny kategoryzacji (klasyfikacji) pytań kwestionariuszowych:

1. **Płaszczyzna merytoryczna** (podział ze względu na sposób konstrukcji – strukturę, format – pytania i/lub odpowiedzi);
2. **Płaszczyzna logiczna** (podział ze względu na funkcje, jakie dane pytania, względnie grupy pytań, spełniają w szerszym kontekście, tj. w odniesieniu do całości kwestionariusza);
3. **Płaszczyzna psychologiczna** (podział ze względu na zamierzone lub niezamierzone, lecz nieuniknione reakcje emocjonalne wywoływane u respondenta).

1. ASPEKT MERYTORYCZNY

Każde z zawartych w kwestionariuszu pytań daje się zaliczyć do jednej z następujących kategorii: **pytań zamkniętych** (skategoryzowanych / pytań o wybór / pytań z wyborem), **pytań otwartych** (nieskategoryzowanych), **pytań półotwartych** (półzamkniętych) oraz **pytań prekategoryzowanych**². Każda z tych kategorii nakłada na badacza oraz ankietera wymogi oraz generuje rozmaite efekty poznawcze (dane różnej jakości i przeznaczenia). Rudymet wszelkich narzędzi badawczych w badaniach ilościowych stanowią pytania otwarte i pytania zamknięte (V. Nachmias, D. Nachmias, 2001). Pytania zamknięte to takie, w których respondent dokonuje wyboru spośród wyszczególnionych wariantów skończonej liczby odpowiedzi. Z kolei w pytaniach otwartych badany udziela odpowiedzi spontanicznie, nie otrzymując do dyspozycji gotowych jej wariantów.

Podział ten jest kategoryzacją w rozumieniu gramatycznym (technicznym) – najczęściej stosowanym. Istnieje krzyżujące się z nim rozumienie: pytania

² Czasami w socjolekcie badaczy termin „pytanie prekategoryzowane” traktowany jest jako synonim pytania zamkniętego. W niniejszym opracowaniu jest rozumiany jako ustanawiający oddzielną kategorię pytań.

zamkniętego i otwartego w sensie logicznym. Pytaniem zamkniętym w sensie logicznym nazwiemy takie pytanie, w którym zbiór możliwych odpowiedzi na nie jasno wskazuje sama treść pytania. Podobnie, pytaniem otwartym w sensie logicznym, będzie takie, w którym liczba i rodzaj możliwych odpowiedzi jest nieskończona. Zwolennikiem stosowania pytań otwartych w badaniach ilościowych (sondażowych) był sam Paul F. Lazarsfeld (1935). W opublikowanym w 1935 roku artykule *The Art. Of Asking Why* wyliczył sześć głównych funkcji pytań otwartych: wyjaśnianie znaczenia odpowiedzi respondenta, wyodrębnienie kluczowych aspektów opinii, ujawnienie przesłanek opinii, pozyskanie odpowiedzi na złożone pytania dotyczące postaw, interpretowanie motywacji i wyjaśnianie zależności statystycznych. Ponadto argumentował, że zarówno pytania otwarte, jak i zamknięte są nieodzowne w każdym, kompleksowym programie badań empirycznych (Lazarsfeld, 1944). Nie wypada przeciwstawiać się autorytetowi, jednak nie mógł brać on pod uwagę innych technik badawczych niż te, gdzie respondent i ankieter stykają się osobiście. W wywiadach telefonicznych, a tym bardziej w ankietach online pytania otwarte stanowią nie lada problem z powodu niedokładności i nieadekwatności pozyskiwanych odpowiedzi. Zalecenie praktyczne w przypadku tych metod badawczych sugeruje, by ograniczać liczbę pytań otwartych w kwestionariuszu do niezbędnego minimum (by nie mogły zostać uznane za wyraz bezradności poznawczej badacza), dokonywać uprzednich badań pilotażowych mających na celu ich wyeliminowanie oraz dokonywać ich zamknięcia (zamiany co najmniej na pytania półotwarte) po zrealizowaniu pewnej części próby (optymalnie ok. 10–25 proc. całości próby) (Hyman, Sierra, 2016).

Nader często stosowanym rodzajem pytania kwestionariuszowego jest **pytanie półotwarte**, to jest zawierające kilka wariantów odpowiedzi (jak w pytaniu zamkniętym), lecz respondent ma możliwość udzielenia odpowiedzi innej niż wyszczególnione (charakterystyczne dla pytania otwartego). Przedłożenie takiej ewentualności respondentowi winno być starannie przemyślane, a cele omówione podczas szkolenia z ankieterami, bowiem nierzadko zdarza się, że osiągnięte efekty poznawcze są niesatysfakcjonujące (lokowanie się w tej kategorii licznych odpowiedzi stanowiących uniki). Natomiast **pytanie prekategoryzowane** to takie, które jest zamknięte dla ankietera, lecz otwarte dla respondenta. Wymienione rodzaje pytań lokują się pomiędzy pytaniami zamkniętymi i otwartymi, wymykając się jednoznacznej klasyfikacji.

W kategoriach merytorycznych istotny z punktu widzenia realizacyjnego i jakości uzyskiwanych danych jest podział ze względu na przedmiot pytania. Wyróżniamy: pytania o fakty, pytania o ocenę faktów oraz pytania o opinie, postulaty i wnioski. W **pytaniach o fakty** badacza interesuje tzw. prawda materialna,

czyli fakty ze sfery obiektywnej. Są to stosunkowo najprostsze do sformułowania pytania, na ogół też sposób ich realizacji, tj. zadawania respondentowi, nie budzi wątpliwości. W **pytaniach o ocenę faktów** badacz zainteresowany jest kwestiami spoza sfery obiektywnej, a więc prawdą psychologiczną. Stосуje się tu bardziej wysublimowane, a jednocześnie narażone na artefakty sposoby pomiaru. Z kolei **pytania o opinie** w wielu opracowaniach zaliczane są do kategorii pytań powyższych, czyli o ocenę faktów, jednakże czasem wyróżniane ze względu na to, że opinie postulatory i wnioski nie muszą u różnych respondentów być tożsame pomimo jednoznacznej oceny danego faktu. Pytania te sprawiają analogiczne trudności realizacyjne.

Kategoria pytań zamkniętych w sensie gramatycznym nie jest jednorodna. W jej ramach można zastosować dwa następujące sposoby klasyfikacji: podział ze względu na liczbę udzielanych odpowiedzi oraz podział ze względu na zastosowany sposób pomiaru. W ramach podziału ze względu na liczbę udzielanych przez respondenta odpowiedzi wyróżniamy następujące dwie kategorie – pytania jednoodpowiedziowe i pytania wieloodpowiedziowe.

Pytania jednoodpowiedziowe to takie, na które respondent może udzielić jednej tylko odpowiedzi. Jeśli spośród zbioru danych odpowiedzi tylko jedna jest logicznie możliwa, pytanie takie nazywamy pytaniem w formie **dysjunkcji naturalnej** bądź **logicznej** (np. pytanie o płeć lub wiek). Jeżeli natomiast logicznie dopuszczalna jest więcej niż jedna odpowiedź, takie pytanie jest **dysjunkcją sztuczną** lub **wymuszoną**. Skrajną formę dysjunkcji naturalnej stanowi **pytanie dychotomiczne (pytanie rozstrzygnięcie lub pytanie w formie alternatywy)**. Pytanie dychotomiczne to takie, w którym respondent posiada możliwość wyboru spośród dwóch wzajemnie wykluczających się odpowiedzi (przy czym należy zdać sobie sprawę, że w pytaniu takim możliwe są zawsze dwie inne reakcje respondenta: po pierwsze, jest to odpowiedź „nie wiem / trudno powiedzieć”, a po drugie, odmowa udzielenia odpowiedzi na pytanie).

Pytania wieloodpowiedziowe to inaczej pytania kafeteryjne lub „z wielocłonową listą alternatyw”. Tu należy poczynić ważne zastrzeżenie: kategoria ta została stworzona jako logiczne dopełnienie powyższej. O ile w kategorii pytań jednoodpowiedziowych mogły znaleźć się tylko takie, na które można udzielić tylko jednej odpowiedzi, o tyle tutaj znajdują się rodzaje pytań dopuszczające udzielenie zarówno jednej, jak i wielu odpowiedzi. Ergo część z nich należy jednocześnie do obu zbiorów, a równie adekwatną w sensie merytorycznym nazwą tego punktu klasyfikacji byłaby „pozostałe pytania dopuszczające udzielenie więcej niż jednej odpowiedzi”. Zdanie ujęte w cudzysłów jest jednocześnie definicją pytania wieloodpowiedziowego. W sensie dosłownym pytaniami wieloodpowiedziowymi

są **pytania w formie koniunkcji** (pytania koniunkcyjne), bowiem są one tak skonstruowane, że respondent może nieodmiennie wybierać więcej niż jedną odpowiedź. Pytania koniunkcyjne mogą przybrać jedną z czterech następujących form:

1. **maksymalnego roszczenia zupełności**, czyli dyspozycji, by respondent wybrał te wszystkie możliwości odpowiedzi, które go dotyczą;
2. **minimalnego roszczenia zupełności**, tj. prośby o wybranie co najmniej jednej z wyszczególnionych odpowiedzi, dopuszczającej jednak wybranie większej ich liczby;
3. **ograniczonego roszczenia zupełności** – żądania wybrania przez respondenta określonej liczby (np. trzech, czterech) odpowiedzi, które go dotyczą;
4. **nieokreślonego roszczenia zupełności** – braku dyspozycji dla respondenta co do liczby możliwych do wybrania odpowiedzi. Jest to kategoria, która nie powinna występować w badaniach kwestionariuszowych, ze względu na potencjalnie szeroki zakres rozrzutu odpowiedzi wskutek dowolnego ich traktowania przez ankierów i respondentów. Domyślnie traktować należy ją jako maksymalne roszczenie zupełności i tak oznaczać.

Logicznym dopełnieniem wyróżnionego w kategorii pytań jednoodpowiedziowych pytania rozstrzygnięcia jest **pytanie dopełnienie**, charakteryzowane jako „rozpoczynające się najczęściej od słowa *który* i żądające wyboru jednej lub kilku spośród wielu odpowiedzi logicznie możliwych” (Babiński, 1984, s. 55).

Drugim rodzajem jest **podział ze względu na zastosowany sposób pomiaru**. W jego ramach wyodrębnimy: pytania kafeteryjne³, pytania z użyciem skali oraz pytania z użyciem rangowania. **Pytanie kafeteryjne** (w rozumieniu wąskim) to takie, w którym respondent dokonuje wyboru poprzez wskazanie odpowiedzi spośród wyszczególnionych w formie listy możliwości.

Z kolei **pytania z użyciem skali** to takie, w których wyszczególnione punkty kafeterii (w rozumieniu szerokim), tworzące kontinuum, leżą w jednym wymiarze. Pytania te możemy podzielić ze względu na sposób opisu czy przedstawienia punktów (itemów) kafeterii (w wąskim rozumieniu) oraz liczbę biegunów. W pierwszym z przypadków wyróżnimy: skalę z opisem słownym, skalę z opisem numerycznym i częściowym słownym, skalę z pełnym opisem słownym i numerycznym oraz skalę z opisem wyłącznie numerycznym. Z realizacyjnego punktu widzenia najlepsze efekty uzyskuje się, gdy skala posiada zarówno opis słowny, jak i numeryczny.

³ Słowo „kafeteria” może być rozumiane dwojako: 1. Szeroko, gdy znaczy dla nas tyle co lista rozwijająca podstawową część pytania, niezależnie od charakteru tej listy; 2. Wąsko, gdy odpowiedzi wyszczególnione w danym pytaniu nie tworzą kontinuum, nie leżą w jednym wymiarze.

Pytania z użyciem skali mogą być jedno- lub dwubiegunowe. **Skale jedno-biegunowe** zawierają tylko część logicznego kontinuum skali, tj. rozciągają się w obszarze od umownego zera tylko w jedną stronę skali, np. ocen pozytywnych lub negatywnych, mają więc charakter jednoznacznie kumulatywny. Używane są bardzo rzadko, np. w przypadku gdy badacz z góry zakłada, że uzyskiwane odpowiedzi na pytanie będą negatywne. Z kolei **skale dwubiegunowe** – rozciągają się po obu stronach umownego zera.

Pytania z użyciem rangowania to takie, w których respondent oznacza stopień ważności, a więc porządkuje wedle pewnego kryterium przedłożone mu elementy. Rangowanie może być **nieprzechodnie**, gdy respondentowi wolno użyć więcej niż raz tej samej rangi, lub **przechodnie**, kiedy nie dopuszcza się dwukrotnego użycia tych samych rang. W tej grupie pytań można wyróżnić **pytania z wyborem preferencyjnym**, w których wybór respondenta dokonywany jest etapowo, wielokrotnie.

Istotny element mający wpływ na jakość zbieranych stanowi poziom dokonywanego pomiaru. Determinuje on właściwości zmiennej: jakość danych, pojemność informacyjną, sposoby przetwarzania danych, a więc zakres i rodzaj konkretnych miar statystycznych możliwych do zastosowania w analizach. W literaturze przedmiotu wyróżnia się cztery poziomy pomiaru zmiennych: nominalny, porządkowy (rangowy), interwałowy (przedziałowy) i ilorazowy (stosunkowy). Podział ten jest powszechnie stosowany w naukach społecznych i przyrodniczych, a wprowadzony został przez amerykańskiego psychologa Stanleya Stevensa (1946). Należy podkreślić, że o ile badacz ma wybór (siłą rzeczy wybór ów pozostaje w naukach społecznych ograniczony, bowiem większość zjawisk społecznych daje się ujmować wyłącznie w kategoriach nominalnych, względnie rangowych) winien bezwzględnie wybierać ilościowe poziomy pomiaru. Zważmy, że obniżanie jakości zmiennej (na przykład zrekodowanie zmiennej z poziomu interwałowego do nominalnego) jest możliwe i proste, o tyle ruch w przeciwnym kierunku jest logicznie niewykonalny⁴.

2. ASPEKT LOGICZNY

Aspekt logiczny odnosi się do służebnej funkcji pełnionej przez pytania dla całości kwestionariusza lub poszczególnych jego bloków, na ogół różnicując zadawanie

⁴ Z wyjątkiem sytuacji, gdy umownie i zwyczajowo potraktujemy skalę typu Likerta jako interwałową.

kolejnych pytań w zależności od udzielonych odpowiedzi. Ten sposób klasyfikacji pozwala na wyróżnienie pytań filtrujących oraz ich szczególnego rodzaju – pytań screeningowych (selekcyjnych), a także pytań do ankietera. **Pytania filtrujące** nazywane są inaczej (lecz rzadko): pytaniami selekcyjnymi lub pytaniami-filtrami. Pytanie filtrujące to takie pytanie, za pomocą którego na początku całej baterii pytań⁵ wydziela się grupę respondentów, która zdaniem badacza winna udzielać odpowiedzi na pytania zawarte w baterii. Pytania uwarunkowane przez filtr nazywane są **pytaniami alternatywnymi**⁶. Pytania filtrujące zawierają więc reguły zadawania kolejnych pytań:

- jeśli reguła nakazuje zadać następne pytanie, nazywamy ją **regułą sekwencji**,
- jeśli reguła nakazuje zadać pytanie inne niż kolejne, nazywamy ją **regułą przeskoku** lub **przejścia**.

Pytania zadawane po zastosowaniu filtra tworzące różne ścieżki kwestionariusza to tzw. **pytania wykluczające się**.

Częstymi błędami jest zamknięcie lub przeciwnie – otwarcie danej baterii pytań alternatywnych na niewłaściwe grupy badanych. Wspomaganie komputerowe umożliwia stosowanie bardzo wysublimowanych reguł filtrów. Można stosować nie tylko proste odpowiedzi na pytania, lecz złożone wskaźniki stanowiące kombinację licznych odpowiedzi, w tym opartych na kalkulacjach ilościowych i tworzonych w toku trwania wywiadu wskaźników. **Pytania screeningowe** (screenery) jest to szczególna forma pytań filtrujących. Umieszczane na początku kwestionariusza są narzędziem selekcji, eliminując osoby nienależące do grupy docelowej (badanej). Z kolei **pytania do ankietera** to takie pytania, na które odpowiedzi musi udzielić sam ankieter, postępując zgodnie z otrzymaną instrukcją.

3. ASPEKT PSYCHOLOGICZNY

Aspekt psychologiczny odnosi się do funkcji modyfikacji emocjonalnego komponentu postawy respondenta wobec całości interakcji związanej z wywiadem lub wycinkowo – wobec określonego pytania. Z tego punktu widzenia wyróżnić można następujące typy pytań: buforujące, drażliwe, kontaktowe, pośrednie oraz

⁵ Bateria pytań to pytania powiązane ze sobą logicznie, odnoszące się do jednego zagadnienia, tematu.

⁶ Niekiedy związek frazeologiczny „pytania alternatywne” bywa używany na określenie zamienników pytań wykorzystywanych przez ankietera, gdy respondent nie zrozumiał przedłożonego mu pytania w wersji pierwotnej.

pytania-pułapki. **Pytania buforujące** to pytania mające neutralizować negatywne napięcie emocjonalne wywołane pytaniem lub pytaniami poprzednimi, zaś badacz przewiduje, że negatywny stan afektywny może mieć wpływ na jakość odpowiedzi na pytania następne. Podobną funkcję pełnią **pytania łączniki** – zapewniają one możliwie łagodne przejście pomiędzy bateriami pytań. Niekoniecznie muszą mieć one gramatyczną formę pytania – czasem są wyłącznie informacją o zakończeniu jednego bloku pytań i przejściu do kolejnego. Nie oznacza to, że pytania zaklasyfikowane jako buforowe są pytaniami niepotrzebnymi, nieistotnymi – po prostu pytania trudne, drażliwe przedziela się niekiedy pytaniami obojętnymi, neutralnymi dla respondenta. Nieco inne znaczenie nadawane jest pojęciu „**pytania-wata**” – ich funkcją jest przeciwdziałanie znużeniu respondenta w kwestionariuszach długich i trudnych, zapewnienie odpoczynku przed następną baterią trudnych pytań. Ich cechą charakterystyczną jest to, że mogą być luźno powiązane z tematyką wywiadu; ich celem jest wywołanie u respondenta miłych, przyjemnych skojarzeń. **Pytania kontaktowe** (pytania „na rozgrzewkę”, pytania rozbiegowe) uczą respondenta jego roli, a więc zaakceptowania sekwencji interakcyjnej pytanie–odpowieź. Udzielona odpowiedź ma znaczenie wtórne, liczy się wdrożenie właściwego wzorca reakcji badanego. Wyróżnia się także kategorię tzw. **pytań rozgrzewających** (treningowych) – uczących respondenta na przykładach spełniania przezeń zadań nakładanych przez pytania trudne, zadawane dopiero po zakończeniu takiej nauki. **Pytania pośrednie** są zadawane nie ze względu na ich funkcję bezpośrednią, czyli chęć uzyskania informacji od respondenta, lecz dlatego, że tak uzyskaną informację traktujemy jako wskaźnik występowania lub natężenia innego zjawiska. Zadajemy je wówczas, gdy pytamy o kwestie drażliwe, zjawiska społecznie nieakceptowane. Logicznym dopełnieniem kategorii pytań pośrednich są **pytania bezpośrednie**. **Pytania-pułapki** należą do kategorii pytań, które nie służą pomiarowi badanego zjawiska, nie są one przez to związane bezpośrednio z tematyką badania, to jest potrzebami informacyjnymi badacza lub są związane z nim pozornie. Ich funkcją jest sprawdzenie stopnia prawdomówności respondenta. Do pytań-pułapek zaliczymy **pytania puste**, czyli takie, na które niemożliwością jest odpowiedzieć prawdziwie, bowiem dotyczą nieistniejących rzeczy, problemów, zjawisk. Pytanie to umożliwia sprawdzenie, w jaki sposób respondent zachowa się w przypadku, gdy zapytamy go o zjawisko, o którym nie wie lub nie słyszał.

Z psychologicznego punktu widzenia warto wyróżnić tzw. **pytania drażliwe**. Są to pytania o informację, która w rozumieniu respondenta nie powinna być udzielana osobom postronnym lub o którą nie powinno się pytać zgodnie z pewnym usankcjonowanym obyczajem. W zależności od zmiennych socjodemograficznych charakteryzujących respondenta pytania drażliwe dotyczyć mogą różnych kwestii:

wykształcenia i wysokości zarobków u mężczyzn z wykształceniem niższym niż średnie, wysokości zarobków u prywatnych przedsiębiorców, wieku u kobiet itd. Nie należy tej kategorii pytań mylić z pytaniami trudnymi, tj. pytaniami dotyczącymi skomplikowanych kwestii, przy których respondent może mieć problemy z ich zrozumieniem i udzieleniem właściwej odpowiedzi. Natomiast **pytania sprawdzające** (walidujące) służą bezpośredniej kontroli respondenta pod kątem prawdy materialnej lub psychologicznej zawartej w jego odpowiedziach. Pytania te mogą zasadniczo przyjmować formę zespołu dwóch lub więcej pytań dublujących się pod względem przedmiotu pytania, lecz nie pod względem treści, umieszczonych w różnych miejscach kwestionariusza. Na etapie analiz zestawia się udzielone przez respondenta odpowiedzi na nie w celach porównawczych.

W omawianym kontekście sytuacyjnym psychologicznego oddziaływania na respondenta stosuje się także czasami tzw. strategię lejka. Przez **strategię lejka** rozumiemy taką baterię pytań, w których każde następne pytanie (oczywiście z wyjątkiem pierwszego, otwierającego ów lejek) jest powiązane z poprzednim i ma coraz mniejszy zakres. Taki układ pytań pozwala na poznanie szczegółowych poglądów respondenta wówczas, gdy nie jest on wystarczająco motywowany do ich udzielenia lub gdy nie chcemy narzucać mu szczegółowym pytaniem własnej perspektywy widzenia badanego problemu. Natomiast w **strategii odwróconego lejka** wąsko sformułowane pytania poprzedzają te szeroko sformułowane. Stosowane są wówczas, gdy respondent nie ma wystarczających informacji do ogólnej oceny danego zagadnienia lub gdy nie jest do niej wystarczająco motywowany.

Warto również przytoczyć egzemplifikacje bardziej wysublimowanych sposobów redukcji napięcia emocjonalnego respondenta w związku z sytuacją wywiadu. Zdarza się, że respondent nie chce dzielić się wiedzą ze względu na fakt, że odpowiedź stawia go złym światłem, wywołuje poczucie wstydu, stanowi jedno ze społecznych tabu lub z innych powodów. Jednym ze sposobów pozyskiwania takich informacji jest technika polegająca na podzieleniu badanych na dwie równoliczne podpróby. Dla każdej z wyodrębnionych podprób konstruujemy dwie kafeterie. Pierwsza z list obejmuje wyłącznie tak zwane zachowania kontrolne, to jest takie wskazania, które mają charakter neutralny dla respondenta. Druga lista obejmuje te same zachowania kontrolne, a także zachowanie będące przedmiotem zainteresowania badacza (a więc takie, które z różnych względów nie jest dla respondenta neutralne). Ważna jest też sama konstrukcja pytania – nie jest to pytanie typowe. Przykład takiej konstrukcji przedstawia tabela 1.

Należy zwrócić uwagę, że istotą tego pytania jest komfort psychiczny dostarczany badanemu, który nie jest zmuszany, by bezpośrednio odpowiadać na

Tabela 1. Pozyskiwanie odpowiedzi na pytania drażliwe – schemat narzędzia pomiaru

Wersja A pytania	Wersja B pytania
Ile z poniższych czynności wykonał/a Pan/ Pani w ciągu ostatniego roku:	Ile z poniższych czynności wykonał/a Pan/ Pani w ciągu ostatniego roku:
<ul style="list-style-type: none"> • Pójście do teatru • Pozostawienie samochodu na płatnym parkingu • Pójście na przyjęcie organizowane przez znajomego z pracy • Wizyta u lekarza specjalisty 	<ul style="list-style-type: none"> • Pójście do teatru • Pozostawienie samochodu na płatnym parkingu • Pójście na przyjęcie organizowane przez znajomego z pracy • Popęnienie myślozbrodni* • Wizyta u lekarza specjalisty

* zachowanie stanowiące przedmiot zainteresowania badacza, pozostałe elementy list (tożsame w wersji A i B mają charakter kontrolny)

Źródło: opracowanie własne.

pytanie o charakterze drażliwym. Po wykonaniu pomiaru podliczamy odsetki odpowiedzi w wersji A oraz w wersji B, a następnie odejmujemy wynik A od B. Różnica będzie stanowić odsetek osób w populacji przejawiających zachowanie mierzone przez badacza. Jest to jednak technika ograniczona o niskiej pojemności informacyjnej – dowiadujemy się co prawda, jaka jest skala określonego zjawiska w populacji, jednak brak informacji na temat tego, który z badanych wskazał określoną odpowiedź, niemożliwe są zatem bardziej złożone analizy w poszczególnych próbkach.

Możliwe jest również zastosowanie innego sposobu dla pozyskania odpowiedzi na pytanie drażliwe, tak zwanej techniki losowej odpowiedzi (*Randomized Response Technique*). W tej technice również tworzona jest para pytań. Pierwsze z pytań ma wydźwięk neutralny, zaś rozkład odpowiedzi na nie winien być badaczowi znany *a priori* lub możliwy do łatwego i wiarygodnego sprawdzenia. Drugie pytanie dotyczy badanej, drażliwej kwestii. W obu pytaniach stosowana jest skala nominalna, dychotomiczna. Oba pytania prezentowane są respondentowi wraz z wyjaśnieniem powodów zastosowania takiej techniki oraz stosowną instrukcją. Akt ów ma znaczenie kluczowe dla powodzenia tej techniki. Respondent pouczany jest, że odpowiada na jedno z tych dwóch pytań (wyboru dokonuje przez rzut monetą), przy czym ankieter nie wie i nie może dowiedzieć się, na które z pytań respondent odpowiadał. Po zebraniu i zestawieniu danych należy od ogólnego odsetka odpowiedzi twierdzących łączyć odsetek odpowiedzi twierdzących na pytanie neutralne. Wynik jest odsetkiem odpowiedzi na pytanie drażliwe (John, Loewenstein, Acquistic i Vosgeraud, 2018).

Tabela 2. Egzemplifikacja techniki losowej odpowiedzi (*Randomized Response Technique*)

Pytanie neutralne	Pytanie drażliwe
Czy urodził(a) się Pan/Pani w miesiącu realizacji badania?	Czy zdarzyło się Pani/Panu kiedykolwiek popełnić myślzbrodnię?
1. Tak	1. Tak
2. Nie	2. Nie

Źródło: opracowanie własne.

Technika ta, w porównaniu z opisaną jako pierwszą, wymaga większej aktywności ze strony badanego – co najmniej zrozumienia istoty oraz uznania go za wiarygodny, to jest nieujawniający odpowiedzi respondenta w pytaniu drażliwym.

4. UWAGI KOŃCOWE O REGUŁACH KONSTRUKCJI NARZĘDZI BADAWCZYCH W BADANIACH ILOŚCIOWYCH

Na podstawie wieloletniej praktyki badawczej i akademickiej wypada sformułować na zakończenie kilka uwag w zakresie konstruowania narzędzi badawczych w badaniach ilościowych, wzięwszy pod uwagę dostrzegane błędy i pomyłki badawcze.

Najważniejszy wydaje się postulat, by kwestionariusz nie stanowił zbioru odrębnych pytań, choćby połączonych jednorodną tematyką. Warto poczynić wysiłki, by pytania mogły być analizowane w formie zbiorczych, syntetycznych wskaźników obejmujących wiele pytań lub całe ich bloki. W tym celu niezbędna jest **refleksja o charakterze teoretycznym** (przegląd istniejących modeli w literaturze przedmiotu) oraz refleksja koncepcyjna (utworzenie na podstawie poszczególnych zmiennych modelu empirycznych wskaźników dających sprowadzić się do wspólnego mianownika i analizować zbiorczo w formie jednej syntetycznej zmiennej). Nierzadko przeoczana jest również konieczność **refleksji o charakterze analitycznym**. Równie częsty, co nierozważny zabieg stanowi ułożenie listy pytań bez uprzednio przygotowanej koncepcji dalszych analiz. Spotyka się również nierzadko tendencję do tworzenia narzędzia badawczego jedynie z zamiarem zastosowania prostych technik statystyki opisowej. Narzędzie badawcze winno być konstruowane z jednoznacznym postanowieniem wykorzystania określonej, możliwie adekwatnej i zaawansowanej, techniki analitycznej lub – jeśli nie wiadomo zawczasu, jakie techniki zostaną użyte – kwestionariusz wywiadu winien być konstruowany na tyle przemyślane, by posiadać walor uniwersalności, umożliwiając potencjalne wdrożenie szerokiego spektrum technik statystycznych.

Należy również przeciwdziałać jednowymiarowemu traktowaniu kwestionariusza wywiadu. Narzędzie badawcze bywa traktowane zbyt wąsko – jako zbiór pytań adresowanych do respondenta z pominięciem instrukcji dla takich kluczowych w procesie pomiaru podmiotów, jak ankieter, analityk czy kontroler (osoba nadzorująca ankieterów podczas pracy i wstępnie walidująca wyniki badania).

Kolejny postulat stanowi **dbałość o komfort badanego**, co bezpośrednio skutkuje podtrzymywaniem jego motywacji do udziału w badaniu, a w efekcie niższym odsetkiem odmów (który może przecież powodować skrzywienie próby). Czynnikiem zwiększającym satysfakcję badanego z udziału w badaniu są właściwy (niezbyt długi) czas interakcji, by uczestnik nie poddał się znużeniu, a w ostateczności nie przerwał pomiaru. Na przykład w badaniach telefonicznych wspomaganych komputerowo wskazuje się jako górną granicę trwania wywiadu kwadrans zarówno w rozmowach wykonywanych na telefony stacjonarne, jak też komórkowe (Vicente, Reis i Santos, 2009). Silną przesłankę skracania trwania wywiadów stanowi jakość pozyskiwanych danych, a w mniejszym stopniu zwiększający się wraz z czasem trwania wywiadu odsetek rozmów przerwanych w trakcie trwania pomiaru (Cape i Philips, 2015). Niektórzy badacze wskazują nawet, że wywiad nie powinien przekraczać 10 minut, zaś badania trwające powyżej 12 minut uznają za „niebezpieczne” (Lester i Wilson, 1995). Z doświadczeń własnych wynika, że przeciętny respondent rozprasza się po upływie około pięciu minut od rozpoczęcia wywiadu i wówczas niezbędne jest podjęcie przez ankietera działań naprawczych, zaś za maksymalną, podręcznikową długość trwania wywiadu uznaje się wywiady dwudziestominutowe (Johnson, Siluk i Taraf, 2014). W warunkach polskich preferowany czas trwania pomiaru to od czterech do pięciu minut (44 proc. wskazań), do 10 minut (25 proc. wskazań) lub od jednej do trzech minut (17 proc. wskazań); w praktyce jednak przeciętny wywiad telefoniczny trwa około 17 minut (IMAS, 2020). Doświadczenia własne dowodzą, że użycie zwrotów i pytań pobudzających aktywność i zainteresowanie respondenta oraz właściwe przeszkolenie ankietera (znajomość oznak spadku zaangażowania w interakcję oraz umiejętność ponownego skupiania i podtrzymywania uwagi badanego) pozwalają na przeprowadzenie wywiadów nawet kilkudziesięciminutowych. Istotne znaczenie mają w kontekście dbałości o komfort respondenta również przejścia pomiędzy poszczególnymi częściami wywiadu oraz bateriami poszczególnych pytań utrzymujące właściwą dynamikę wywiadu (zastosowanie pytań buforujących). Elementem wartym przemyślenia jest kolejność pytań oraz kolejność baterii pytań, która winna być dyktowana względami poznawczymi, lecz również przesłankami związanymi z komfortem badanego. Nieodmiennie warto zwracać uwagę na język, w jakim formułowane jest narzędzie badawcze. Zdarza się, że

sformułowania w narzędziu badawczym bywają pozostawione w języku badacza wynikającym z teorii, a więc niezrozumiałym dla badanego.

Ogólne porady dotyczące narzędzia badawczego nie kończą listy zaleceń. Porównywalną wagę wypada przykładać do konstruowania narzędzia badawczego na poziomie poszczególnych pytań. Do najczęściej popełnianych należą **błędy logiczne** – jest to formułowanie pytań o tzw. podwójnej treści (błąd multiplikacji), czyli zawierających dwa lub więcej pytań jednocześnie. Zbliżony problem stanowią pytania o tak zwanej podwójnej negacji, to jest zawierające w zdaniu dwa przeczenia i choć poprawne logicznie, to jednak trudne do zrozumienia dla badanego, a wskutek tego generujące artefakty. Błędy w obszarze formy i treści pytań również skutkują doniosłymi, negatywnymi konsekwencjami. Zalecenie jest tyleż proste, co trudne w realizacji: konsekwentne stosowanie zasady „jedno zagadnienie = jedno pytanie” oraz rezygnacja z przeczeń w pytaniach (partykuła „nie”). W szczególności niebezpieczeństwo niosą: pytania zawierające terminy wieloznaczne (nieprecyzyjne), pytania zbyt długie (naruszające „złotą regułę 20 słów”), niewłaściwie zadane pytania trudne (bez stosownego wyluszczenia zagadnienia respondentowi), pytania drażliwe z pominięciem właściwego wstępu oraz – co nader częstsze – pytania zakładające fałszywe znawstwo (przeceniające wiedzę badanego), a także pytania zbyt precyzyjne (przeceniające pamięć respondenta). A zatem pytania należy konstruować tak, by badany ponosił jak najmniejszy wysiłek umysłowy przy udzielaniu na nie odpowiedzi. Pamiętać należy również o tym, by badanemu umożliwić uniknięcie odpowiedzi poprzez umieszczenie odpowiedzi „nie wiem, trudno powiedzieć” oraz „odmowa odpowiedzi” w kafeteriach. Ważny element to również forma i treść kafeterii. Wystrzegajmy się pytań o zbyt długiej kafeterii, a jeśli to niemożliwe – kafeterię winniśmy randomizować (prezentować każdemu z badanych w losowej kolejności). Istotny bywa również brak rozłączności kafeterii (nakładanie się na siebie znaczeń poszczególnych itemów) oraz brak wyczerpalności (kafeteria nie zawiera elementów, które powinna).

Tworzenie narzędzia pomiaru w badaniu ilościowym jest, jak sformułowano w tytule, tyleż nauką, co sztuką, bowiem zadośćuczynienie wyżej wskazywanym zasadom formalnych nie gwarantuje stworzenia doskonałego lub co najmniej dobrego narzędzia badawczego; w równej mierze liczy się koncepcja – jeśli nie błyskotliwa, to co najmniej elegancka – mająca źródło w twórczym przetworzeniu teoretycznych aspektów literatury przedmiotu.

Bibliografia

- Apanowicz, J. (2002). *Metodologia ogólna*. Gdynia: Bernardinum.
- Babiński, G. (1984). Pytania kwestionariuszowe: podstawowe podziały i typologie. W: J. Wasilewski (red. nauk.), *Wybrane zagadnienia metodologiczno-teoretyczne badań socjologicznych*. Kraków: Uniwersytet Jagielloński.
- Cape, P., Philips, K. (2015). Questionnaire Length and Fatigue Effects: The Latest Thinking and Practical Solutions, *Survey Sampling International*. Pobrane z: <https://www.surveysampling.com/site/assets/files/1586/questionnaire-length-and-fatigue-effects-the-latest-thinking-and-practical-solutions.pdf>.
- Durrant, G. B., Groves, R. M., Staetsky, L., Steele, F. (2010). Effects of Interviewer Attitudes and Behaviors on Refusal in Household Surveys. *Public Opinion Quarterly*, 74 (1), s. 1–36. DOI: 10.1093/poq/nfp098.
- Gallup, G., Rae, S. (1940). *The Pulse of Democracy*. New York: Simon & Schuster.
- Groves, R. M. (2006). Nonresponse Rates and Nonresponse Bias in Household Surveys. *The Public Opinion Quarterly*, 70 (5), s. 646–675. DOI: 10.1093/poq/nfl033.
- Gruszczyński, L. A. (2003). *Kwestionariusze w socjologii. Budowa narzędzi surveyowych*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.
- Hox, J. J., De Leeuw, E. D. (2002). The Influence of Interviewers' Attitude and Behavior on Household Survey Nonresponse: an International Comparison. W: R. M. Groves, D. A. Dillman, J. L. Eltinge, R. J. A. Little (red. nauk.) *Survey Nonresponse*. New York: Wiley.
- Hyman, M., Sierra, J. (2016). Open- versus close-ended survey questions. *NMSU Business Outlook*, 14 (2), s. 1–5.
- IMAS International Poland. (2009). *Researchers create long questionnaires, but respondents expect that telephone opinion polls will last only a few minutes*. Pobrane z: <http://imas.pl/en/wp-content/uploads/sites/2/2013/06/Expected-length-of-the-telephone-interview.pdf>.
- Jäckle, A., Lynn, P., Sinibaldi, J., Tipping, S. (2013). The Effect of Interviewer Experience, Attitudes, Personality and Skills on Respondent Co-operation with Face-to-Face Surveys. *Survey Research Methods*, 7 (1), s. 1–15. DOI: 10.18148/srm/2013.v7i1.4736.
- John, L. K., Loewenstein, G., Acquisti, A., Vosgeraud, J. (2018). When and why randomized response techniques (fail to) elicit the truth. *Organizational Behavior and Human Decision Processes*. DOI: 10.1016/j.obhdp.2018.07.004.
- Johnson, E. P., Siluk, L., Taraf, S. (2014). Techniques for fusing survey modules: Respondent matching and data imputation. *International Journal of Market Research*, 56 (4), s. 425–442. DOI: 10.2501/IJMR-2014-031.
- Kahneman, D., Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47 (2), s. 263–292. DOI: 10.2307/1914185.
- Krok, E. (2015). Budowa kwestionariusza ankietowego a wyniki badań. *Zeszyty Naukowe Uniwersytetu Szczecińskiego*, 874 (37), s. 55–73. DOI: 10.18276/si.2015.37-05.
- Krzewińska, A. (2006). Psychologia poznawcza Norberta Schwartza i jej zastosowanie w warunkach polskich. Prezentacja wyników badania własnego. *Analizy i próby technik badawczych w socjologii*, T. 11. Łódź: Łódzkie Towarzystwo Naukowe.
- Lazarsfeld, P. F. (1935). The Art of Asking WHY in Marketing Research: Three Principles

- Underlying the Formulation of Questionnaires. *National Marketing Review*, 1 (1), s. 26–38.
- Lazarsfeld, P. F. (1944). The Controversy over Detailed Interviews: An Offer for Negotiation. *Public Opinion Quarterly*, 8 (1), s. 38–60. DOI: 10.1086/265666.
- vol. 8, s. 38–60.
- Lester, A., Wilson, I. (1995). *Surveying businesses by telephone – a case study of methodology*. Materiały International Conference on Survey Measurement and Process Quality, American Statistical Association.
- Lutyńska, K. (1978). Ankieterzy i badacze: z badań nad wpływem ankieterskim. *Przegląd Socjologiczny*, 30, s. 143–173.
- Lutyńska, K. (1984). *Wywiad kwestionariuszowy. Przygotowanie i sprawdzanie narzędzia badawczego*. Wrocław: Ossolineum.
- Lyberg, L., Dean, P. (1992). *Methods for Reducing Nonresponse Rates: A Review*. Annual Meeting of the American Association for Public Opinion Research. St. Petersburg.
- Międzynarodowy Kodeks Badań Rynku i Badań Społecznych ICC/ESOMAR. (2008). Pobrane z: [https://www.ptbriio.pl/files/Miedzynarodowy_Kodeks_Badan_Rynku_i_Badan_Sp HYPHERLINK „https://www.ptbriio.pl/files/Miedzynarodowy_Kodeks_Badan_Rynku_i_Badan_Spolecznych.pdf”olecznych.pdf](https://www.ptbriio.pl/files/Miedzynarodowy_Kodeks_Badan_Rynku_i_Badan_Sp_HYPERLINK_„https://www.ptbriio.pl/files/Miedzynarodowy_Kodeks_Badan_Rynku_i_Badan_Spolecznych.pdf”olecznych.pdf).
- Morton-Williams, J. (1993). *Interviewer Approaches*. Dartmouth: Aldershot.
- Nachmias, V., Nachmias D. (2001). *Metody badawcze w naukach społecznych*. Warszawa: Wydawnictwo Zysk i S-ka.
- National Opinion Research Center. (1953). *Isolation, Measurement, and Control of Interviewer Effect. A Systematic Study of Sources of Error in the Empirical Study of Attitudes, Opinions, and Other Aspects of Human Behavior*. Chicago: University of Chicago.
- Sawiński, Z., Sztabiński, P.B., Sztabiński, F. (2000). (red. nauk.). *Podręcznik ankietera*. Warszawa: IFiS PAN.
- Stevens, S. S. (1946). On the Theory of Scales of Measurement. *Science*, 103 (2684), s. 677–680.
- Stoop, I., Billiet, J., Koch, A., Fitzgerald, R. (2010). *Improving Survey Response. Lessons learned from the European Social Survey*. Chichester: Wiley.
- Vicente, P., Reis, E., Santos, M. (2009). Using mobile phones for survey research. A comparison with fixed phones. *International Journal of Market Research*, 51 (5), s. 613–634. DOI:10.2501/S1470785309200852.
- Waksberg, J. (1987). Sampling Methods for Random Digit Dialing. *Journal of the American Statistical Association*, 73 (361), s. 40–46. DOI: 10.1080/01621459.1978.10479995.
- Zagańczyk, A. (2014). Zasady konstruowania kwestionariusza ankiety. *BiTP*, 33 (1), s. 37–43.