

**Radosław Mącik**

Uniwersytet Marii Curie-Skłodowskiej w Lublinie

**Małgorzata Korba**

Zakład Ubezpieczeń Społecznych Oddział w Lublinie

---

## **WIARYGODNOŚĆ POMIARU W BADANIACH *MIXED-MODE*: PORÓWNANIE EFEKTÓW STOSOWANIA PAPI I CAWI**

---

**Streszczenie:** W artykule omówiono wyniki porównania dwu niezależnych badań wybranych efektów metodologicznych zastosowania kwestionariusza ankiety internetowej (CAWI) oraz klasycznej ankiety papierowej (PAPI). Na tle przeglądu literatury zaprezentowano wyniki autorskich badań porównawczych w zakresie rzetelności pomiaru z użyciem skal Likerta i semantycznej, a także analizy kompletności odpowiedzi pytań otwartych. Wnioskiem z przeprowadzonej analizy jest porównywalna jakość pomiaru za pomocą obu sposobów zbierania danych – skale są wystarczająco rzetelne, a odpowiedzi na pytania – otwarte bogate treściowo. Techniczne możliwości CAWI, szczególnie rotacja pozycji skal i wymuszanie udzielenia odpowiedzi, mogą wpływać na uzyskiwanie danych o lepszej jakości za pomocą ankiet internetowych (przy porównywalnym sposobie doboru próby i adekwatnej jej strukturze).

**Słowa kluczowe:** wiarygodność pomiaru, *mixed-mode*, PAPI, CAWI.

### **1. Wstęp**

Wobec spadających stóp zwrotu kwestionariuszy ankietowych dla zwiększenia stopnia realizacji próby coraz częściej stosowane jest tzw. podejście *mixed-mode*, tj. dostarczania kwestionariusza respondentowi w różny sposób (np. tradycyjnie pocztą lub bezpośrednio: przez telefon lub Internet), po to, by mógł wybrać najwygodniejszy sposób.

Takie podejście może budzić jednak pewne wątpliwości metodologiczne, szczególnie dotyczące możliwości agregacji danych zebranych w różny sposób. Niniejsza praca zawiera zwięzłą analizę literatury w tym zakresie i prezentuje, jako studia przypadków, porównanie miar i wskaźników stosowanych do oceny wiarygodności pomiaru otrzymanych w dwu niezależnych studiach wykonanych technikami PAPI i CAWI: jednym – na dwu równoległych dużych próbach o zróżnicowanej strukturze demograficznej – oraz drugim – na dwu równoległych małych próbach jednolitych demograficznie.

## 2. Podejście *mixed-mode* w badaniach kwestionariuszowych

Podejście *mixed-mode* (nazywane również *multiple-mode*) w badaniach kwestionariuszowych polega na łącznym zastosowaniu różnych sposobów zapewnienia kontaktu z badanym w ramach jednego projektu badawczego, np. ankiety pocztowej, internetowej (CAWI), samodzielnie wypełnianej na papierze (PAPSI) lub wywiadu ankieterskiego – tradycyjnego (PAPI) albo wspomaganego komputerowo (CAPI), czy też wywiadu telefonicznego (CATI) [17]. Wśród technik używanych w tym podejściu łączone są często sytuacje pomiaru bez udziału ankietera z sytuacjami wywiadu, przy czym relatywnie często stosowany zestaw technik to PAPI/CAPI, CAWI i CATI.

Procedury oparte na podejściu *mixed-mode* mogą prowadzić do:

- równoległego zbierania danych za pomocą różnych sposobów (najczęściej zgodnie z preferencją osoby badanej, której pozwala się dokonać wyboru sposobu dostarczenia kwestionariusza, co *nota bene* według D.A. Dillmana może skutkować taką samą stopą partycypacji w badaniu jak przy użyciu pojedynczej metody, ale dawać badanemu większą satysfakcję wynikającą z możliwości wyboru dogodnej dla siebie opcji [8, s. 3];
- sekwencyjnej procedury kontaktów z respondentami, w której przejście do kolejnego sposobu dostarczenia kwestionariusza wynika z niepowodzenia zastosowania bardziej preferowanych przez badacza technik, np. najpierw ankiety pocztowej, następnie internetowej, a w ostateczności wywiadu telefonicznego, w którym to przypadku łatwiej o większe stopy zwrotu kwestionariuszy [8, s. 3-4; 2, s. 248].

Głównym celem zastosowania podejścia *mixed-mode* jest chęć otrzymania danych o lepszej jakości – głównie z powodu uzyskiwania lepszych stóp realizacji próby [16, s. 70] – i po niższym koszcie niż z użyciem pojedynczej techniki zbierania danych.

Należy pamiętać, że każda ze wspomnianych technik ma swoje immanentne zalety i wady, a jej użycie może prowadzić do uzyskania różnych efektów, np. z powodu obecności ankietera lub jej braku, różnego stopnia skupienia uwagi w trakcie udziału w badaniu, a także wynikające z różnego formatowania pytań i różnych konwencji ich zadawania w określonych technikach. Należy zdawać sobie sprawę, iż wykorzystanie różniących się technik pomiaru oznaczać będzie konieczność poradzenia sobie przynajmniej z częścią z tych problemów.

Dostępna literatura na temat metodologicznych uwarunkowań zastosowania podejścia *mixed-mode* w badaniach kwestionariuszowych nie jest zbyt obszerna (przeгляд dostępnych autorowi bibliograficznych baz danych oraz tzw. kwalifikowanych źródeł internetowych pozwala zidentyfikować najwyżej kilkadziesiąt dość zróżnicowanych tematycznie prac o charakterze metodologicznym, w tym tylko pojedyncze przykłady pozycji książkowych, np. [9; 6]), mimo że zainteresowanie tą tematyką sięga połowy lat 90. XX w. Obejmuje ona przede wszystkim artykuły naukowe

o charakterze eksperymentalnym, których autorzy porównują konsekwencje metodologiczne użycia wspomnianego podejścia w konkretnym zastosowaniu, najczęściej porównując ze sobą efekty łączenia 2 lub 3 sposobów zbierania danych, np. prace takich autorów, jak: Cabanoglu, Warde i Moreo [3], Meckel, Walters i Baugh [15], Bäckström i Nilsson [1], Wygant, Olsen, Call i Curtin [19] czy też – przede wszystkim – D.A. Dillmana i jego współpracowników [8; 9].

W celu uniknięcia przynajmniej części problemów związanych z błędami pomiaru w sytuacjach stosowania podejścia *mixed-mode* D.A. Dillman zaproponował zunifikowane zasady konstrukcji kwestionariuszy w każdym ze sposobów zbierania danych, nazywając je podejściem *uni-mode* [6, s. 232-240], m.in. zwracając uwagę na konieczność:

- posługiwania się tymi samymi wariantami odpowiedzi dla pytań i skal we wszystkich sposobach zbierania danych,
- unikania zmian podstawowej struktury pytania pomiędzy wersjami (np. kafeteria w wersji papierowej zamieniona na listę rozwijalną lub szereg pytań skalowanych „tak/nie” w badaniu internetowym),
- ograniczenia (np. do 5 lub 7) kategorii odpowiedzi dla skal złożonych, przy czym wyraźnie zalecane jest stosowanie skal ze słownym opisem wszystkich kategorii odpowiedzi, a nie jedynie końców skali,
- posługiwania się równoważnymi instrukcjami sterującymi pomiarem, szczególnie warunkowymi przejściami do innych miejsc kwestionariusza,
- posługiwania się przetestowanymi i równoważnymi instrukcjami dla ankierów w sposobach zbierania danych wymagających ich obecności.

### **3. Porównanie głównych cech kwestionariuszy papierowych (PAPI) i internetowych (CAWI)**

Pomiędzy tradycyjnymi ankietami papierowymi a ankietami internetowymi można wskazać wiele różnic dotyczących: zwrotności, czasu badania, ponoszonych kosztów, możliwości uzyskania reprezentatywności, jakości uzyskanych danych, możliwości zastosowania udogodnień w projektowaniu, zbieraniu informacji oraz analizie danych.

Dla badania kwestionariuszem internetowym (CAWI) losowy dobór próby narzęcza pewnych trudności, ponieważ populacja internautów różni się znacząco od populacji generalnej (nadal z Internetu korzystają głównie ludzie młodzi, lepiej wykształceni i zamożniejsi), a ponadto nie istnieje wykaz populacji internautów, chyba że badanie dotyczy szczególnej zawężonej populacji, jak grupa pracowników danego przedsiębiorstwa, członków pewnej społeczności internetowej itd. Liczba możliwych do zbadania segmentów stale jednak rośnie. Dla ankiety papierowej problem reprezentatywności istnieje w zasadzie jedynie w kontekście właściwego doboru próby do określonego badania – dostępnych segmentów populacji jest znacznie więcej, choć mają ograniczony geograficznie zasięg. Ankieta internetowa ma nad

kwestionariuszem papierowym tę przewagę, że jest z nią łatwiej dotrzeć do grupy osób o szczególnych, trudnodostępnych cechach, np. populacji osób transseksualnych czy hodujących rzadko spotykane zwierzęta. Ankiety internetowe charakteryzują się w zasadzie nieograniczonym zasięgiem, co różni je od ankiet bezpośrednich z papierowym kwestionariuszem (PAPI).

Drugim bardzo istotnym czynnikiem różniącym oba wspomniane sposoby zbierania danych jest czas realizacji projektu badawczego. Ankieta internetowa cechuje się zdecydowanie szybszą możliwością zebrania wyników badania niż ankiety papierowe, nawet w ciągu doby [18, s. 105]. Zwykle mija około 1-2 tygodni zanim uda się zebrać wystarczającą liczbę wypełnionych kwestionariuszy, jeżeli nie są stosowane żadne przypomnienia ani dodatkowe zachęty do wypełnienia ankiety internetowej, natomiast przy zastosowaniu motywatorów ten czas skraca się do 3-4 dni [19]. Dla ankiety internetowej faza drukowania, powielania i dystrybucji ankiet do respondenta nie istnieje. W przypadku Internetu, by opublikować przygotowany kwestionariusz, wystarczy jedno kliknięcie. Oszczędność czasu daje również automatyczne kodowanie w systemie CAWI. Ankieta internetowa ma też dla respondenta tę zaletę, że nie wymaga odpowiedzi natychmiastowej, przez co badany wypełnia ją w dogodnym dla siebie czasie i może jej poświęcić więcej uwagi. Zauważono jednak, że z aktywnością w sieci wiąże się krótszy czas koncentracji na jednej czynności: „uczestnicy badań internetowych wykazują mniej cierpliwości niż respondenci w badaniach prowadzonych bezpośrednio lub przez telefon i często tracą zainteresowanie badaniem po 25-30 pytaniach. Czas potrzebny na udzielenie odpowiedzi na te pytania nie powinien przekraczać 16 minut” [13, s. 156].

Wywiad kwestionariuszowy daje badaczowi bezpośredni kontakt z ankietowanym, w przeciwieństwie do badania internetowego. Z bezpośrednim kontaktem z respondentami wiąże się jednak pewna niedogodność – korzystanie z pracy ankietatorów sprawia, że badacz nie ma pełnej kontroli nad przebiegiem badania. Taką kontrolę daje natomiast ankieta internetowa [12]. Badanie w sieci pozwala również szybko zareagować na nadesłaną ankietę, np. wysłać podziękowanie za wypełniony kwestionariusz.

Kolejną istotną kwestią w porównaniach CAWI i PAPI jest możliwa do uzyskania stopa zwrotu kwestionariuszy. Niestety kilkanaście lat badań nie przyniosło jednoznacznej odpowiedzi na pytanie, która z metod gwarantuje większą stopę zwrotu: „Niektóre ankiety internetowe osiągały zwrotność na poziomie aż 87 procent, inne z kolei były w stanie zapewnić dramatycznie niski procent zwrotu” [12], czasem jedynie kilkuprocentowy. Część badaczy twierdzi, że „kolejne ankiety internetowe charakteryzują się coraz niższym poziomem zwrotności, zwłaszcza wtedy, gdy badania mają charakter ogólny. Nieco lepsze rezultaty przynoszą internetowe badania profilowane” [13, s. 175].

Wskazuje się, że jakość zebranych danych, określana ich ilością, kompletnością oraz zawartością informacyjną, jest wyższa dla badań internetowych [14, s. 107-111]. Dzieje się tak z kilku powodów. Kwestionariusz internetowy ma mechanizmy

redukujące błędy takie, jak dodatkowe instrukcje, automatyczne sprawdzanie poprawności danych, ogranicza też obciążenie odpowiedzi wpływem ankietera. Badanie zaprojektowane w sieci pozwala też zindywidualizować kwestionariusz, a co za tym idzie – pozyskać od respondenta więcej szczegółowych informacji, nie obciążając go pytaniami niepotrzebnymi. Za pomocą pytań filtrujących i przejść logicznych pomiędzy poszczególnymi ekranami kwestionariusza badacz może kierować respondenta do pytań, na które będzie on w stanie udzielić wyczerpującej odpowiedzi dającej użyteczne informacje.

Ankiety internetowe dają również szansę dodawania do kwestionariusza nie tylko obrazów (co jest możliwe i w ankiecie papierowej, choć jest to kosztowne), ale także dźwięków, filmów i prezentacji. Ankiety CAWI umożliwiają automatyczne rotowanie odpowiedzi, co zmniejsza ich obciążenie i utrudnia zniekształcanie wyników badania respondentom, którzy bezrefleksyjnie udzielają odpowiedzi na pytania kafeteryjne (zaznaczając np. za każdym razem pierwszą odpowiedź, by jak najszybciej skończyć badanie). W PAPI natomiast papier, jako nośnik stały, nie pozwala na rotowanie odpowiedzi, ponieważ generowałoby to zbyt wysokie koszty (przygotowanie wielu wersji tego samego kwestionariusza, bardziej pracochłonne kodowanie danych). Użyte nieumiejętnie instrumenty zastosowane w ankiecie internetowej mogą jednak powodować powstawanie dodatkowych błędów. Tak dzieje się na przykład, gdy badacz nie zatroszczy się o to, by rozwijana lista odpowiedzi miała w pierwszym widocznym dla użytkownika polu informację neutralną, a nie jedną z odpowiedzi.

Podnoszonym często argumentem za intensyfikacją badań internetowych są niskie koszty ich prowadzenia. Najbardziej obie techniki różnicują koszty zmienne. Wśród nich dla PAPI należy wyróżnić przede wszystkim relatywnie kosztowne: druk kwestionariuszy, wynagrodzenie dla ankieterów, prace koordynacyjne oraz koszt kodowania odpowiedzi. Ankiety internetowe generują natomiast niewielkie koszty zmienne, dotyczą one przede wszystkim ewentualnych zachęt do wypełnienia kwestionariusza, np. umieszczenia bannerów odsyłających do ankiety na wybranych witrynach internetowych, kuponów rabatowych lub darmowych produktów i usług udostępnianych w zamian za wypełnienie ankiety. Ankiety internetowe wymagają natomiast poniesienia wyższych kosztów stałych niż w tradycyjnym podejściu PAPI. Wstępne analizy polskich badaczy pozwoliły określić próg graniczny opłacalności różnych technik ankietowych. Wynika z nich, że „badania, w których uczestniczy więcej niż 60 respondentów, bardziej opłaca się prowadzić w internecie niż z wykorzystaniem [tradycyjnej – dop. aut.] poczty (...). Porównanie kosztów ankiety pocztowej i ankiety rozsyłanej pocztą elektroniczną, również z uwzględnieniem dodatkowego kontaktu, wskazuje, że próg opłacalności wynosi około 15 (...). Progi opłacalności ankiet internetowych ukształtowały się zatem na stosunkowo niskim poziomie” [13, s. 281].

## 4. Wiarygodność pomiaru w badaniach PAPI i CAWI – przypadek 1

Tradycyjnie, w sposób uproszczony, wiarygodność pomiaru określa się przez pryzmat trafności i rzetelności pomiaru. Porównując wiarygodność pomiaru w tym przypadku posłużono się dwoma dużymi, ogólnopolskimi, równoległymi próbami, o praktycznie takiej samej charakterystyce demograficznej respondentów, w których dane zbierano na podstawie tego samego kwestionariusza wykonanego odpowiednio jako CAWI (próba 1,  $n = 1100$ ) i PAPI (próba 2,  $n = 1067$ ).

Ze względu na ograniczenia objętościowe niniejszego opracowania zaprezentowano szczegółowo wyniki oceny zgodności wewnętrznej poszczególnych skal złożonych (typu Likerta oraz semantycznej) jako sposobu określenia rzetelności pomiaru. Przeprowadzone wcześniej eksploracyjne analizy czynnikowe dały w obu przypadkach praktycznie identyczne rezultaty, pozwalając wyciągnąć wniosek o wysokiej trafności czynnikowej skal w obu przypadkach. Ponadto dokonano oceny kompletności odpowiedzi na zawarte w kwestionariuszach pytania otwarte.

Otrzymane rezultaty wskazują, że w sytuacji nierotowania kolejności stwierdzeń w skalach złożonych w wersji CAWI (by upodobnić pomiar do PAPI) uzyskiwano podobną zgodność wewnętrzną skal składowych (tab. 1-3). Dla skal Likerta na 15 analizowanych skal w 7 przypadkach nieznacznie lepszą rzetelność uzyskano, używając PAPI, w 2 sytuacjach uzyskano identyczne rezultaty, a w 6 – nieco lepszą rzetelność dawało użycie CAWI. Natomiast dla skali semantycznej (porównano trzykrotne użycie tej samej skali w różnych kontekstach) w 3 na 3 przypadkach użycie CAWI pozwalało na uzyskanie lepszej rzetelności.

**Tabela 1.** Zgodność wewnętrzna skal składowych modelu TAM/TTF<sup>a</sup> (skala typu Likerta)

Konstrukt	Liczba pozycji	$\alpha$ -Cronbacha	
		ankieta internetowa (CAWI)	ankieta papierowa (PAPI)
Postrzegana użyteczność technologii	5	0,82	0,84
Postrzegana łatwość użycia technologii	3	0,79	0,84
Intencja użycia technologii	4	0,68	0,70
Faktyczne użycie technologii	5	0,78	0,75
Zgodność technologii z celem użycia	7	0,90	0,89

<sup>a</sup> Model TAM/TTF to model akceptacji technologii autorstwa F.D. Davisa [4], wzbogacony o skalę TTF – zgodności technologii z celem użycia, zaproponowaną przez Goodhue i Thompsona [10]. W badaniu posługiwano się własną adaptacją skal składowych modelu.

**Tabela 2.** Zgodność wewnętrzną skal składowych kwestionariusza PCS<sup>a</sup> (skala typu Likerta)

Konstrukt	Liczba pozycji	$\alpha$ -Cronbacha	
		ankieta internetowa (CAWI)	ankieta papierowa (PAPI)
Styl perfekcjonistyczny	3	0,67	0,64
Styl wrażliwości na markę	3	0,70	0,70
Styl wrażliwości na modę/nowości	3	0,79	0,80
Styl hedonistyczny/rekreacyjny	3	0,85	0,84
Styl wrażliwości na cenę	3	0,65	0,66
Styl impulsywny	3	0,61	0,62
Styl wrażliwości na nadmiar informacji	3	0,70	0,71
Styl rutynowy, lojalny markom/miejscom zakupu	3	0,71	0,70
Styl kompulsywny	3	0,80	0,80

<sup>a</sup> PCS (*Profile of Consumer Styles*) jest krótkim kwestionariuszem diagnozującym profil stylów podejmowania decyzji zakupowych przez konsumenta opracowanym przez Sproles i Kendall [16]. W badaniu posługiwano się własną rekonstrukcją kwestionariusza.

Źródło: badania własne.

Dodatkowo zwracano uwagę na trafność danych uzyskanych z pytań otwartych, dokonując analizy treści, ale też uwzględniając proste wskaźniki jakości uzyskanych danych w postaci odsetka braków odpowiedzi (udzielenie odpowiedzi nie było wymuszane) oraz średniej długości odpowiedzi w teście zdań niedokończonych oraz pytaniach otwartych o strukturze wskazania rankingu 3 czynników (por. tab. 4).

**Tabela 3.** Zgodność wewnętrzną pozostałych skal złożonych (przypadek 1)

Skala	Typ skali	Liczba pozycji	$\alpha$ -Cronbacha	
			ankieta internetowa (CAWI)	ankieta papierowa (PAPI)
Postrzeganie zakupów w sklepach konwencjonalnych (P4)	semantyczna <sup>a</sup>	11	0,62	0,62
Postrzeganie zakupów w sklepach internetowych (P11)	semantyczna <sup>a</sup>	11	0,76	0,75
Postrzeganie zakupów na aukcjach internetowych (P12)	semantyczna <sup>a</sup>	11	0,78	0,76
Samoocena zmian w zachowaniach zakupowych pod wpływem technologii internetowych (P22)	Likerta	13	0,83	0,70

<sup>a</sup> Stosowano te same określenia końców skali – używano wersji dwubiegunowej określeń.

Źródło: badania własne.



**Tabela 4.** Wybrane charakterystyki odpowiedzi na pytania otwarte (przypadek 1)

Zmienna	Ankieta internetowa (CAWI)		Ankieta papierowa (PAPI)		Uwagi
	% braków odpowiedzi	średnia długość wypowiedzi (znaków)	% braków odpowiedzi	średnia długość wypowiedzi (znaków)	
P22_1	29,4	14,2	3,9	14,8	test zdań niedokończonych
P22_2	31,5	15,9	8,0	16,1	
P22_3	34,0	22,6	12,9	21,7	
P22_4	28,7	16,1	4,4	14,8	
P22_5	36,1	19,3	16,1	17,7	
P22_6	35,5	19,3	12,7	19,8	
P22_7	32,0	22,0	23,9	19,4	
P23_1	32,4	15,2	14,9	15,4	pierwsza cecha
P23_2	50,0	16,4	50,7	14,1	druga cecha
P23_3	63,3	17,0	77,6	13,7	trzecia cecha
P23_4	32,8	21,8	20,2	20,3	pierwsza cecha
P23_5	60,5	23,0	80,7	18,6	druga cecha
P23_6	73,6	20,4	93,4	18,9	trzecia cecha
P23_7	34,1	16,6	26,3	16,0	pierwsza cecha
P23_8	60,1	16,5	69,3	13,2	druga cecha
P23_9	71,8	15,5	87,6	14,5	trzecia cecha
P23_10	33,7	20,0	30,2	21,3	pierwsza cecha
P23_11	62,3	21,3	80,2	18,7	druga cecha
P23_12	73,2	20,8	92,7	18,8	trzecia cecha

Źródło: badania własne.

W teście zdań niedokończonych CAWI dało większą liczbę braków odpowiedzi przy podobnej lub większej długości wypowiedzi udzielonych. Natomiast w rankingu cech uzyskiwano w CAWI mniejszą liczbę braków odpowiedzi dla 2 i 3 pozycji w rankingu, a większą – dla pozycji 1, przy czym wypowiedzi były konsekwentnie dłuższe w CAWI.

Porównując uzyskane w obu próbach (przypadek 1) wskaźniki świadczące o rzetelności i trafności pomiaru, należy zauważyć, iż pomiar za pomocą ankiety internetowej okazał się w wystarczającym stopniu ekwiwalentny w stosunku do klasycznych technik wykorzystujących kwestionariusz w kwestii zarówno zgodności wewnętrznej, jak i trafności czynnikowej. Podobnie należy ocenić ekwiwalencję pomiaru dla pytań otwartych.



## 5. Wiarygodność pomiaru w badaniach PAPI i CAWI – przypadek 2

W tym przypadku potencjalne różnice między wynikami otrzymywanymi za pomocą CAWI i PAPI miało wskazać badanie na dwóch równoległych małych próbach studentów (o liczebności odpowiednio: CAWI:  $n = 41$  i PAPI:  $n = 39$ ) o jednolitej charakterystyce demograficznej pod względem wieku, płci, miejsca zamieszkania, a także poziomu umiejętności komputerowych, częstotliwości korzystania z Internetu i sytuacji materialnej. Obie grupy wypełniały kwestionariusz o takiej samej treści, przy czym pierwsza wypełniała kwestionariusz dystrybuowany na papierze, a druga – kwestionariusz wypełniany drogą internetową (z rotowaniem stwierdzeń w skalach i wymuszeniem udzielenia odpowiedzi). Podstawowe różnice zaobserwowane między grupami dotyczyły rzetelności zastosowanych skal, długości odpowiedzi na pytania otwarte i ich oryginalności.

Istotą opisywanego badania nie było stworzenie skal o najwyższej możliwej rzetelności, ale porównanie rzetelności między tymi samymi skalami użytymi w kwestionariuszu CAWI i PAPI. Wyniki dotyczące rzetelności poszczególnych skal (dyferencjału semantycznego i skali Likerta) zawiera tab. 5 – w tym przypadku rzetelność mierzona współczynnikiem  $\alpha$ -Cronbacha była we wszystkich 4 przypadkach wyraźnie wyższa dla ankiety internetowej. Potencjalnym czynnikiem wywołującym większą rzetelność skal była w tej sytuacji rotacja stwierdzeń skali.

**Tabela 5.** Zgodność wewnętrzną skal złożonych (przypadek 2)

Skala	Typ skali	Liczba pozycji	$\alpha$ -Cronbacha	
			ankieta internetowa (CAWI)	ankieta papierowa (PAPI)
Odczuwane korzyści z uczestnictwa w społecznościach internetowych	Likerta	4	0,604	0,305
Odczuwane niedogodności z uczestnictwa w społecznościach internetowych	Likerta	4	0,584	0,141
Pozytywne cechy Internetu	semantyczna	4	0,575	0,328
Negatywne cechy Internetu	semantyczna	5	0,561	– <sup>a</sup>

<sup>a</sup> Uzyskano wartość ujemną dla PAPI – prawdopodobnie badani przez PAPI w przypadku jednego ze stwierdzeń konsekwentnie nie zauważali właściwego kierunku skalowania.

Źródło: badania własne.

**Tabela 6.** Wybrane charakterystyki odpowiedzi na pytania otwarte (przypadek 2)

Pytanie	Statystyka	Liczba słów			Liczba znaków		
		CAWI	PAPI	różnica (%)	CAWI	PAPI	różnica (%)
Pytanie 1	Mediana (znaki) Średnia (słowa)	5,6	5,3	+ 6	41	34	+ 21
	Maksimum	30	12	+ 150	242	95	+ 155
	Minimum	1	1	–	9	10	– 10
Pytanie 2	Mediana (znaki) Średnia (słowa)	15,7	15,0	+ 5	94	70	+ 34
	Maksimum	66	34	+ 94	397	217	+ 83
	Minimum	2	2	–	10	9	+ 11

Źródło: badania własne.

W odniesieniu do użytych w przypadku 2 pytań otwartych między odpowiedziami respondentów w CAWI i PAPI nie ma znaczącej różnicy w kwestii liczby użytych słów. Co jednak warto zauważyć, dla ankiety internetowej w obu pytaniach są to wartości wyższe. W pytaniach otwartych zawartych w kwestionariuszu na stronie WWW można też stwierdzić większe zróżnicowanie długości odpowiedzi – najwyższe wartości dla kwestionariuszy w sieci przewyższają maksima długości odpowiedzi w kwestionariuszu ankiety bezpośredniej nawet 2,5 razy. Wyższe różnice w liczbie liter odpowiedzi poszczególnych respondentów pozwalają przypuszczać, że słownictwo w odpowiedziach z CAWI jest bardziej złożone niż w wypowiedziach na papierze [5].

Stwierdzenia w pytaniach otwartych w kwestionariuszu internetowym cechują się większą oryginalnością. Respondenci wypełniający kwestionariusz CAWI zapisali więcej unikatowych stwierdzeń w odpowiedzi na pytanie o udostępnianie innym internautom plików chronionych prawami autorskimi. W ankiecie papierowej padło około 7 oryginalnych stwierdzeń, podczas gdy w kwestionariuszu internetowym ankietowani zawarli około 9 unikalnych opinii. Oznacza to, że pytania otwarte w CAWI dają w efekcie większą wartość informacyjną.

## 6. Główne wnioski

Podsumowując różnice wiarygodności pomiaru między CAWI i PAPI, można wskazać kilka wniosków – pamiętając jednak, że prezentowane w pracy wyniki nie powinny być uogólniane, ponieważ bazują na studiach przypadków.

Przedstawione równoległe dane z dwu różnych projektów badawczych pozwoliły potwierdzić tezę, że ta sama skala w badaniu CAWI i PAPI nie musi się charakteryzować taką samą wiarygodnością. O ile dla ocen trafności czynnikowej sposób

kontakty z respondentem (CAWI vs. PAPI) nie miał znaczenia, to ogólnie wyższą rzetelnością cechowały się skale w ankiecie internetowej.

Warto też zwrócić uwagę na wyniki dotyczące długości i oryginalności odpowiedzi na pytania otwarte. Badanie wskazuje na większą użyteczność informacji zebranych za pomocą kwestionariusza w sieci. Odpowiedzi grup posługujących się CAWI są dłuższe, bardziej złożone i zawierają bardziej zróżnicowane treści.

Po przeanalizowaniu kilku czynników związanych z różnicami między kwestionariuszem internetowym a papierowym można stwierdzić, że wiarygodność pomiaru za pomocą CAWI często jest co najmniej taka sama, jeśli nie lepsza niż w PAPI zarówno dla skal złożonych, jak i pytań otwartych.

## Literatura

- [1] Bäckström C., Nilsson C., *Mixed Mode Surveying. A Comparison of Paper-questionnaires And Web-questionnaires*, 2002, <http://www.modsurvey.org/text/mixedmodesurveying.pdf>, 26.06.2009.
- [2] Biemer P.P., Groves R.M., Lyberg L.E., *Measurement Errors in Surveys*, Wiley, 2004.
- [3] Cobanoglu C., Warde B., Moreo P.J., *A comparison of mail, fax and web-based survey methods*, "International Journal of Market Research" 2001, vol. 43, no 4, s. 441-452.
- [4] Davis F.D., *Perceived usefulness, perceived ease of use, and user acceptance of information technology*, "MIS Quarterly" 1989, vol. 13, no 3, s. 319-339.
- [5] Descombe M., *Open-ended questions and online surveys: the mode effect in relation to length*, 7<sup>th</sup> International Conference on Social Science Methodology 2008, <http://www.rc332008.unina.it/index.html>, 25.06.2009.
- [6] Dillman D.A., *Mail and Internet Surveys: The Tailored Design Method*, 2<sup>nd</sup> Edition, John Wiley & Sons, New York 2007.
- [7] Dillman D.A., *Mail and Other Self-Administered Surveys in the 21st Century: The Beginning of a New Era*, 1998, <http://survey.sesrc.wsu.edu/dillman/papers/svys21st.pdf>, 1.12.2009.
- [8] Dillman D.A., Phelps G., Tortora R., Swift K., Kohrell J., Berck J., *Response Rate and Measurement Differences in Mixed Mode Surveys* (Draft Paper), 2001, [http://survey.sesrc.wsu.edu/dillman/papers/Mixed%20Mode%20ppr%20\\_with%20Gallup\\_%20POQ.pdf](http://survey.sesrc.wsu.edu/dillman/papers/Mixed%20Mode%20ppr%20_with%20Gallup_%20POQ.pdf), 2.12.2009.
- [9] Dillman D.A., Smyth J.D., Christian L.M., *Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method*, 3<sup>rd</sup> ed., John Wiley & Sons, New York 2008.
- [10] Goodhue D.L., Thompson R.L., *Task-technology fit and individual performance*, "MIS Quarterly" 1995, vol. 19, no 2, s. 213-236.
- [11] Haysletta M.M., Wildemuth B.M., *Pixels or pencils? The relative effectiveness of Web-based versus paper surveys*, "Library & Information Science Research" 2004, vol. 26, s. 73-93.
- [12] Kaniewska-Sęba A., Leszczyński G., Pilarczyk B., *Badania marketingowe na rynku business-to-business*, Oficyna Ekonomiczna, Kraków 2006.
- [13] Karcz K., Bajdak A. (red.), *Badania marketingowe w internecie*, AE, Katowice 2005.
- [14] Maćcik R., *Wykorzystanie Internetu w badaniach marketingowych*, UMCS, Lublin 2005, s. 107-111.
- [15] Meckel M., Walters D., Baugh P., *Mixed-mode Surveys Using Mail and Web Questionnaires*, "The Electronic Journal of Business Research Methodology" 2005, vol. 3, issue 1, s. 69-80.
- [16] Sproles G.B., Kendall E.L., *A methodology for profiling consumer decision-making styles*, "The Journal of Consumer Affairs" 1986, vol. 20, no 2 (Winter), s. 267-279.
- [17] SPSS Inc., *Survey Research in a Multi-Modal World. Opportunities and Challenges*, [http://www.spss.com/events/e\\_id\\_2652/Multi-Modal\\_World\\_Final\\_Slides\\_8.5.08.pdf](http://www.spss.com/events/e_id_2652/Multi-Modal_World_Final_Slides_8.5.08.pdf), 1.12.2009.

- [18] Wielki J., *Elektroniczny marketing poprzez internet. Reengineering procesu marketingowego*, Wydawnictwo Naukowe PWN, Warszawa-Wrocław 2000.
- [19] Wygant S., Olsen D., Call V., Curtin J., *Comparative Analyses of Parallel Paper, Phone and Web Surveys: Some Effects of Reminder, Incentive and Mode*, <http://assess.byu.edu/assessmentweb/unitreview/iaa/presentations>, 17.04.2009.

### **CREDIBILITY OF MEASUREMENT IN MIXED-MODE RESEARCH: THE COMPARISON OF EFFECTS OF PAPI AND CAWI APPLICATION**

**Summary:** The paper presents selected results of direct comparison for two pieces of independent research of chosen methodological effects of CAWI and PAPI utilization. Against a background of literature survey the authors present the results of their own comparative research in the field of measurement reliability with the application of Likert and semantic scales as well as the analysis of completeness of answers for open questions. The main conclusion of the analysis is comparable quality of measurement using both ways of collecting data – the scales are reliable enough and the answers for open questions are content rich. Technical possibilities in CAWI – especially the position rotation of scales and forcing responses – may contribute to obtaining data about better quality by means of Internet questionnaire.