

Radosław Mącik

Uniwersytet Marii Curie-Skłodowskiej w Lublinie
e-mail: radoslaw.macik@umcs.pl

**OCENA FORM I ZAKRESU INFORMACJI
O METODYCE BADAWCZEJ
W POLSKICH CZASOPISMACH NAUKOWYCH
DOTYCZĄCYCH MARKETINGU I KONSUMPCJI**

**FORMS AND SCOPE OF RESEARCH METHODOLOGY
INFORMATION IN POLISH SCIENTIFIC JOURNALS
FOR MARKETING AND CONSUMPTION**

DOI: 10.15611/ekt.2017.4.05

JEL Classification: C10, C18, C80, C83

Streszczenie: Dokonano oceny form i zakresu raportowania informacji metodologicznych 100 artykułów dotyczących marketingu i konsumpcji opublikowanych w 3 polskich czasopismach naukowych w 2016 r. Wykorzystano podejście eksperckie oraz drzewa klasyfikacyjne. Analizowane abstrakty zawierają zwykle skąpe informacje o metodzie i próbie. W 3/4 tekstów brakowało pytań lub hipotez badawczych. Informacje o próbie ograniczono typowo do liczebności i sposobu doboru, pomijając opis populacji, strukturę próby i poziom jej realizacji. Dobrze opisywano zbieranie danych, rzadko raportując rzetelność (w 19% tekstów) i trafność pomiaru (7%). Rzadkością była dyskusja wyników (12%). Co drugi artykuł spełniał minimalne wymagania przejrzystości badania, a za wzorcowe uznano 2 teksty. 60% badań nie było replikowalne. Wskazano implikacje dla autorów, czasopism i recenzentów.

Słowa kluczowe: raportowanie informacji metodologicznych, dobór próby, wiarygodność pomiaru, trafność, rzetelność, analiza danych, replikowalność analiz.

Summary: Paper evaluates forms and scope of reporting methodological information in 100 articles on marketing and consumption published in 3 Polish scientific journals in 2016 with the use expert approach and classification trees. Analysed abstracts usually contain scant information about the method and the sample. In 3/4 of the texts, research questions or hypotheses are lacking. The information about the sample is typically limited to the sample size and sampling method, omitting reference to the population, the sample structure and the response rates. Authors are describing data collection well, although they rarely report measurements reliability (in 19% of texts) and validity (7%). The discussion of results is a rarity (12%). Half of the texts meet the minimum requirements for the transparency of the study, and two them being considered the highest standard. 60% of the texts are not replicable. The implications for authors, journals and reviewers are indicated.

Keywords: methodological information reporting, sample selection, validity and reliability of measurement, data analysis, replicability of research.

1. Wstęp

Wyniki badań naukowych prezentowane w czasopismach i innych publikacjach naukowych powinny być przedstawione w sposób transparentny, tj. taki, by umożliwić ich odbiorcy krytyczną ocenę procesu badawczego, jego realizacji i analizy wyników. Pożądaną cechą raportowania wyników badań jest także ich przedstawienie, włączając niezbędne informacje o metodzie badania, stosowanych miarach oraz dobranej próbie, by można było konkretne badanie zreplikować i/lub wykorzystać jego wyniki w coraz powszechniej prowadzonych metaanalizach. Pozwala to na lepszą generalizację wyników badań i chroni przed rozpowszechnieniem wyników badań przeprowadzonych z naruszeniem rygorów metodologicznych.

Genezą tekstu, będącego rozwinięciem wersją wystąpienia w formie referatu (*invited lecture*) wygłoszonego podczas XXI Warsztatów metodologicznych im. Profesora Stefana Mynarskiego w Lublinie, których tematem przewodnim były: Dane – źródła i pochodzenie a wiedza, są obserwacje i doświadczenia autora występującego w różnych rolach w procesie wydawniczym. Zasadniczym celem artykułu jest zaprezentowanie wyników analizy eksperckiej celowo dobranej grupy publikacji pod kątem raportowania informacji metodologicznych, by określić ogólny poziom przejrzystości opisywanych procedur badawczych i zakres informacji o metodzie badania przekazywanych przez autorów. W dalszej kolejności celem jest wypracowanie praktycznych wskazówek co do pożądaných sposobów raportowania informacji metodologicznych w artykułach naukowych oraz wzbudzenie dyskusji naukowej na ten temat. Poza obszarem rozważań są faktyczne uzyskane w badaniach wyniki.

Obserwacja wąskiego zakresu zwyczajów publikacyjnych, tj. praktyk raportowania informacji metodologicznych, grupy naukowców zajmujących się badaniami marketingu, konsumpcji i zachowań konsumentów na podstawie systematycznego przeglądu artykułów opublikowanych w polskich czasopismach, okazała się wyzwaniem niezmiernie ciekawym, ale tylko częściowo budującym. Na podstawie celowo dobranej próby 100 publikacji z 3 czasopism, w których publikowane są teksty naukowe ze wskazanego obszaru tematycznego, dokonano próby oceny zakresu raportowania informacji o metodyce prezentowanych w tekstach badań, w tym o stosowanych metodach analizy danych, odnosząc się do dobrych praktyk publikacyjnych, również z innych dziedzin nauki.

Obecne i nadchodzące zmiany polityki naukowej, w tym oceny efektywności środków przeznaczanych na naukę, także w postaci analiz parametrycznych i nauko-metrycznych, doprowadziły do zmiany zwyczajów publikacyjnych znacznej grupy polskich naukowców. Niewątpliwie słuszne tendencje do zwiększenia umiędzynarodowienia dorobku nauki polskiej poprzez promowanie publikowania w renomowanych czasopismach zebranych na tzw. liście A MNiSW (niesłusznie nazywanej listą filadelfijską) nie dają niestety jednakowych szans publikacyjnych przedstawicielom różnych dyscyplin naukowych. Zazwyczaj wysokie wymagania w zakresie merytorycznym i językowym zbiegają się bowiem z dużą konkurencją autorów z całego

świata. Dlatego też w strategiach budowania dorobku publikacyjnego lepiej punktowane czasopisma polskie z tzw. listy B nadal odgrywają i odgrywać będą znaczną rolę, chyba że w ocenie parametrycznej publikacje krajowe przestaną być uwzględniane. W dziedzinie nauk ekonomicznych liczba punktowanych czasopism naukowych na liście B jest bowiem tylko o około połowę mniejsza od wskazanych na liście A czasopism posiadających współczynnik *impact factor*, o publikację w których konkurują autorzy z niemal całego świata.

Należy zauważyć, że różne dziedziny i dyscypliny naukowe wypracowały różne standardy publikacyjne, przy czym w kilku z nich poziom formalizacji wymagań jest dość wysoki, łącznie z koniecznością stosowania obszernych wytycznych zawartych w podręcznikach dla autorów, np. w dyscyplinie psychologii [APA 2010], czy też rozbudowanych wytycznych dla autorów w poszczególnych czasopismach, np. w badaniach medycznych. W dziedzinie nauk ekonomicznych przeważa model wytycznych w poszczególnych czasopismach. W przypadku poddanych analizie czasopism polskich zakres wymagań metodycznych określonych w ten sposób jest bardzo niewielki i ogranicza się w zasadzie do wymagań w zakresie ogólnie pojmowanej struktury artykułu. Rzadkie są również opracowania naukowe sugerujące konkretne standardy publikacji wyników badań określonego typu – np. M. Szreder postuluje przyjęcie proponowanego przez siebie standardu w prezentowaniu wyników badań ankietowych, zawierającego informacje niezbędne do oceny wiarygodności i precyzji badania, przy czym postulat ten dotyczy przede wszystkim prezentowania badań sondażowych w mediach, w tym sondaży wyborczych [Szreder 2010].

2. Wymagania metodyczne wobec publikowanych badań empirycznych zawarte w instrukcjach dla autorów wybranych czasopism

Wymagania metodyczne wobec publikowanych badań empirycznych zawarte w instrukcjach dla autorów publikujących w analizowanych czasopismach są tylko pozornie sformalizowane, ponieważ *explicite* wskazany zakres wymaganych informacji jest zwykle lakoniczny (zwykle tak, by pasował bez modyfikacji do różnych typów artykułów, w tym tekstów przeglądowych). Zidentyfikowane wymagania dla uwzględnionych w analizie trzech czasopism zostały zamieszczone w tab. 1.

W większości sytuacji wytyczne dla autorów są bardzo szczegółowe w warstwie technicznej, sugerując użycie określonych sposobów prezentacji treści, w tym wyników badań (formatowanie i wymagania techniczne wobec tabel i wykresów, a także dokumentowania źródeł literaturowych poprzez wskazanie stylu cytowań i jego dokładnych przykładów). Natomiast w warstwie metodologicznej wymagania te są często domyślne lub niespecyficzne. O ile konieczność wyraźnego wskazania celu artykułu jest zrozumiała, o tyle w żadnym z analizowanych czasopism od autorów nie wymagano zamieszczenia pytań i/lub hipotez badawczych. W dwu z trzech cza-

sopism pewien wąski zakres informacji metodologicznych jest wymagany w streszczeniu, ale w jednym z nich na tym się kończą wskazania dla autorów – w dalszym opisie nie ma żadnych konkretnych wymagań. Jedno z czasopism formalnie wymagało wyodrębnionej części metodycznej w strukturze tekstu, a kolejne szczegółowego raportowania informacji dotyczących populacji, sposobu doboru próby i jej struktury, a także wskaźnika odmów i sposobu traktowania braków danych – co również praktycznie sprowadza się do wyodrębnienia części tekstu te kwestie opisujących.

Tabela 1. Miejsce i zakres informacji metodologicznych wymaganych w instrukcjach dla autorów

Wyszczególnienie:	Czasopismo		
	A	B	C
Wymagane informacje metodologiczne w streszczeniu	Nie	Tak (rodzaj wykorzystywanej metody badawczej/ podejścia badawczego)	Tak (identyfikacja obiektu badań, istota stosowanej metody)
Wymagana wyodrębniona część metodyczna	Nie	Tak (wyróżnione za pomocą śródtytułów omówienie metodyki badań)	Nie
Szczegółowe wymagania co do raportowania informacji metodologicznych	Tak (określenie badanej populacji i operatu do wyboru jednostek do próby; wskazanie techniki doboru próby, oszacowanie w miarę możliwości błędu pomiaru, podanie liczebności próby i alokacji warstw (przy warstwowaniu); podanie wskaźnika odmów udziału w badaniu; wskazanie sposobu imputacji braków danych)	Nie	Nie

Źródło: opracowanie własne na podstawie dokumentów instrukcji dla autorów i plików powiązanych czasopism wybranych do analizy.

W żadnym z czasopism nie wymagano wskazania danych umożliwiających ocenę wiarygodności pomiaru – co prawda można to zrealizować tylko w części artykułów badawczych posługujących się złożonymi skalami pomiarowymi, ale takie podejście jest bardzo rozpowszechnione wśród badaczy. Podobnie też pozostawiono do decyzji autorów zamieszczenie innych informacji, w tym tych, które mogą warunkować replikację badań przez innych badaczy.

3. Metoda

Na potrzeby artykułu dokonano celowego doboru czasopism naukowych publikujących artykuły dotyczące marketingu, konsumpcji i zachowań konsumentów. Biorąc pod uwagę cel analizy, wybrano trzy z nich, kierując się ich zawartością merytoryczną oraz przyznaną punktacją w obrębie listy B MNiSW. Zdecydowano, aby nie wskazywać tytułów wybranych czasopism, posługując się kodami: A, B i C, ponieważ celem tekstu nie jest ocena czasopism jako takich, a praktyk autorów w nich publikujących. Wybrane czasopisma są punktowane w zakresie 9-12 punktów, co sytuuje je w środku rozkładu punktacji czasopism z listy B dla nauk ekonomicznych [Rozkosz, Kulczycki 2016]. Dwa z nich są miesięcznikami, jedno – publikuje rocznie kilkadziesiąt tomów z całej dziedziny nauk ekonomicznych, a kilka z nich dotyczy wybranej tematyki. Analizie poddano łącznie 100 tekstów, które ukazały się w 2016 r., losując numery z rocznika 2016 w taki sposób, by liczba analizowanych tekstów była proporcjonalna do ogólnej objętości rocznika (lub tematycznych numerów w roczniku). W ten sposób wybrano do analiz 41 tekstów z czasopisma A, 35 – z B i 24 z C. Jedynie w czasopiśmie A w wylosowanych numerach znalazły się teksty w języku angielskim i było ich 5 w analizowanych 100.

Wybrane artykuły poddano analizie eksperckiej, posługując się łącznie 39 zmiennymi nominalnymi i porządkowymi oraz analizą treści. Ze względu na charakter zmiennych zasadnicza analiza sprowadza się do klasyfikacji poszczególnych tekstów na podstawie jakościowo traktowanych kryteriów oraz zliczania częstości dla wyróżnionych kategorii. Tam, gdzie było to wskazane, poszerzono informacje z klasyfikacji o spostrzeżenia z analizy treści. Prezentowane wyniki odnoszą się do próby i nie pretendują do generalizacji do populacji wszystkich artykułów na wskazane tematy opublikowanych w polskich czasopismach w 2016 r.

Klasyfikując artykuły według typów, przyjęto podział na artykuły badawcze (z podtypami: opis, eksploracja i eksplanacja), przeglądy literatury (podtypy: artykuł przeglądowy i metaanaliza), artykuły koncepcyjne i studia przypadku. Struktura analizowanych tekstów (rys. 1) wskazuje na duży udział artykułów przeglądowych (ok. 1/3 próby) i dominację wśród artykułów badawczych tekstów skoncentrowanych na prostym opisie wyników badań (46 przypadków z 58). Artykuły o charakterze eksplanacji, metaanalizy oraz studium przypadku w próbie są raczej wyjątkami, podobnie artykuły koncepcyjne – prezentujące nowe koncepcje teoretyczne lub propozycje pomiaru zjawisk dotychczas w dany sposób niebadanych. Struktury tekstów według typów w poszczególnych czasopismach są dość podobne, chociaż czasopismo C publikuje nieco więcej tekstów badawczych (63% *versus* 54-56%). Wśród analizowanych tekstów dominują teksty jednoautorskie (ponad ¾ analizowanych, najmniej w czasopiśmie C – 69%), ponadto w czasopismach A i B teksty wielu autorów to zazwyczaj efekty współpracy międzyuczelnianej (różne afiliacje), a w czasopiśmie C – współpracy wewnętrznej (ta sama afiliacja obu autorów – nie było w tej grupie tekstów o większej ich liczbie).



Rys. 1. Struktura analizowanych tekstów według przyjętej typologii

Źródło: opracowanie własne.

Wśród artykułów badawczych (włączając do tej grupy studia przypadku oraz metaanalizę) – łącznie 61 tekstów – dominowało klasyczne podejście ilościowe (2/3 z tej grupy), pozostałe teksty oparte były na prostych analizach źródeł wtórnych lub stosowały podejście jakościowe. Wyjątkowo autorzy decydowali się na podejście mieszane, formalny eksperyment lub etnografię.

4. Wyniki

Analizując ilość i zakres informacji metodologicznych, w próbie publikacji przyjęto kolejność zgodną z układem typowego artykułu naukowego, tj. poczynając od zawartości abstraktu, przez wyodrębnione części opisujące metodę(y) badania i stosowaną próbę, aż po zasadniczą część z wynikami przeprowadzonych analiz. Podstawę procentowania stanowi w tej części artykułu grupa tekstów z pominięciem artykułów przeglądowych i koncepcyjnych, tj. 61 publikacji (chyba że wskazano inaczej).

4.1. Zawartość abstraktów

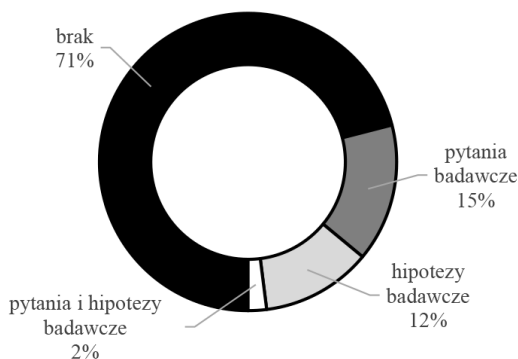
W abstrakcie przeszło połowa analizowanych tekstów z tej grupy (51%) zawierała bardzo skąpe informacje metodologiczne – zwykle była to metoda badania lub liczebność badanej próby, ewentualnie łącznie ze wskazaniem sposobu jej doboru. Szerszy zakres wspomnianych informacji zawierało 26% poddanych analizie tekstów, w większości przypadków były to 3-4 informacje, odpowiednio o: metodzie badania, liczebności i sposobie doboru próby, badanej grupie (opis) oraz stosowanych metodach analizy danych. W pozostałej części publikacji (23%) abstrakt nie

zawierał żadnych informacji metodologicznych. Natrafiono również na sytuacje nietypowe: jeden z tekstów zawierał informacje o liczebności próby i metodzie badania tylko w abstrakcie, a w kolejnym informacja w abstrakcie była w rażący sposób niespójna z zawartością tekstu głównego. Autorzy stosunkowo luźno traktowali wytyczne analizowanych czasopism w zakresie pożądanej zawartości abstraktu, a nawet jego struktury.

4.2. Informacje o metodzie

Prawie połowa (48%) tekstów miała w swojej strukturze wyodrębnioną część metodologiczną; często dlatego, że wymagana była ona przez czasopisma albo prezentowała korzystne dla autora informacje. W artykułach, gdzie takiej części brak, spekulatywnie można założyć, że do przyczyn tego braku należy brak wymagań w tym zakresie ze strony czasopisma lub niedbałość/brak świadomości autora(ów), ale też w przypadku tekstów słabszych przekonanie, że nie ma, czym się chwalić, lub szkoda autorowi poświęcić miejsce przy ograniczonej objętości tekstu.

Niezależnie od wyodrębnienia części metodologicznej sprawdzono, czy autorzy w swoich tekstach postawili pytania i hipotezy badawcze (rys. 2).



Rys. 2. Obecność pytań i hipotez badawczych w analizowanych tekstach

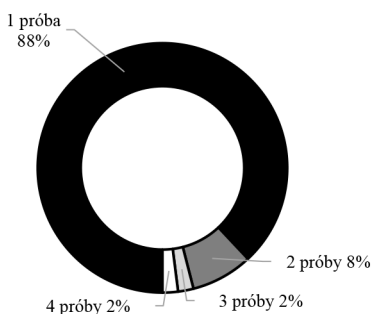
Źródło: opracowanie własne ($n = 61$).

W zdecydowanej większości analizowanych publikacji brak jest zarówno pytań, jak i hipotez badawczych, co sugeruje brak chęci autorów do ich weryfikacji i potwierdza skupienie na prostym opisie. Warto w tym miejscu zadać pytanie, czy wynika to z trudności formułowania pytań, a szczególnie hipotez, czy też z trudności weryfikacji lub obawy o zmniejszenie szans publikacyjnych w przypadku weryfikacji negatywnej, czy może z przekonania, że elementy te są niepotrzebne, szczególnie jeśli opisano cel pracy (co jest z reguły egzekwowane w czasopismach)?

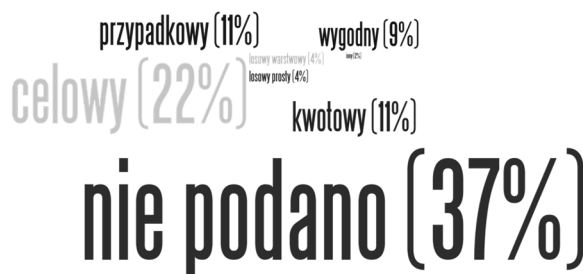
4.3. Informacje o próbie

Dobrze identyfikowalne są w analizowanych publikacjach informacje o stosowanej próbie, szczególnie na tle bardzo szczegółowych informacji o populacjach, z których próbę(y) pobierano. Informacje o stosowanych próbach zawarto na rys. 3.

a) liczba prób



b) sposób doboru próby



c) liczebności prób stosowane w badaniach ilościowych (zakres 22-2547 jednostek)



Rys. 3. Podstawowe informacje o stosowanych w analizowanych tekstach próbach

Źródło: opracowanie własne ($n = 61$).

Analizowane artykuły prezentowały typowo badania oparte na jednej próbie, jedynie ok. 12% z nich oparte było na 2-4 próbach. W badaniach ilościowych deklarowano liczebność prób pomiędzy 22 a 2547 jednostkami. W badaniach jakościowych odpowiednio: 3-50. W 37% przypadków nie udało się znaleźć w tekście artykułu, jaki sposób doboru próby zastosowano. Tam, gdzie zostało to określone, zwykle korzystano z nielosowych metod doboru próby – najczęściej posługiwano się doбором celowym, a relatywnie często kwotowym lub przypadkowym. Próby losowe wykorzystano jedynie w pojedynczych przypadkach, co sugeruje trudności z ich zastosowaniem, prawdopodobnie ze względu na dostęp do wykazów populacji, koszt dotarcia do respondenta, a także efektywność kontaktów z jednostkami wylosowanymi do próby. W efekcie stosowane próby formalnie rzadko mogą być uznane za reprezentatywne w sensie statystycznym, a wyniki z nich uzyskiwane odnoszone do populacji.

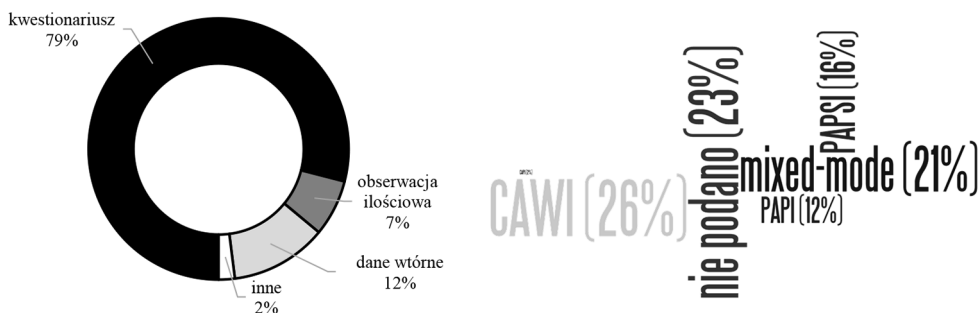
W zdecydowanej większości przypadków jednostką próby był konsument (72%), rzadziej badano przedsiębiorstwa (14%) lub gospodarstwa domowe (6%),

incydentalnie zaś jednostkami próby były: witryna internetowa i jednostka administracji. Struktura ta wiąże się z łatwością dotarcia do indywidualnych konsumentów (szczególnie w porównaniu z przedsiębiorstwami) i względną łatwością ich badania w porównaniu z całym gospodarstwami domowymi. W znacznej części tekstów (44%) nie podano żadnych informacji o strukturze próby, nawet w podziale na najbardziej oczywiste zmienne demograficzne, np. płeć. W podobnej części artykułów (39%) zamieszczono w tekście podstawowe informacje z tego zakresu, a w jedynie 17% przypadków artykuł zawierał tablice lub wykresy prezentujące strukturę próby.

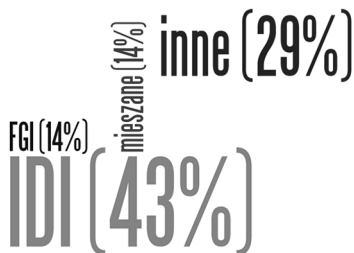
4.4. Informacje o metodzie zbierania danych

Wśród deklarowanych sposobów zbierania danych ilościowych dominowały badania kwestionariuszowe (z dominacją zbierania danych przez Internet – CAWI, oraz podejścia mixed-mode – łączącego kilka sposobów udostępnienia kwestionariuszy, czyli sposobów ograniczających koszt badania lub maksymalizujących efektywność kontaktów z respondentami (rys. 4)). Z kolei teksty jakościowe posługiwały się najczęściej pogłębionymi wywiadami indywidualnymi (43%). Dziwi nieco niska popularność zogniskowanych wywiadów grupowych – FGI (14%), powszechnie wykorzystywanych w badaniach stosowanych (rys. 4).

a) sposoby zbierania danych ilościowych ($n = 54$) b) sposoby kontaktu z respondentami ($n = 48$)



c) sposoby zbierania danych jakościowych ($n = 7$ – bardzo mała liczebność)



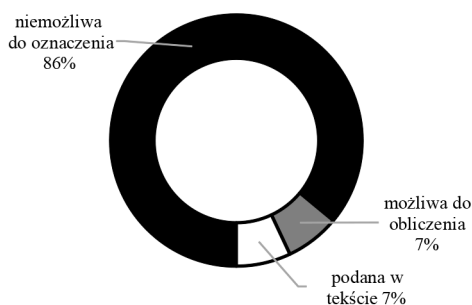
Rys. 4. Stosowane sposoby zbierania danych ilościowych i jakościowych

Źródło: opracowanie własne.

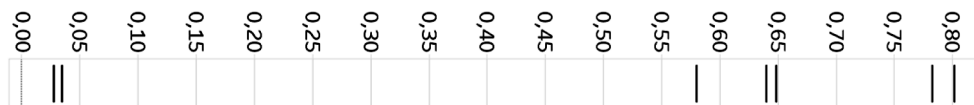
4.5. Informacje o poziomie realizacji próby

Poziom realizacji prób jest dużą niewiadomą (rys. 5) – można go ustalić bezpośrednio lub pośrednio (obliczyć na podstawie podanych informacji) jedynie w 14% artykułów, w których posługiwano się różnego typu kwestionariuszami (tj. w 7 przypadkach).

a) stopa poziomu realizacji próby ($n = 48$)



b) podane lub obliczone stopy zwrotu kwestionariuszy ($n = 7$)



Rys. 5. Informacje o stopie realizacji próby w analizowanych tekstach

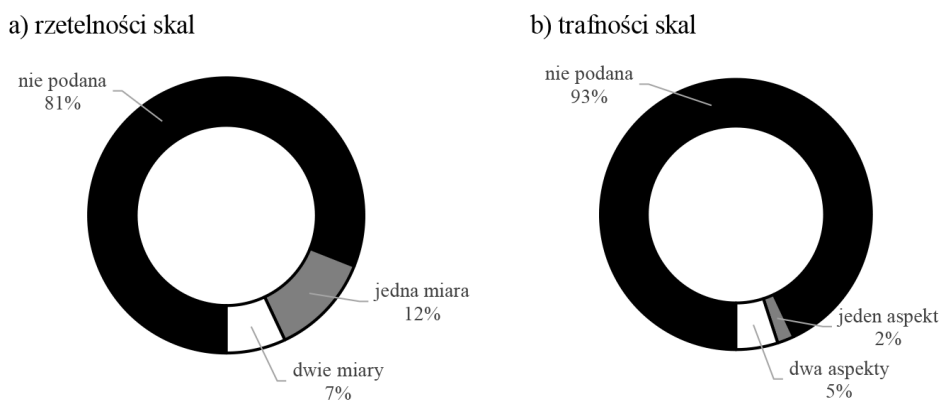
Źródło: opracowanie własne.

W przypadku stopy realizacji próby w dwu przypadkach uzyskano wartości wysokie – nieco poniżej 80%, w trzech – przeciętne (ok. 55-65%), a w dwu – skrajnie niskie – rzędu 3%. Brak informacji o stopie realizacji próby w większości tekstów może sugerować kilka sytuacji – po pierwsze: np. przy ogólnodostępnych kwestionariuszach, szczególnie internetowych, nie można oznaczyć liczby respondentów, którzy otrzymali kwestionariusz i z tego powodu nie można obliczyć poziomu realizacji próby – inaczej niż np. w badaniu z zaproszeniem pocztą elektroniczną, po drugie – dla prób nielosowych rejestruje się często tylko kontakty zrealizowane, np. w doborze kwotowym nie wszyscy badacze odróżniają odmowy udziału w badaniu od niezakwalifikowania się uczestnika ze względu na niespełnienie kryteriów rekrutacji, po trzecie – może to być efekt obawy o publikowalność tekstów z bardzo niską realizacją próby, gdzie faktyczną metodą doboru staje się dobór przypadkowy, a badania są *de facto* ochotnicy.

4.6. Informacje o skalach pomiarowych i wiarygodności pomiaru

Ważną kwestią jest zamieszczanie w artykułach badawczych używanych skal pomiarowych, ponieważ znajomość treści itemów skal pozwala na pełną replikację badania. W ok. 1/3 przypadków nie opisano skal w żaden sposób pozwalający na identyfikację używanych itemów (np. podając je w postaci nazw zmiennych, bez treści), w 28% tekstów posługiwano się wymiarami skal (np. otrzymanymi w eksploracyjnej analizie czynnikowej lub we wcześniejszych badaniach), a w co piątym tekście podano brzmienie skal tylko częściowo (np. tylko zmodyfikowane itemy). Pełną treść skal pomiarowych zawarto w tylko 20% analizowanych tekstów.

Osobnym problemem jest rzadkie raportowanie informacji o wiarygodności pomiaru. Rzetelność i szczególnie trafność skal to informacje rzadko ujawniane przez autorów analizowanych artykułów (rys. 6).



Rys. 6. Informacje o rzetelności i trafności stosowanych skal

Źródło: opracowanie własne ($n = 48$).

Żadnego wskaźnika rzetelności pomiaru nie podano w 4/5 tekstów! W pozostałych przypadkach posługiwano się zazwyczaj współczynnikiem alfa-Cronbacha, a alfa-Cronbacha i współczynnik rzetelności łącznej (CR) były podane tylko w tekstach, gdzie stosowano analizy PLS-PM (modelowanie ścieżkowe oparte na algorytmach częściowych najmniejszych kwadratów). Jeszcze gorzej sytuacja wygląda, jeśli chodzi o informacje o trafności pomiaru – brakowało jej aż w 93% analizowanych publikacji. Trafność zbieżną podano w 3 przypadkach, a różnicową – w dwu. Dwa aspekty trafności, podobnie jak dwie miary rzetelności, raportowali jedynie autorzy tekstów wykorzystujący podejście analityczne PLS-PM.

4.7. Informacje o stosowanych metodach analizy danych i opisie wyników

Jeśli chodzi o stosowane techniki analizy danych ilościowych, w analizowanych tekstach dominuje prosty opis statystyczny (w 60%) przypadków, analizę wariancji zastosowano w 8% przypadków, modele ścieżkowe (głównie w wersji PLS-PM) w 6% tekstów, a inne techniki analityczne były stosowane rzadko (rys. 7). Więcej niż jedną metodę zastosowano w 6% tekstów. Warto zauważyć, że stosowane techniki na ogół są proste i dostarczają wyników mało pogłębionych, zastosowanie podejść modelowych jest rzadkie, a żaden z autorów nie przedstawił w swoim tekście wartości współczynników.



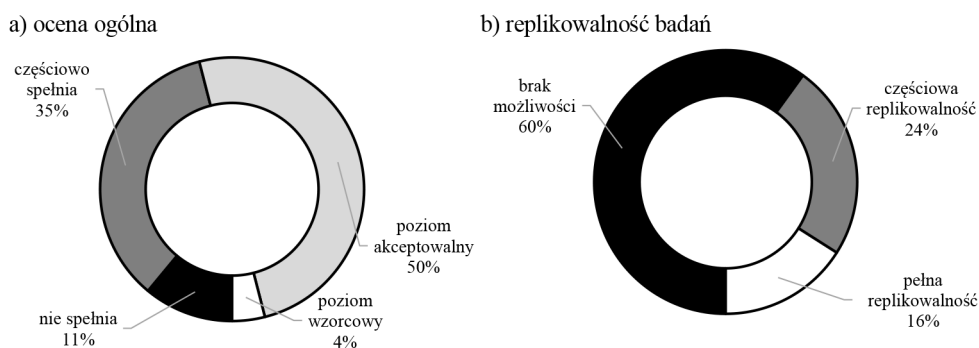
Rys. 7. Stosowane techniki analizy danych – teksty ilościowe

Źródło: opracowanie własne ($n = 48$).

Większość analizowanych tekstów (73%) charakteryzowała się szczegółowym opisem uzyskanych wyników, na poziomie zgodnym ze stosowanymi metodami analizy danych, a w kolejnej $\frac{1}{4}$ artykułów wyniki opisano lakonicznie, skupiając się na przekazie w formie tablic i wykresów. W próbie znalazł się również jeden tekst, gdzie przedstawiono w syntetyczny sposób wyniki badania, ale w tekście ich wcale nie opisano – treść nie nawiązywała do nich bezpośrednio. Wysoki poziom szczegółowości opisu wyników nie koresponduje jednak z pożądaną dyskusją otrzymanych wyników z innymi badaniami na ten sam lub podobny temat. W 88% analizowanych tekstów nie było wyodrębnionej w strukturze dyskusji wyników, a w treści pojawiały się co najwyżej pojedyncze odniesienia do prac badawczych innych autorów. W 5% przypadków dyskusja była skrótowa, a tylko w 7% miała charakter rozwinięty i dość obszerny. Nie jest to sytuacja pożądana – czytelnik nie może wtedy odnieść wyników z czytanego tekstu do prac innych badaczy i poddać je własnej ocenie. Co prawda w wielu tekstach przedstawiano skrótowo wyniki innych badań w sekcji przeglądu literatury (nawet jeśli nie była tak nazwana), natomiast nie zastępuje to formalnej dyskusji wyników.

5. Podsumowanie

Poziom raportowania informacji metodologicznych w analizowanych tekstach jest zróżnicowany i przeciętnie dość niski. Biorąc pod uwagę łączne spełnienie kryteriów oceny eksperckiej (rys. 8), dokładnie połowę tekstów (50%) uznano za reprezentujące poziom akceptowalny, uwzględniając zastosowane metody badawcze i analityczne (generalnie raczej proste i znane od wielu lat), a tylko 4% z nich za publikacje wzorcowe w tym zakresie (wszystkie stosujące modele ścieżkowe PLS-PM). Jednocześnie nieco ponad 1/3 tekstów tylko częściowo spełnia kryteria oceny, a 11% publikacji nie spełnia ich w żadnym stopniu, ponieważ informacje o metodyce przeprowadzonego badania są obecne w tekście w stopniu szczątkowym, niepozwalającym na odniesienie się do uzyskanych wyników.



Rys. 8. Oceny podsumowujące – teksty ilościowe

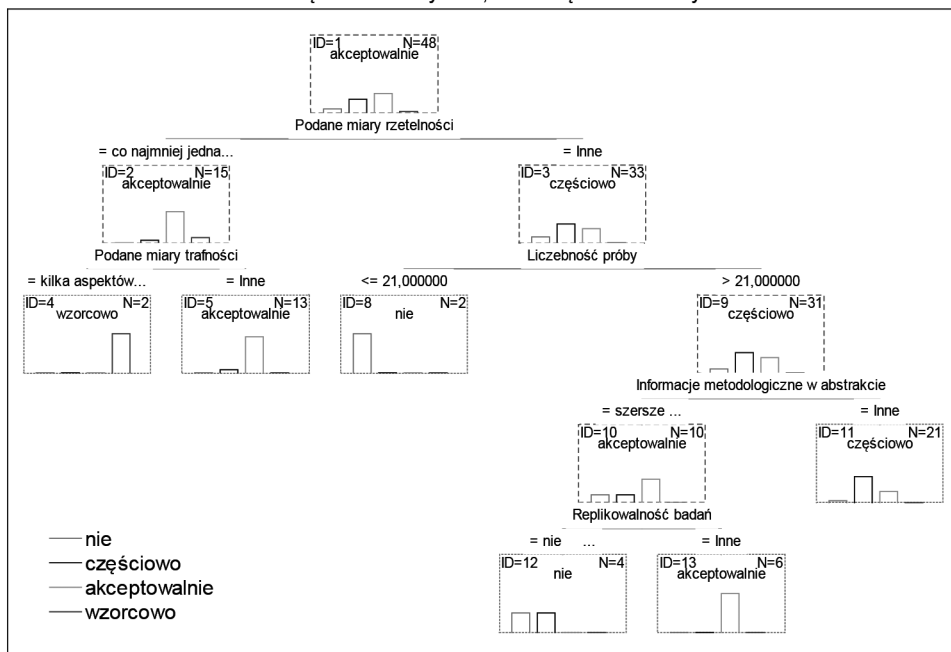
Źródło: opracowanie własne ($n = 48$).

Analizowane artykuły są w niewielkim stopniu replikowalne. Pełną replikowalnością, ze względu na zaprezentowane informacje o metodzie badania i dołączone skale, charakteryzowało się jedynie 16% prezentowanych badań. Około 1/4 tekstów charakteryzuje częściowa replikowalność – albo brakuje w nich pewnych informacji o badanej populacji i próbie, albo nie wszystkie itemy skal (czy też całe skale) są bezpośrednio dostępne w tekście, chociaż można do nich dotrzeć, np. odwołując się do innych publikacji autora lub do oryginalnych tekstów wprowadzających do dyskursu naukowego narzędzia przejęte w całości przez autorów analizowanych opracowań. Niestety aż 60% prezentowanych badań nie da się wcale replikować – ich autorzy nie zamieścili w ich treści potrzebnych w tym celu informacji. W efekcie nie tylko trudno replikować już opublikowane badania, które przez to mają bardziej charakter przyczynków niż badań rozwijających określone nurty, ale też trudno uwzględnić takie badania w metaanalizach.

Ponieważ użyto wielu kryteriów oceny, a liczebność zbioru poddanych ocenie tekstów jest niewielka, wskazanie, które z nich w największym stopniu wpływają na

ocenę końcową (dla tekstów, w których wykorzystano metody ilościowe), było możliwe za pomocą technik drzew klasyfikacyjnych (C&RT). Wyniki klasyfikacji (na podstawie 11 zmiennych, w tym 2 ilościowych, por. rys. 11) dla tekstów, w których wykorzystano podejście ilościowe, przedstawiono na rys. 9.

Drzewo 3 dla jakościowej zmiennej objaśnianej: Spełnienie kryteriów oceny
Liczba węzłów dzielonych: 5, liczba węzłów końcowych: 6



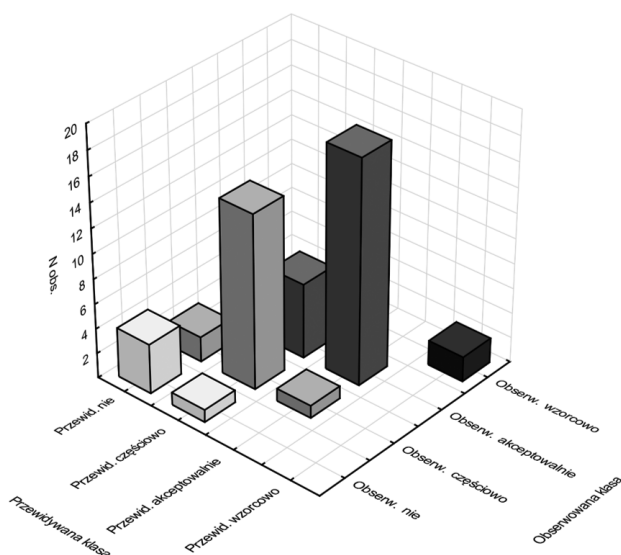
Rys. 9. Przyjęte drzewo klasyfikacyjne – teksty ilościowe

Źródło: opracowanie własne ($n = 48$).

Klasyfikacja tekstu jako *wzorcowego* w sensie raportowania w tekście informacji metodologicznych następowała, jeśli podano w nim co najmniej jedną miarę rzetelności pomiaru i kilka aspektów trafności (co najmniej trafność zbieżną i różnicową). Teksty sklasyfikowane jako *akceptowalne* charakteryzowały się podaniem przynajmniej jednej miary rzetelności pomiaru lub liczebnością próby większą niż 21 jednostek, szerszym zakresem informacji metodologicznych w abstrakcie i przynajmniej częściową możliwością replikacji badania. Teksty *częściowo akceptowalne* oprócz liczebności próby powyżej 21 miały skąpe informacje metodologiczne w abstrakcie lub odnotowano ich brak. Do klasyfikacji tekstu jako nieakceptowalnego w ocenianym zakresie prowadziło zbadanie próby mniejszej od 21 jednostek lub jej równej albo przy próbie większej, szerszym zakresem informacji metodologicznej w abstrakcie – brak replikowalności prezentowanych badań.

Prezentowane drzewo klasyfikacyjne zostało wybrane na podstawie V -krotnego sprawdzianu krzyżowego jako optymalne, tj. drzewo najmniejszej wielkości (najmniej złożone), którego koszty sprawdzianu krzyżowego nie różnią się znacznie od minimalnych kosztów sprawdzianu krzyżowego. Dla poszczególnych klas zmiennej objaśnianej *Spełnienie kryterium oceny* poprawność klasyfikacji wyniosła: klasa *nie* – 80%, klasa *częściowo* – 82,3%, klasa *akceptowalnie* – 75%, klasa *wzorcowo* – 100%. Macierz klasyfikacji przedstawiono na rys. 10. Najczęstszym błędem klasyfikacji było zaliczenie tekstu do przewidywanej kategorii *częściowo* przy obserwowanej kategorii *akceptowalnie*.

Macierz klasyfikacji 3
Zmienna zależna: Spełnienie kryteriów oceny
Opcje: Jakościowa zależna, Drzewo numer 3, Próba do analizy



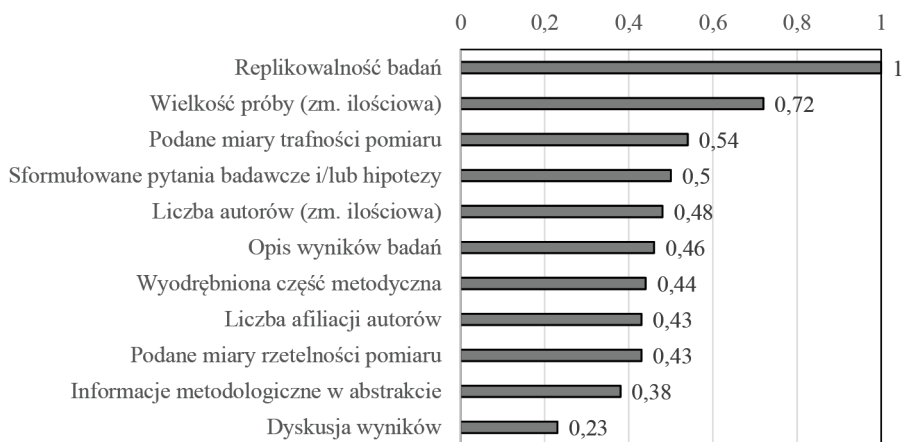
Rys. 10. Macierz klasyfikacji – teksty ilościowe

Źródło: opracowanie własne ($n = 48$).

Wśród predyktorów klasyfikacji względnie największą wagowość miały: replikowalność badań (3 poziomy), wielkość próby (zm. ilościowa), podane miary trafności pomiaru. Z kolei najmniejsza wagowość w klasyfikacji dotyczyła zakresu informacji metodologicznych w abstrakcie oraz pogłębienia dyskusji wyników badań własnych z innymi publikacjami.

Dla tekstów prezentujących wyniki badań jakościowych (w próbie zidentyfikowano 10 takich przypadków) oceny dokonywano według mniejszej liczby kryteriów. Otrzymano podobny do tekstów ilościowych odsetek ocen *akceptowalnych* (tj. połowę), mniejszy odsetek tekstów zaklasyfikowano jako *częściowo akceptowalne*,

a więcej niż w poprzednim przypadku – 30% – uznano za *niespełniające kryteriów* raportowania informacji metodologicznych. Główne kryteria sprzyjające wyższej ocenie dla tej grupy tekstów to pogłębiony opis zastosowanej metody badawczej (zwykle jednak bez odrębnej części metodycznej) oraz obszernie przedstawienie wyników przeprowadzonych badań.



Rys. 11. Względna ważność predyktorów w klasyfikacji – teksty ilościowe

Źródło: opracowanie własne ($n = 48$).

Niezbyt optymistyczne wyniki przedstawionych analiz uprawniają do postawienia pytań o przyczyny opisanego stanu. Czy jest to efekt relatywnie niskich wymagań czasopism w połączeniu z dość ściśle przestrzeganyymi wymaganiami dotyczącymi objętości publikowanych tekstów? Czy sytuacja wynika z chęci publikowania dużo, by maksymalizować liczbę punktów do oceny parametrycznej jednostki, a krótkie teksty raczej nie pozwalają dokładnie opisać wszystkich aspektów metodyki badań? Czy też wreszcie wynika to ze zróżnicowanej świadomości metodologicznej autorów? W świetle dążenia do doskonałości naukowej obecnej w polityce naukowej na te i podobne pytania warto spróbować udzielić odpowiedzi. Raczej pesymistyczna konkluzja, wskazująca wiele słabości i braków w opisie metodyki badań w publikowanych tekstach, nie ulega złagodzeniu, jeśli porównamy uzyskane rezultaty z podobnymi analizami, np. analizą publikacji w renomowanych czasopismach medycznych z 2011 r. [Bennett i in. 2011]. W przywołanym tekście analiza 117 kwestionariuszowych badań medycznych wskazała na zamieszczenie kwestionariusza lub zasadniczych jego pytań/skał w 35% artykułów, podanie danych o rzetelności lub trafności pomiaru jedynie w 19% przypadków, a stopę zwrotu kwestionariuszy w 25% tekstów. Po 11% artykułów zawierało odniesienie do reprezentatywności próby oraz opis postępowania z brakami danych [Bennett i in. 2011, s. 1].

Przeprowadzona analiza ma ograniczenia wynikające z głównie opisowego charakteru kryteriów oceny, a także doboru tekstów i ich źródeł. Autor zamierza poszerzyć zakres analiz w kolejnej publikacji, w której uwzględniony zostanie szerszy zakres czasowy i tematyczny publikacji. Niemniej jednak zauważone zwyczaje publikacyjne są na tyle wyraźne, że upoważniają do sformułowania wielu zaleceń o praktycznym charakterze, zarówno dla autorów, jak i dla redakcji czasopism, których stosowanie przyczynić się może do wyższego poziomu przejrzystości raportowania wyników badań empirycznych.

6. Implikacje praktyczne

Przeprowadzona analiza ma szerokie implikacje praktyczne, w tym dla redakcji czasopism, recenzentów i autorów tekstów naukowych.

Redakcje czasopism powinny rozważyć doprecyzowanie zaleceń dla autorów w zakresie wymagań raportowania informacji o metodyce badania, wskazując co najmniej listę wymagań co do opisu populacji i próby w formie wymaganej w czasopiśmie A, ale także zalecając dla tekstów badawczych strukturę z wyodrębnionymi częściami „Metoda”/„Próba”/„Miary”, a dla artykułów wykorzystujących skale złożone przynajmniej podstawowe miary rzetelności i trafności pomiaru dla używanych konstruktów. Dodatkowo istotne jest dopuszczenie publikacji artykułów dłuższych niż 20-22 tys. znaków ze spacjami, by nie powodować dysonansu u autorów – co da się zmieścić w tekście o ograniczonej objętości: więcej informacji o metodzie czy więcej wyników badania. W stosunku do recenzentów można sformułować prośbę, by bardziej rygorystycznie sprawdzali spełnienie zaleceń dla autorów w zakresie informowania o metodyce przeprowadzonych badań i wiarygodności pomiaru.

Tabela 2. Zalecenia dla autorów w zakresie raportowania informacji metodologicznych

Obszar informacji	Pożądany zakres informacji
Streszczenie	Podstawowe informacje o badanej populacji i próbie oraz metodzie badania
Metoda	Informacje o metodzie pomiaru wraz z uzasadnieniem wyboru metody
Próba	Określenie badanej populacji i sposobu doboru próby wraz oszacowaniem w miarę możliwości błędu pomiaru, podanie liczebności próby i alokacji warstw (przy warstwowaniu); podanie wskaźnika odmów udziału w badaniu
Miary	Opis konstruktów używanych w badaniu wraz ze wskazaniem skal dla ich pomiaru, jeśli są to skale nowe, z dołączeniem brzmienia ich itemów oraz ich wymiarowości na podstawie analiz czynnikowych (co najmniej EFA, zalecane CFA), podanie przynajmniej jednej miary rzetelności pomiaru, podanie miar trafności zbieżnej i różnicowej, wskazanie sposobu imputacji braków danych
Wyniki	Miary efektu wpływu zmiennych niezależnych na zmienną zależną (jeśli to możliwe) lub możliwość ich wyliczenia; inne informacje umożliwiające replikację wyników badania

Źródło: opracowanie własne.

W dalszej kolejności na podstawie przeprowadzonych analiz sformułowano kilka zaleceń dla autorów tekstów, których uwzględnienie w swojej praktyce publikacyjnej doprowadzi do większej przejrzystości raportowania procesu badawczego w artykułach naukowych – zalecenia zebrano w tab. 2.

Informacje metodologiczne w artykule badawczym powinny pozwolić na łatwą replikację badania, a jeśli to możliwe, powinny zawierać miary efektu wpływu zmiennych niezależnych na zmienną zależną lub umożliwiać ich wyliczenie w celu ich wykorzystania w metaanalizach.

Odrębnym zadaniem wydaje się przygotowanie zwięzłego podręcznika dobrych praktyk raportowania informacji metodologicznych w różnego typu artykułach badawczych, nie tylko z uwzględnieniem powszechnie stosowanych kwestionariuszowych badań ilościowych, ale też podejść jakościowych czy też studiów przypadku. Podręcznik ten powinien być szeroko upowszechniony szczególnie wśród najmłodszych adeptów nauki: początkujących doktorantów czy też studentów ostatnich lat studiów angażujących się w działalność naukową, np. przez uczestnictwo w studenckich kołach naukowych.

Literatura

- APA, 2010, *Publication Manual of the American Psychological Association*, Sixth Edition. Available at: <http://www.apastyle.org/manual/index.aspx> [Accessed June 26, 2017].
- Bennett C. i in., 2011, *Reporting guidelines for survey research: an analysis of published guidance and reporting practices*, PLOS Medicine, 8(8), p. e1001069.
- Rozkosz E.A., Kulczycki E., 2016, *Polskie czasopisma z nauk humanistycznych i nauk społecznych w liczbach*, Poznań.
- Szreder M., 2010, *O potrzebie określenia standardu w prezentacji badań sondażowych*, Marketing i Rynek, nr 9, s. 9-12.