

Próby w badaniach psychologicznych – wnioski i rekomendacje płynące z przeglądu literatury

Joanna Świdowska¹

Akademia Pedagogiki Specjalnej im. M. Grzegorzewskiej, Instytut Psychologii

<https://orcid.org/0000-0001-7439-2973>

Magdalena Puchalska

Akademia Pedagogiki Specjalnej im. M. Grzegorzewskiej, Instytut Psychologii

<https://orcid.org/0000-0001-6116-4656>

Maciej Grzegorzczyk

Uniwersytet Kardynała Stefana Wyszyńskiego w Warszawie, Instytutu Psychologii

<https://orcid.org/0009-0004-6325-7767>

Zuzanna Szalek

Uniwersytet Kardynała Stefana Wyszyńskiego w Warszawie, Instytutu Psychologii

<https://orcid.org/0009-0008-0699-1374>

Natalia Nadolna

Uniwersytet Kardynała Stefana Wyszyńskiego w Warszawie, Instytutu Psychologii

<https://orcid.org/0009-0009-1047-4377>

Martyna Andryszkiewicz

Uniwersytet Łódzki, Instytutu Psychologii

<https://orcid.org/0009-0008-7433-6013>

Maria Jastrzębska

Uniwersytet Kardynała Stefana Wyszyńskiego w Warszawie, Instytutu Psychologii

<https://orcid.org/0009-0005-6479-7287>

Przemysław Szablowski

<https://orcid.org/0009-0008-7728-9339>

Maja Śmigrodzka

Akademia Pedagogiki Specjalnej im. M. Grzegorzewskiej, Instytut Psychologii

<https://orcid.org/0009-0005-9517-5253>

¹ Adres do korespondencji: jswiderska@aps.edu.pl.

Streszczenie

Cel: Celem przeglądu była identyfikacja elementów występujących w opisach procedury badania i próby oraz sformułowanie na tej bazie rekomendacji dla autorów.

Metoda: Badanie przeprowadzono zgodnie ze standardem PRISMA. Łącznie przeanalizowano 305 artykułów polskich badaczy z czterech czasopism, z czego 230 zostało poddanych krytycznej ocenie. Ostateczne analizy dotyczyły 295 opisów badań.

Wyniki: Analizy wykazały, że najczęściej badacze raportują: procedurę badania, zastosowaną metodę doboru i wielkość próby, formę badania, a z cech socjodemograficznych: płeć i wiek.

Konkluzje: Wnioski z analiz i doświadczenia z przeglądu pozwoliły sformułować rekomendacje w zakresie wyczerpującego metodologicznie opisu procedury i próby.

Słowa kluczowe: dobór próby, opis procedury, jakość danych, generalizacja, przegląd literatury

Psychologia jest nauką empiryczną, która zajmuje się zachowaniem człowieka i jego procesami psychicznymi (Zimbardo i Gerrig, 2012). Oznacza to, że psychologowie badają określoną liczbę osób i na drodze wnioskowania indukcyjnego starają się opisać, wyjaśnić i przewidywać działania jednostki oraz mechanizmy, które za nimi stoją. Aby zmaksymalizować zakres generalizacji wniosków, badania powinny być prowadzone na próbach reprezentatywnych.

Odpowiedni dobór próby jest zatem kluczowy do przeprowadzenia trafnego badania, a niedostateczna dbałość w tej kwestii może nieść za sobą negatywne konsekwencje. Do najsłynniejszych metodologicznych pomyłek należy m.in. porażka magazynu *Literary Digest* w sondażach przedwyborczych z 1936 roku, który przewidywał sukces Landona (Babbie, 2021). Według Squire'a (1988) do błędnej predykcji przyczyniły się: 1) niski wskaźnik zwrotu ankiet (2 mln z 10) i 2) źle dobrana próba, która – jak zauważa Babbie (2021) – składała się z przedstawicieli zamożnej części społeczeństwa.

Niemniej opisywana porażka przyniosła korzyści Gallupowi, który stosując kwotowy dobór próby, prawidłowo przewidział ówczesny wynik wyborów. Niestety, metoda ta nie sprawdziła się 9 lat później, gdy wraz z Crossleyem i Roperem prognozowali wygraną Deweya (Lusinchi, 2018). Planując dobór kwotowy, nie uwzględnili przemiany struktury społecznej spowodowanej migracjami ze wsi do miast. Niedbałość w doborze próby może prowadzić do fałszywych wyników i negatywnych konsekwencji dla samych badaczy, o czym świadczy historia *Literary Digest*, które ostatecznie upadło w dwa lata po nieszczęśliwych wyborach (Babbie, 2021). Z kolei wydarzenia z 1946 roku zostały skomentowane przez *New Yorker* jako „całkowity upadek badań opinii publicznej” (Lusinchi, 2018, s. 3).

Wśród innych znanych przykładów znajdują się: raport Kinseya dotyczący seksualności Amerykanów (Maslow i Sakoda, 1952) czy badania Leary'ego i Alperta nad skutecznością terapii z wykorzystaniem LSD w zmniejszaniu ryzyka recydywy (Brzeziński, 2019).

Biorąc pod uwagę powyższe doniesienia, podjęliśmy się przeglądu artykułów naukowych opisujących badania ilościowe autorstwa psychologów afiliowanych przy polskich jednostkach badawczych. Miał on na celu sprawdzenie szczegółowości opisu procedury doboru i samej próby, gdyż są to elementy kluczowe do przeprowadzenia replikacji.

Metody doboru próby

W literaturze wyróżnia się probabilistyczne i nieprobabilistyczne metody doboru. Te pierwsze oparte są na zasadzie, że każda osoba z populacji ma taką samą szansę na wzięcie udziału w badaniu i została do niego wylosowana, co z kolei pozwala na uzyskanie reprezentatywnej próby. Nieprobabilistyczne metody doboru nie gwarantują natomiast, że otrzymana próba będzie odwzorowywać docelową populację pod względem interesujących badacza cech, tym samym należy podchodzić bardzo ostrożnie do uogólniania wniosków z takich badań (Shaughnessy i in., 2020).

W literaturze (Brzeziński, 2019; Sarstedt i in., 2018) wyróżnia się następujące podstawowe metody losowego doboru: prosty, warstwowy (ang. *stratified sampling*) i grupowy (ang. *cluster sampling*). Do ich przeprowadzenia niezbędne jest posiadanie operatu losowania (listy wszystkich elementów z populacji) i generatora liczb losowych.

Istnieje wiele nieprobabilistycznych metod doboru próby. Jednym z takich podejść jest zastosowanie technik, w których badacze sami wybierają respondentów albo na podstawie danych kryteriów (dobór celowy i kwotowy), albo robią to w sposób przypadkowy (ang. *accidental sampling*). Ta ostatnia metoda nie jest polecana, ponieważ nie podlega kontroli żadna zmienna populacji, co uniemożliwia generalizację wniosków (Brzeziński, 2019).

Celowy dobór polega na wybraniu osób, które według badaczy lub oceny eksperckiej są typowymi przedstawicielami docelowej populacji. Oznacza to, że respondenci spełniają wcześniej określone kryteria przynależności, np. wykonywany zawód, miejsce zamieszkania. Udoskonaleniem tej metody jest dobór kwotowy (ang. *quota sampling*), w którym na wstępie określa się rozkłady interesujących badaczy cech populacji. Celem doboru staje się odtworzenie takich samych proporcji tych zmiennych w próbie (Sarstedt i in., 2018).

W badaniach społecznych często wykorzystuje się wygodny dobór próby (ang. *convenience sampling*), w którym selekcja uczestników jest podyktowana możliwościami samych badaczy, a nie przesłankami teoretycznymi. Wybierane są osoby, które łatwo zlokalizować lub zrekrutować, co w konsekwencji prowadzi do niedoreprezentowania pewnych warstw populacji. Próby tak zebrane są tendencyjne i narażone na wiele błędów, a ponadto często składają się z ochotników (Baker i in., 2013). Ochotnicy to osoby, które same zgłaszają się do udziału w badaniach, odpowiadając na zamieszczane ogłoszenia. Tworzą oni tendencyjną próbę, różniącą się od ogólnej populacji, m.in. wyższym poziomem wykształcenia, inteligencji i potrzeby aprobaty społecznej (Brzeziński, 2019). W porównaniu z innymi metodami doboru próby składające się z ochotników są bardziej narażone na tendencyjne braki w odpowiedziach (ang. *non-response bias*) i błąd samoselekcji (Cheung i in., 2017).

W przypadku populacji ukrytych lub trudno dostępnych, a także w sytuacji, gdy członkowie populacji mogą nie chcieć dzielić się wrażliwymi informacjami z obcymi, np. chorzy na AIDS, przydatna jest metoda kuli śnieżnej. Polega ona na wybraniu przez badaczy niewielkiej liczby reprezentantów populacji, a następnie poproszeniu ich o rekrutację kolejnych osób. Proces ten wykorzystuje sieci społeczne respondentów. Warto podkreślić, że reprezentatywność pierwszych przebadanych osób jest kluczowa dla jakości próby (Salganik i Heckathorn, 2004).

Osobną kwestię stanowi dobór próby w badaniach online, np. poprzez umieszczanie linków na portalach. Respondentami są osoby, mające dostęp do Internetu, które w danym momencie odwiedziły odpowiednią stronę internetową i podjęły decyzję o udziale w badaniu. Badacz nie ma tutaj pełnej kontroli nad selekcją konkretnych osób. Takie próby mogą być tendencyjne i podatne na błędy wynikające z wykorzystania populacji osób korzystających z Internetu oraz z błędu samoselekcji (Bethlehem, 2010).

Wielkość próby a moc testu statystycznego

Dobór próby adekwatnej wielkości jest kluczowy dla uzasadnienia twierdzeń dotyczących populacji, z której to została pobrana, i dla trafności zewnętrznej badania. Możliwe sposoby oszacowania jej liczebności zależą od metody doboru i estymowanego parametru populacyjnego (Brzeziński, 2019). W literaturze podkreśla się, że próba nie powinna być zbyt mała, gdyż obniża to rzetelność badania, ale również nie powinna być zbyt duża, gdyż zwiększa ryzyko popełnienia błędu I rodzaju. Zatem oba oszacowania są obciążone i prowadzą do zaniżenia lub zawyżenia parametrów populacyjnych (Sapra, 2022).

Wielkość próby jest także kluczowa dla osiągnięcia odpowiedniej mocy testu statystycznego, czyli prawdopodobieństwa wykazania faktycznie występujących różnic lub związków (Singh i Masuku, 2014). Wyróżnia się dwie główne metody szacowania: 1) *a priori* – wielkość próby szacowana jest przed badaniem na podstawie: mocy testu, poziomu istotności i wielkości efektu; oraz 2) *post hoc* – istotą jest określenie, czy wielkość użytej próby minimalizuje ryzyko popełnienia błędu II rodzaju, przy określonym poziomie istotności i przyjętej wielkości efektu. Nie są to jedyne metody, a wymienienie wszystkich wykracza poza problematykę podjętą w niniejszym tekście (zob. Faul i in., 2007).

Znaczna część z determinant wielkości próby (np. przyjęty poziom ufności, hipotezy determinujące, stosowane testy statystyczne, rodzaj szacowanego parametru) jest silnie zależna od decyzji badacza. Dlatego wielkość próby powinna być decyzją przemyślaną i uzasadnioną, a nie przypadkową.

Kwestie etyczne, prawne i wynagradzanie osób badanych

Począwszy od doboru próby aż do zakończenia badania psycholog powinien przede wszystkim kierować się wolnością i godnością każdego uczestnika, a nie pokusą publikacji przełomowego wyniku (Brzeziński, 2019). Udział w badaniu

powinien być dobrowolny i anonimowy, poprzedzony wyrażeniem świadomej zgody. W ostatnich latach etyczne aspekty prowadzonych badań zyskały na znaczeniu, co w efekcie doprowadziło do wymogu pozyskania zgody odpowiedniej komisji ds. etyki badań przed ich rozpoczęciem.

Zmianie uległy także uwarunkowania prawne związane ze zbieraniem i przechowywaniem danych, szczególnie dotyczących informacji wrażliwych. Ważną do poruszenia kwestią jest również gratyfikacja respondentów. Pomijając kontrowersje związane z jej stosowaniem (zob. Brase, 2009), warto zdawać sobie sprawę z prawnych i finansowych aspektów takiej decyzji. Okazuje się, że lepszym rozwiązaniem jest zastosowanie nagród (nawet mniejszych) dla wszystkich niż losowe rozdzielanie większych nagród tylko dla części respondentów. W tym drugim przypadku Urząd Skarbowy ma prawo ocenić tę metodę jako loterię promocyjną, według definicji zawartej w art. 2 ustawy o grach hazardowych (2009)². Loterie promocyjne zgodnie z art. 7 wspomnianej regulacji „mogą być urządzone, na podstawie udzielonego zezwolenia”, którego według art. 32 „udziela dyrektor izby administracji skarbowej”.

Poruszone problemy związane z doбором prób do badań psychologicznych pozwoliły na sformułowanie następujących pytań badawczych:

1. Jakie informacje na temat próby podają polscy badacze, a czego w nich brakuje?
2. Jakie informacje o procedurze badania i próbie powinny być podane, aby opisy można było uznać za dokładne i wyczerpujące?
3. Jakie powinny być standardy w zakresie opisu procedury badania i próby?

Metoda

W celu zapewnienia odpowiednich wymogów naukowych przeprowadzonego przeglądu literatury zastosowano standard PRISMA (Moher i in., 2009). Nakłada on wymóg podania następujących informacji: liczbę zidentyfikowanych i odrzuconych publikacji, przyjętych kryteriów włączania i wykluczania oraz przebiegu procesu wyboru artykułów.

Identyfikacja artykułów

Ze względu na problem badawczy ukierunkowany na psychologiczne badania empiryczne polskich autorów przeprowadzenie przeglądu poprzez inspekcję stanu baz danych uznaliśmy za nieoptymalne rozwiązanie. Dlatego postanowiliśmy w pierwszym kroku wyłonić czasopisma, które je publikują. Kluczowe było,

² „Loterie promocyjne, w których uczestniczy się przez nabycie towaru, usługi lub innego dowodu udziału w grze i tym samym nieodpłatnie uczestniczy się w loterii, a podmiot zarządzający loterię oferuje wygrane pieniężne lub rzeczowe” (Ustawa o grach hazardowych, 2009, art. 2.1).

aby tematyka ukazujących się w nich badań była różnorodna, dotarcie do tekstów krajowych i zagranicznych oraz uwzględnienie czasopism o różnym prestiżu w celu ograniczenia efektu tendencyjnego wyboru źródeł. Na drodze dyskusji wewnątrz zespołu i konsultacji z grupą pięciu ekspertów (publikujących psychologów) wybraliśmy następujące czasopisma: *Personality and Individual Differences* (PAiD; 100 punktów na liście MEiN (Komunikat Ministra Edukacji i Nauki, 2023), 4,3 IF), *Frontiers in Psychology* (*Frontiers*; 70 punktów, 3,8 IF), *Roczniki Psychologiczne* (*Roczniki*; 70 punktów), *Polskie Forum Psychologiczne* (PFP; 40 punktów).

Kryteria włączenia i wykluczenia

Artykuły zakwalifikowane do analiz musiały spełniać następujące kryteria: 1) przedstawiać ilościowe badania empiryczne, 2) zostać opublikowane jako recenzowane artykuły oryginalne, 3) zostać opublikowane w latach 2021–2022, 4) pierwszy autor musiał być afiliowany przy polskiej instytucji naukowej, 5) być napisane w języku polskim lub angielskim.

Nie było ograniczeń dotyczących liczby autorów, afiliacji pozostałych autorów (w publikacjach wieloautorskich), tematyki, liczby opisywanych badań, pochodzenia osób badanych i modelu badawczego.

Kryterium wykluczenia był brak pomiaru latentnych konstruktów psychologicznych lub brak dostępu do pełnej treści artykułu.

Proces wyboru artykułów do analizy

Na podstawie dokonanego przeglądu czasopism wyłonionych zostało 305 artykułów. Po wstępnej selekcji opartej na tytułach i abstraktach wykluczone zostały 63 artykuły, które nie spełniały kryteriów włączenia – były to głównie przeglądy literatury/metaanalizy (9), artykuły teoretyczne (25), badania jakościowe (8) i inne teksty nieempiryczne (21). Z pozostałych 242 tekstów kolejnych 7 wykluczono ze względu na brak dostępu do tekstu. Po krytycznej ocenie treści artykułów odrzucono kolejnych 5, w których nie mierzono zmiennych psychologicznych. Ostatecznie szczegółowy przegląd obejmował 230 artykułów. Ich lista została zamieszczona w tabeli w Załączniku 1 otwartym repozytorium (zob. https://osf.io/k4qsw/?view_only=dc40b39f6a9746b9a5781a2d8c5bfa67).

W celu zapewnienia zgodności oceny artykułów autorzy przeszli specjalne szkolenie, polegające na wykonaniu procedury kodowania na czterech tekstach. Ponadto, po samodzielnym opracowaniu przydzielonej części artykułów, w całym zespole konsultowano napotkane wątpliwości.

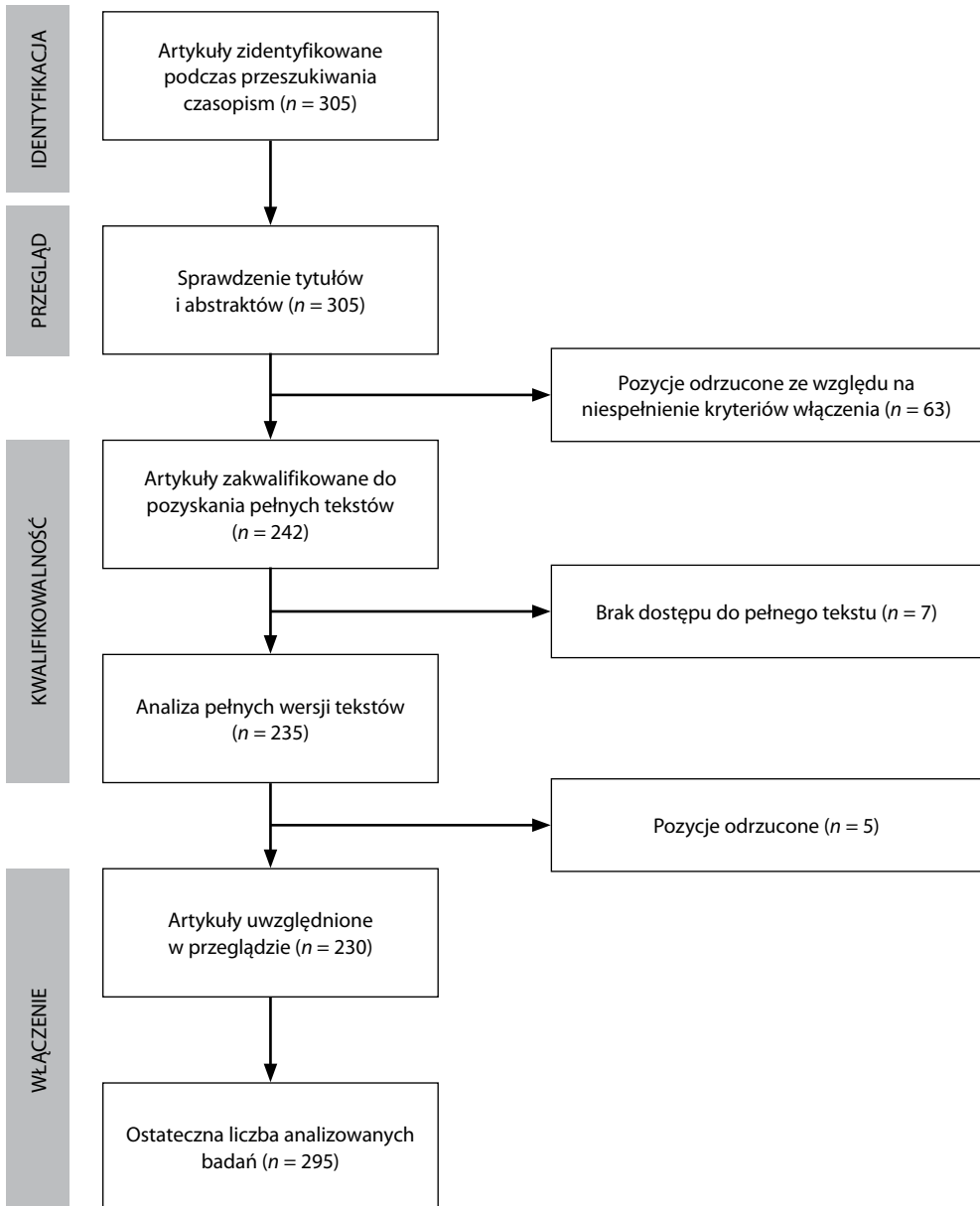
Modalnie jeden tekst przedstawiał opis pojedynczego badania ($n = 185$), a w 30 pracach opisywano ciąg dwóch badań na różnych próbach, w 11 – trzech, w 3 – czterech, zaś jedno dotyczyło cyklu pięciu badań. Ponieważ każde z raportowanych w artykułach badań zawierało informacje o respondentach, zostały one włączone do analiz jako odrębne pozycje. Innymi słowy, jeżeli w artykule

omówione były dwa badania, to opis próby każdego z nich kodowany był osobno. W efekcie wyłoniliśmy 295 opisów prób.

Etapy procesu wyłaniania danych przedstawiono na rysunku 1.

Rysunek 1

Przebieg procedury wyboru artykułów



Kodowane zmienne

Wytyczne opracowania raportu z przeprowadzonych badań empirycznych nie zawierają jednoznacznych kryteriów dotyczących charakterystyki respondentów. Dlatego przed przystąpieniem do kodowania wyróżniliśmy pożądane elementy opisu prób, które bazując na naszych dotychczasowych doświadczeniach i wiedzy metodologicznej, powinny zostać podane, aby można było dokonać oceny jakości zaprezentowanych wniosków czy próby replikacji badania. Należą do nich następujące informacje:

- 1) *szczegółowy opis procedury badania*, dotyczący czynności wykonywanych przez uczestników, liczby i kolejności kwestionariuszy oraz innych informacji specyficznych dla opisywanego badania;
- 2) *docelowa populacja*, na którą badacze planują uogólniać wnioski;
- 3) zastosowana *metoda doboru badanych*;
- 4) *forma badania* (stacjonarna czy online);
- 5) *kryteria włączenia uczestników badania*;
- 6) *kryteria wykluczania respondentów*;
- 7) *odsetek osób wykluczonych* (czy podano lub był możliwy do wyliczenia);
- 8) *wielkość próby*, czyli ostateczna liczba badanych osób;
- 9) *uzasadnienie wielkości próby*;
- 10) *kiedy badanie zostało przeprowadzone*;
- 11) *kto prowadził badanie*;
- 12) *wynagrodzenie dla osób badanych*;
- 13) *dane demograficzne*, które obejmują informacje dotyczące:
 - a) *płci*,
 - b) *wieku* – w przypadku pomiaru ilościowego podanego jako średnia z odchyleniem standardowym, a dla pomiaru porządkowego – podanego jako frekwencja osób należących do danej grupy wiekowej;
 - c) *wykształcenia*; w przypadku grup specyficznych, takich jak dzieci i uczniowie, uznawaliśmy, że jego poziom jest podany *implicite*,
 - d) *narodowości* podawanej *explicite*, gdyż wśród mieszkańców Polski, studentów i osób wypełniających kwestionariusze w języku polskim mogą znaleźć się obcokrajowcy posługujący się biegle tym językiem.

Ponadto uwzględnialiśmy także informacje o tym, czy był to *eksperyment* i czy badanie było prowadzone *z udziałem studentów* (w całości lub częściowo). Szczegółowe podsumowanie wymienionych zmiennych i sposobu ich kodowania zostało zamieszczone w tabeli w Załączniku 2 otwartym repozytorium (zob. https://osf.io/k4qsw/?view_only=dc40b39f6a9746b9a5781a2d8c5bfa67).

Rezultaty

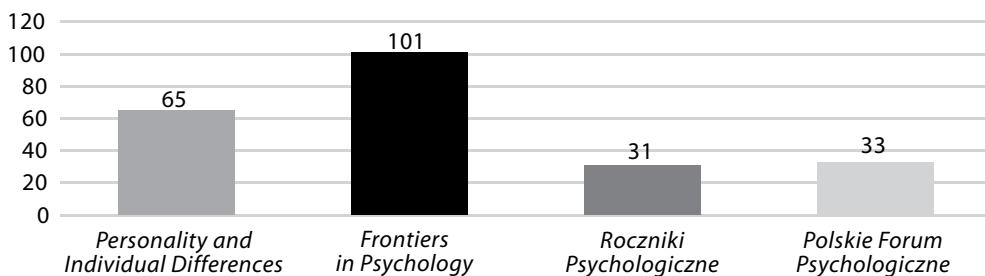
Ogólna charakterystyka artykułów

Rysunek 2 (s. 43) przedstawia udział każdego z czasopism w analizowanej bazie artykułów. Najwięcej tekstów zostało opublikowanych we *Frontiers* (43,9%),

zaś najmniej – w *Rocznikach* (13,5%). Tabela 1 przedstawia liczbę badań opisywanych w artykułach.

Rysunek 2

Liczba artykułów uwzględnionych w analizach w podziale na czasopisma ($N = 230$)



Wśród opisywanych badań wystąpiło 45 eksperymentów (15,3%), a 9 (3,1%) było badaniami longitudinalnymi. W pozostałych opisywano związki między zmiennymi lub porównywano różne populacje. Podejmowana przez autorów problematyka była bardzo różnorodna. Można zauważyć wyróżniający się odsetek badań poświęconych COVID-19 ($n = 33$; 11,2%), co jest naturalnym zjawiskiem dla analizowanego okresu. Na drugim miejscu znalazły się adaptacje lub propozycje nowych narzędzi ($n = 17$; 5,8%).

Tabela 1

Liczba badań opisywanych w artykułach w podziale na czasopisma ($N = 295$)

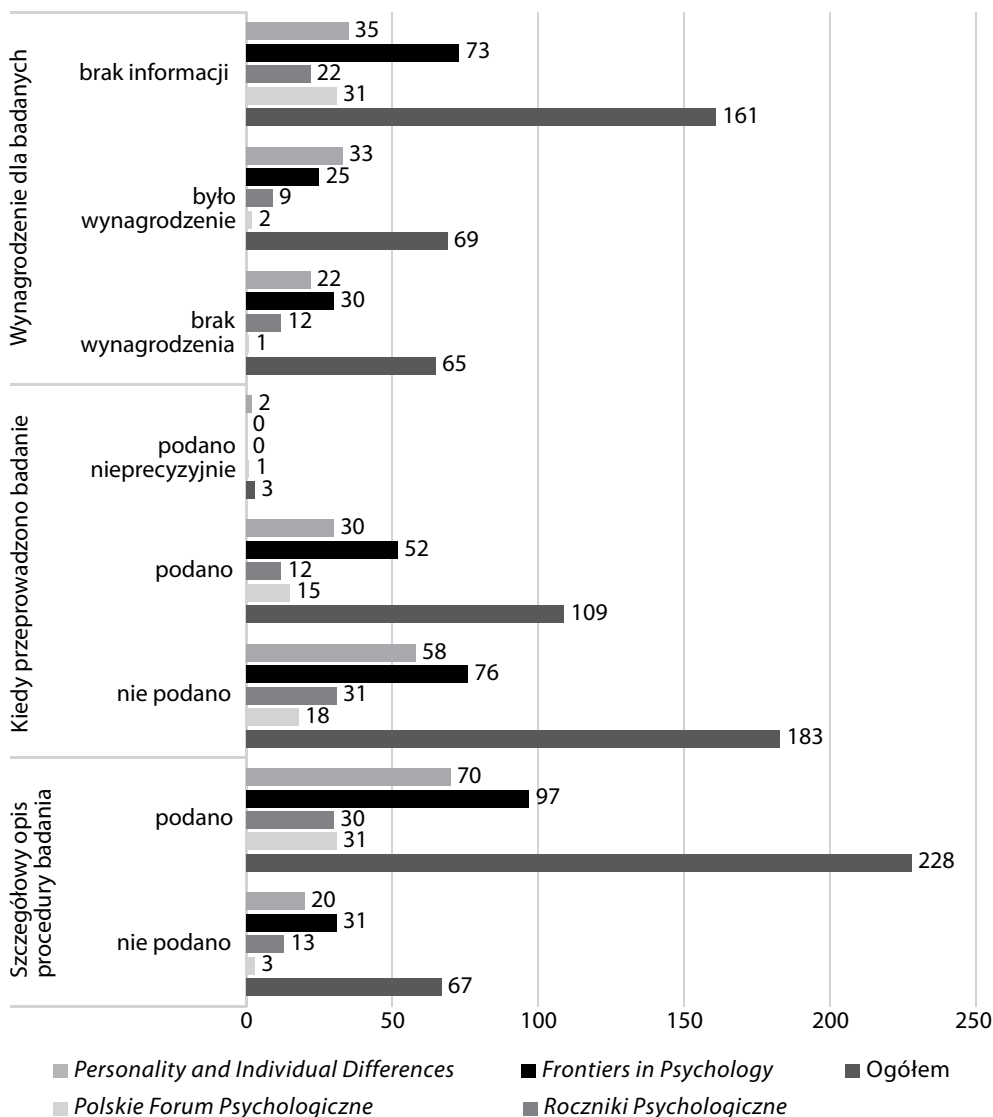
Liczba badań w artykule	<i>Personality and Individual Differences</i>		<i>Frontiers in Psychology</i>		<i>Roczniki Psychologiczne</i>		<i>Polskie Forum Psychologiczne</i>		Ogółem	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
1	46	70,8	86	85,1	21	68,8	32	97,0	185	80,5
2	14	21,5	7	6,9	8	25,0	1	3,0	30	13,0
3	4	6,2	5	5,0	2	6,3			11	4,8
4	1	1,5	2	2,0					3	1,3
5			1	1,0					1	0,4
N	90	30,5	128	43,4	43	14,6	34	11,5	295	100,0

Szczegółowy opis procedury badania, termin badania, kto je prowadził i o wynagrodzeniu uczestników

Rysunek 3 przedstawia, czy w opisach wystąpiły: a) szczegółowe dane na temat procedury badania, pozwalające na ich późniejszą replikację; b) informacje o tym, kiedy zostało przeprowadzone oraz c) informacje o wynagrodzeniu dla badanych.

Rysunek 3

Obecność informacji o szczegółowej procedurze i terminie badania oraz o wynagrodzeniu dla respondentów w podziale na czasopisma (N = 295)



W 77,3% badań niezależnie od czasopisma podawano wystarczające informacje, umożliwiające odtworzenie podejmowanych przez badaczy kroków, $\chi^2(3) = 5,3; p = 0,151$. Jeżeli ich brakowało, to głównie z powodu zastosowania doboru celowego (ze względu na zainteresowanie specyficzną populacją, np. przewlekle chorzy) i poprzestanie jedynie na prezentacji narzędzi pomiarowych ($n = 16; 23,9\%$) lub otrzymanie danych od firmy zewnętrznej (panelu badawczego; $n = 13; 19,4\%$). Biorąc pod uwagę formę badania, to w tych prowadzonych online odsetek szczegółowych opisów jest mniejszy ($n = 89; 74,2\%$) niż w stacjonarnych ($n = 122; 87,8\%$) i mieszanych ($n = 12; 92,3\%$).

Informacja o terminie, w jakim odbyło się badanie, pojawiała się jedynie w 38,0% tekstów ($n = 112$). Jest to drugie kryterium, które wypadło podobnie we wszystkich czasopismach ($\chi^2(6) = 7,9; p = 0,249$). Prawdopodobnie jest to efekt procesu publikacyjnego – długiego czasu oczekiwania na decyzje czasopisma oraz braku pewności co do tego, że w pierwszym czasopiśmie tekst zostanie opublikowany.

Informacje o nagradzaniu badanych znajdowały się w 45,4% badań ($n = 134$), aczkolwiek frekwencja ich występowania różniła się między czasopismami, $\chi^2(6) = 31,5; p < 0,001$. Najczęściej raportowano o tym w artykułach w *PaiD* ($n = 55; 61,1\%$), w których respondenci na ogół (36,7%) otrzymywali wynagrodzenie za udział (m.in. w formie voucherów, punktów, upominków). Po przeciwnej stronie znalazły się badania w *PPF* – tam jedynie trzy razy (8,8%) odwoływano się do kwestii nagradzania respondentów.

Tabela 2

Informacje o tym, kto prowadził badanie w podziale na czasopisma (N = 295)

Prowadzący badanie	<i>Person- ality and Individual Differences</i>		<i>Frontiers in Psychol- ogy</i>		<i>Roczniki Psycholo- giczne</i>		<i>Polskie Forum Psycholo- giczne</i>		Ogółem	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Badacze/autorzy	11	12,2	25	19,5	1	2,3	1	2,9	38	12,9
Firma badawcza	33	36,7	3	2,3	4	9,3	0	0	40	13,6
Studenci	2	2,2	1	0,8	1	2,3	1	2,9	5	1,7
Pracownicy placówki (np. psychologowie, pielęgniarki, pra- cownicy więzienia)	3	3,3	7	5,5	0	0	0	0	10	3,4
Specjaliści	2	2,2	4	3,1	0	0	4	11,8	10	3,4
Asystenci naukowci	5	5,6	10	7,8	2	4,7	1	2,9	18	6,1
Brak informacji	34	37,8	78	60,9	35	81,4	27	79,4	174	59,0

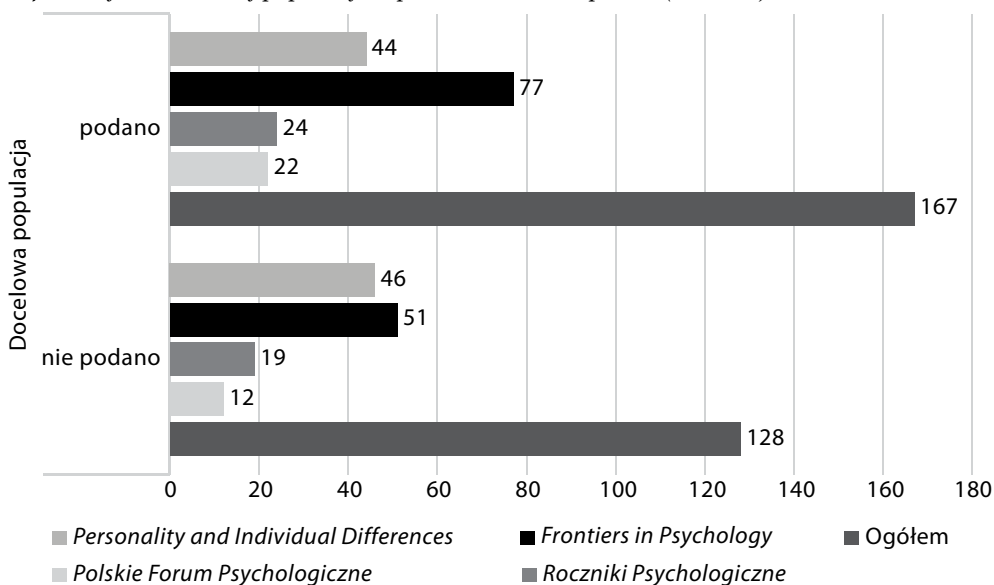
Również występowanie informacji o tym, kto prowadził badanie (tabela 2, s. 45), istotnie różniło się między czasopismami, $\chi^2(6) = 31,5; p < 0,001$. Modalnym przypadkiem (59,0%) był brak informacji. Jest to niepokojący sygnał, ale prawdopodobną przyczyną tak wysokiego odsetka jest bezosobowa forma artykułu naukowego, a w języku angielskim – stosowana strona bierna. Dodatkową uwagę zwraca popularność paneli badawczych (m.in. Ariadna, Prolificm, Pollster) w publikacjach PaiD (36,7%).

Docelowa populacja, metoda doboru próby i forma badania

W 167 badaniach (56,6%) autorzy przeprowadzali je w odniesieniu do konkretnej populacji (np. adaptując narzędzie; zob. Nickel, 2021) lub uogólniali wnioski na szerszą zbiorowość (rysunek 4). Nie stwierdzono istotnych statystycznie różnic między czasopismami w tym zakresie, $\chi^2(3) = 3,8; p = 0,289$.

Rysunek 4

Informacje o docelowej populacji w podziale na czasopisma ($N = 295$)



Również obecność informacji o zastosowanej metodzie doboru próby utrzymywała się na zbliżonym poziomie ok. 70% we wszystkich czasopismach, $\chi^2(3) = 0,3; p = 0,959$. Niemniej metodą doboru dającą największe szanse na uzyskanie próby reprezentatywnej jest dobór losowy. Ten wystąpił jedynie w pięciu przypadkach (1,7%) opisanych we *Frontiers*. W opisach dominowały: dobór celowy (24,4%), wygodny (16,6%) oraz panele badawcze (13,6%). Należy zwrócić uwagę, że w 27,8% badań informacja o metodzie doboru nie została podana. Występujące ich rodzaje w podziale na czasopisma przedstawiono w tabeli 3.

Tabela 3*Metody doboru próby występujące w badaniach w podziale na czasopisma (N = 295)*

Metoda doboru	<i>Person- ality and Individual Differences</i>		<i>Frontiers in Psy- chology</i>		<i>Roczniki Psycholo- giczne</i>		<i>Polskie Forum Psycholo- giczne</i>		Ogółem	
	<i>n</i>	<i>%</i>	<i>n</i>	<i>%</i>	<i>n</i>	<i>%</i>	<i>n</i>	<i>%</i>	<i>n</i>	<i>%</i>
Dobór losowy	0	0	5	3,9	0	0	0	0	5	1,7
Panel badawczy	33	36,7	4	3,1	3	7,0	0	0	40	13,6
Dobór celowy	9	10,0	41	32,0	11	25,6	11	32,4	72	24,4
Dobór kwotowy	1	1,1	0	0	0	0	1	2,9	2	0,7
Ochotnicy	3	3,3	4	3,1	2	4,7	4	11,8	13	4,4
Kula śnieżna	3	3,3	6	4,7	0	0	5	14,7	14	4,7
Dobór wygodny	13	14,4	25	19,5	8	18,6	3	8,8	49	16,6
Dobór seryjny	0	0	3	2,3	0	0	1	2,9	4	1,4
MTurk	0	0	4	3,1	0	0	0	0	4	1,4
Dobór przypadkowy	2	2,2	2	1,6	6	14,0	0	0	10	3,4
Brak informacji	26	28,9	34	26,6	13	30,2	9	26,5	82	27,8

W zdecydowanej większości badań (92,2%), niezależnie od czasopisma, znajdowała się informacja o tym, w jakiej formie było ono prowadzone: stacjonarnej, online i mieszanej (tabela 4), $\chi^2(3) = 2,6$; $p = 0,453$. Modalną formą we wszystkich czasopismach, z wyjątkiem PaiD, były badania stacjonarne, stanowiące ponad połowę przypadków. Natomiast we wspomnianym periodyku dominowały badania online (57,8%), co może wynikać z faktu, że prowadzono je w czasie pandemii.

Tabela 4*Forma badania w podziale na czasopisma (N = 295)*

Forma badania	<i>Person- ality and Individual Differences</i>		<i>Frontiers in Psychology</i>		<i>Roczniki Psycholo- giczne</i>		<i>Polskie Forum Psycholo- giczne</i>		Ogółem	
	<i>n</i>	<i>%</i>	<i>n</i>	<i>%</i>	<i>n</i>	<i>%</i>	<i>n</i>	<i>%</i>	<i>n</i>	<i>%</i>
Stacjonarne	27	30,0	71	55,5	21	48,8	20	58,8	139	47,1
Online	52	57,8	41	32,0	16	37,2	11	32,4	120	40,7
Mieszane	2	2,2	9	7,0	1	2,3	1	2,9	13	4,4
Brak informacji	9	10,0	7	5,5	5	11,6	2	5,9	23	7,8

Wielkość próby i jej uzasadnienie, kryteria włączenia i wykluczenia oraz odsetek osób wykluczonych

Liczebność prób była bardzo zróżnicowana i wahała się od 10 do 7108 osób. W tabeli 5 przedstawiono podstawowe statystyki charakteryzujące rozkład tej zmiennej dla całego zbioru danych i w podziale na czasopisma. Otrzymane rozkłady okazały się mocno prawoskośne, co oznacza przewagę prób liczebności mniejszej od średniej. W konsekwencji, a także ze względu na dużą rozbieżność we frekwencji badań z tych czasopism: 1) mediana wydaje się lepszą miarą charakteryzującą te rozkłady oraz 2) zastosowanie testu nieparametrycznego do porównania wielkości prób między periodykami byłoby obciążone większym ryzykiem błędu ze względu na różny stopień zniekształcenia rozkładów.

Tabela 5

Charakterystyka rozkładów wielkości prób w podziale na czasopisma (N = 295)

	<i>Personality and Individual Differences</i>	<i>Frontiers in Psychology</i>	<i>Roczniki Psychologiczne</i>	<i>Polskie Forum Psychologiczne</i>	Ogółem
<i>M</i>	482,4	252,7	313,6	450,6	417,5
[95% CI]	[403,9; 561,7]	[176,3; 329,0]	[185,6; 441,7]	[290,1; 611,1]	[341,2; 493,8]
SD	378,6	218,7	416,0	917,5	666,1
Me	356,0	145,0	200,0	251,5	258,0
Q	437,0	417,0	192,0	337,0	385,0
Min	84,0	20,0	40,0	10,0	10,0
Max	1961,0	893,0	2390,0	7108,0	7108,0

Chcąc czytelniej przedstawić rozkład analizowanej zmiennej i umożliwić eksplorację potencjalnych źródeł różnic w wielkości porównywanych prób, dane ilościowe zostały przekształcone w zmienną porządkową (tabela 6, s. 49). Okazuje się, że rodzaj periodyku istotnie różnicuje odsetek występowania poszczególnych kategorii, $\chi^2(12) = 36,3$; $p < 0,001$. W przypadku PaiD modalnie badano od 251 do 500 osób (31,1%), a w *Rocznikach* najczęściej – od 131 do 250 respondentów (37,2%). We *Frontiers* i PFP najwięcej było badań na małych próbach, liczących do 130 osób.

Ponadto podjęto próbę ustalenia, co jeszcze może różnicować wielkość prób. Zaobserwowano, że szczegółowy opis badania [$\chi^2(4) = 9,8$; $p < 0,05$], kto je

przewodził [$\chi^2(4) = 13,0; p < 0,05$] i kiedy do niego doszło [$\chi^2(8) = 26,8; p < 0,001$] występowały częściej w przypadku większych prób (liczących od 251 do 1000 osób) niż mniejszych (do 250 osób). Szczegółowe wyniki znajdują się w Załączniku 3 w otwartym repozytorium (zob. https://osf.io/k4qsw/?view_only=dc40b39f6a9746-b9a5781a2d8c5bfa67).

Tabela 6

Klasy wielkości prób w podziale na czasopisma (N = 295)

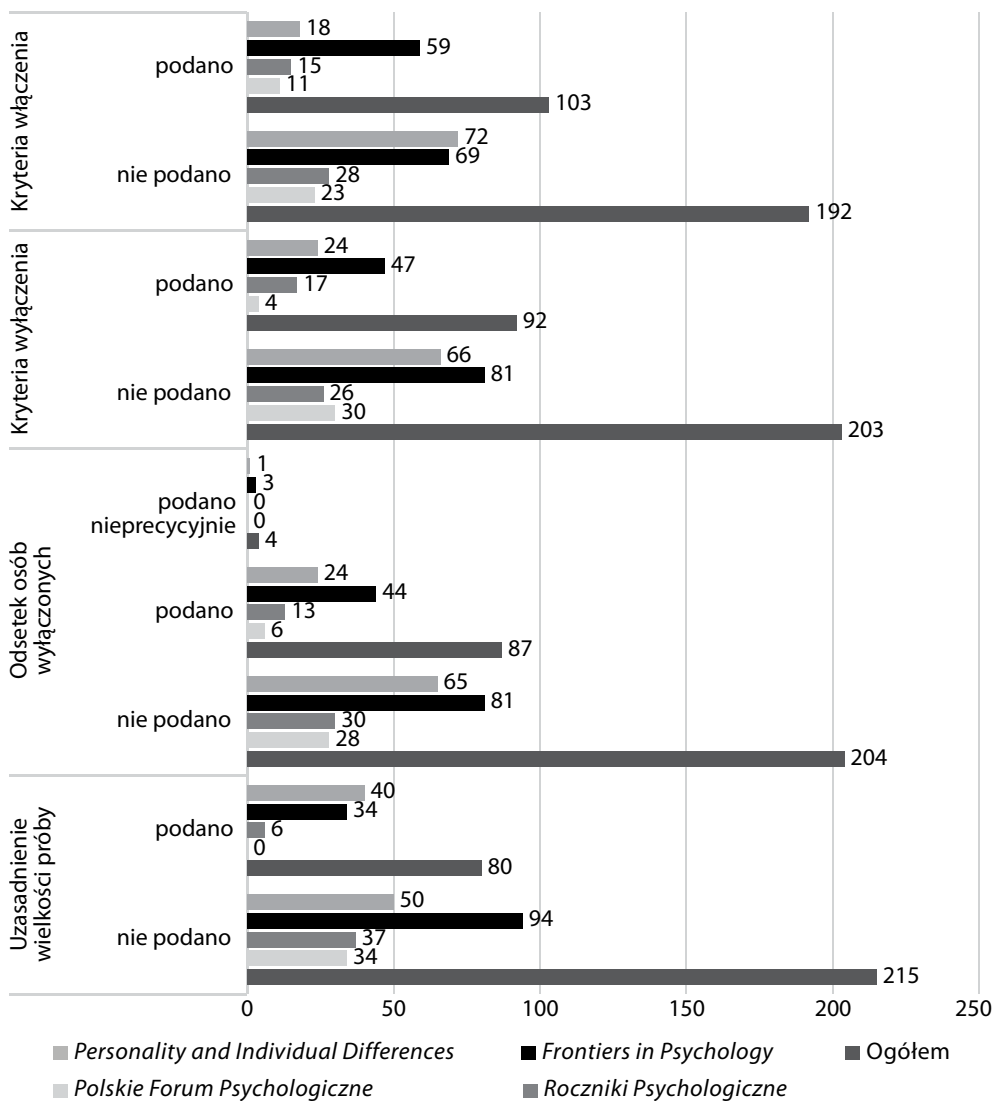
Wielkość próby	<i>Personality and Individual Differences</i>		<i>Frontiers in Psychology</i>		<i>Roczniki Psychologiczne</i>		<i>Polskie Forum Psychologiczne</i>		Ogółem	
	n	%	n	%	n	%	n	%	n	%
Próby małe (do 130 osób)	11	12,2	46	35,9	14	32,6	15	44,1	86	29,2
Próby od 131 do 250 osób	19	21,1	18	14,1	16	37,2	5	14,7	58	19,7
Próby od 251 do 500 osób	28	31,1	34	26,6	7	16,3	11	32,4	80	27,1
Próby od 501 do 1000 osób	22	24,4	21	16,4	4	9,3	3	8,8	50	16,9
Próby powyżej 1001 osób	10	11,1	9	7,0	2	4,7	0	0	21	7,1

W ok. jednej trzeciej badań podawano kryteria włączenia i wykluczenia respondentów. W przypadku tych pierwszych najczęściej raportowano je w artykułach publikowanych we *Frontiers* (46,1%), zaś najrzadziej – w *PaiD* (20,0%), $\chi^2(3) = 15,9; p < 0,001$. Natomiast kryteria wykluczenia najczęściej pojawiały się w *Rocznikach* (39,5%), a najrzadziej – w *PFP* (11,8%), $\chi^2(3) = 10,1; p < 0,05$. Z kolei informacja o odsetku lub liczbie osób wykluczonych pojawiała się w ok. 29,5% badań, niezależnie od czasopisma, $\chi^2(6) = 6,5; p = 0,367$.

Uzasadnienie wielkości próby pojawiało się w 27,1% badań (rysunek 5, s. 50). W tym aspekcie polskie czasopisma wypadły gorzej na tle czasopism zagranicznych. W *PFP* autorzy ani razu nie podejmowali tego wątku, a w *Rocznikach* – wystąpił jedynie sześć razy (14,0%). Najczęściej informacja ta pojawiła się we *Frontiers* (44,4%).

Rysunek 5

Informacje o kryteriach włączenia i wyłączenia, odsetku osób wykluczonych oraz uzasadnieniu wielkości próby w podziale na czasopisma ($N = 295$)



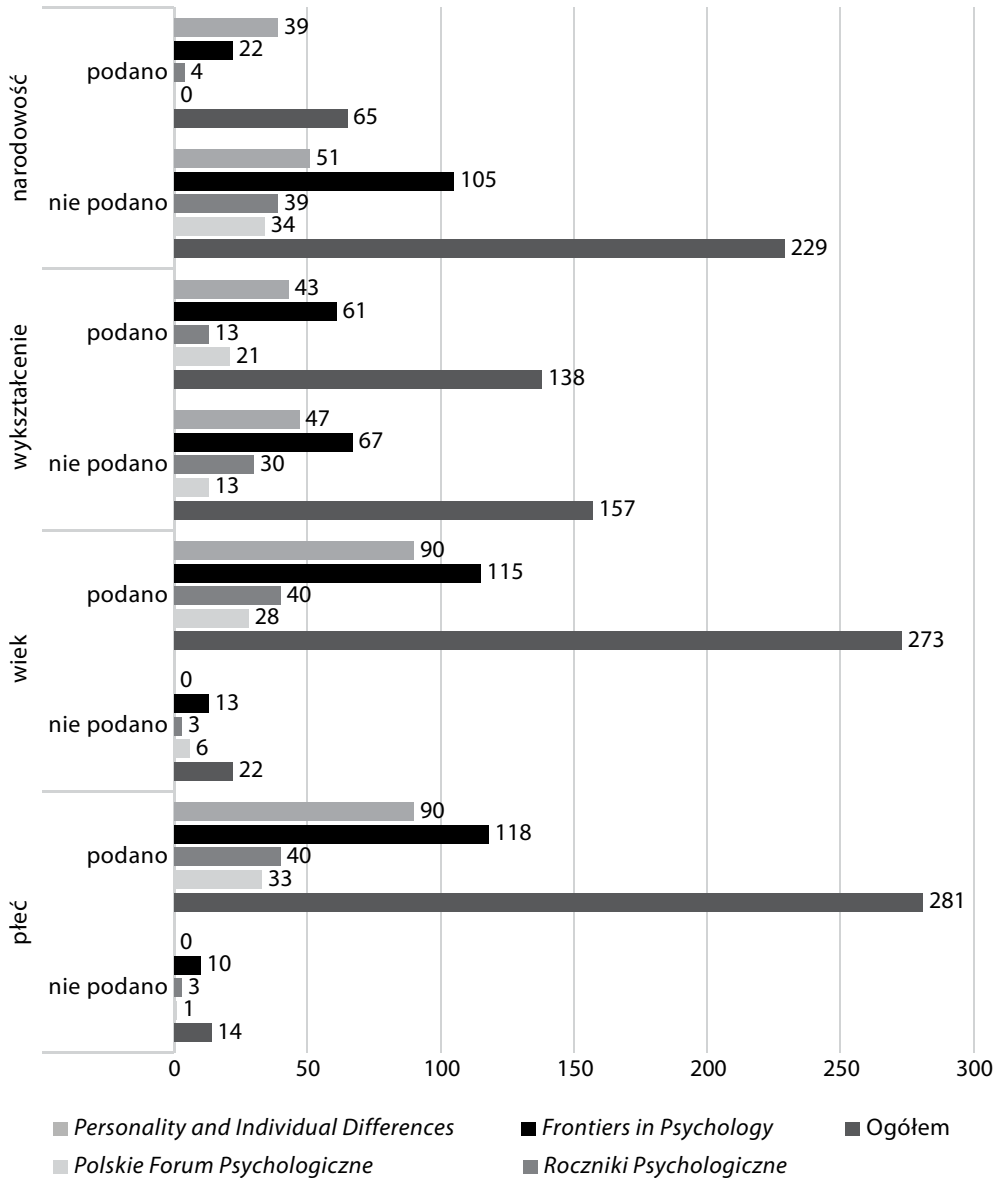
Dane demograficzne respondentów: płeć, wiek, wykształcenie, narodowość

W przypadku charakterystyki badanych uwzględniono najbardziej fundamentalne informacje: płeć, wiek, poziom wykształcenia i narodowość (rysunek 6, s. 51). Mimo że informacje o płci i o wieku występowały w ponad 90% opisów, to okazały

się istotnie różnić między periodykami. W przypadku płci za występowanie rozbieżności odpowiada głównie *PaiD*, w którym informacje te były podawane zawsze, $\chi^2(3) = 7,9; p < 0,05$. Podobnie przy podawaniu wieku, któremu towarzyszyło najczęstsze pomijanie tego elementu w opisach w PFP (17,6%), $\chi^2(3) = 13,7; p < 0,01$.

Rysunek 6

Informacje o danych demograficznych w podziale na czasopisma (N = 295)



Poziom wykształcenia najczęściej występował w artykułach PFP (61,8%), a najrzadziej – we *Frontiers* (30,2%), $\chi^2(3) = 7,9$; $p < 0,05$. Informacja taka pojawiła się w 46,8% tekstów przy stosowaniu liberalnego kryterium (tzn. w przypadku artykułów opisujących niemowlęta, przedszkolaków i uczniów przyjęto, że te informacje występują). Najrzadziej podawaną charakterystyką respondentów, niezależnie od czasopisma, była podawana wprost narodowość (22,1%). W publikacjach w Paid informację tę raportowano najczęściej (43,3%), natomiast w PFP nie wspomniano o tym ani razu, $\chi^2(3) = 39,0$; $p < 0,001$. Należy zwrócić uwagę, że językiem polskim mogą posługiwać się również obcokrajowcy.

Dyskusja

Celem niniejszego przeglądu była identyfikacja elementów występujących w opisach procedury badania, w szczególności dotyczących samej próby, a także sformułowanie na tej bazie rekomendacji dla utworów.

Spśród analizowanych przez nas elementów opisu najczęściej występowały: opis procedury umożliwiający jej replikację, zastosowana metoda doboru i wielkości próby, forma badania, a z cech socjodemograficznych: płeć i wiek. Jest to skromna liczba informacji, które wymagają komentarza.

Jak wykazały analizy, w 77,3% opisów badań uwzględniono informacje o przebiegu ich procedury. W badaniach stacjonarnych powinna być zawarta informacja, gdzie i w jakich okolicznościach zostały one przeprowadzone, np. w trakcie pracy/zajęć, podczas przerwy, indywidualnie/grupowo. Dodatkowo ważne jest nie tylko wylistowanie zastosowanych kwestionariuszy, lecz także podanie kolejności ich wypełniania (efekt kolejności; Bayat i in., 2023) oraz tego, jak zostały zebrane dane socjodemograficzne, np. ankieta własna. Te informacje szczególnie rzadko pojawiały się w badaniach online. W połączeniu ze wzrostem ich popularności wskazuje to na potrzebę identyfikacji informacji, które należałoby podawać.

Niemniej procedura badania to zarówno szczegółowy opis zadań uczestnika, jak i podsumowanie, co jest przedmiotem zainteresowania badacza (populacja czy zjawisko), w jakim okresie zbierane były dane, jak długo i przez kogo. Informacje te występowały zdecydowanie zbyt rzadko, a przecież każda z nich ma znaczenie dla metodologicznej oceny jakości wniosków.

Co więcej, chcąc odnaleźć informacje o przebiegu badania i respondentach, czytelnik musi uzbroić się w cierpliwość. Okazuje się, że sekcje artykułów poświęcone tej tematyce (np. *Próba*, *Procedura badania*) nie są jedynymi, gdzie umieszcza się te informacje. Często pojawiały się one w innych miejscach, np. *Narzędziach* czy *Limitacjach*. Zdarzyło się, że niektóre dane na temat próby odnotowano tylko w *Abstrakcie*. Utrudnia to zrozumienie, jak faktycznie przebiegał proces zbierania danych, co w konsekwencji może uniemożliwiać ocenę trafności wewnętrznej badania i jego replikację.

Mimo że zaledwie w 27,8% badań nie udało się zidentyfikować metody doboru próby, nie był to niestety efekt precyzji opisów. Badacze nie zawsze nazywali ją *explicitie*, ale na podstawie opisów można je było wywnioskować. Zdarzały się

też niespójności między deklarowaną metodą a późniejszą procedurą. Sytuacja komplikowała się jeszcze bardziej, gdy opisywano wiele różnych sposobów docierania do badanych. Często raportowano to jako metodę kuli śnieżnej. Należy podkreślić, że tylko jeśli większość osób została zwerbowana przez samych respondentów, można mówić o tej technice. Sama prośba o rozsyłanie informacji do kolejnych osób nie jest wystarczająca.

W ponad połowie badań autorzy odnosili się do konkretnej populacji lub uogólniali na nią swoje wnioski. Takie postępowanie jest zasadne, gdy dana próba jest reprezentatywnym odpowiednikiem populacji, a jedynym doborem, który umożliwi jej uzyskanie, jest dobór losowy. Niestety, ten wystąpił jedynie w pięciu badaniach. Z nieprobabilistycznych metod doboru najbliższy losowemu (Brzeziński, 2019) jest dobór kwotowy, który zastosowano tylko dwa razy.

W opisach badań online często pojawiała się informacja wyłącznie o umieszczeniu linku na stronach internetowych czy Facebooku, bez sprecyzowania dokładnego miejsca. Różnica pomiędzy przypadkowym doborem, gdy linki są umieszczane *wszędzie*, a celowym doborem, gdy wyszukiwane są *konkretne grupy* zrzeszające osoby o danych zainteresowaniach/zawodzie, wpływa na możliwość replikacji badań i na ocenę trafności zewnętrznej prezentowanych wniosków.

Coraz więcej uwagi poświęca się liczebności próby w związku z możliwością ujawnienia efektu o określonej wielkości oraz kontrolą błędów I i II rodzaju (Adam, 2020; Lakens, 2022). Tej informacji, w połączeniu z precyzyjnie podanymi kryteriami włączenia i wykluczenia oraz odsetkiem/liczbą osób wykluczonych, brakuje – szczególnie w badaniach publikowanych w krajowych czasopismach.

Ostatnim kryterium były dane socjodemograficzne. Zgodnie z założeniami informacje o płci i wieku pojawiały się niemal we wszystkich badaniach. Rzadziej występowały wzmianki o poziomie wykształcenia, a niemal w ogóle nie podawano informacji o narodowości, nawet w czasopismach międzynarodowych. Warto zaznaczyć, że zarówno próby z udziałem studentów, jak i osób mówiących po polsku (bez wzmianki, że jest to ich język ojczysty), nie muszą składać się wyłącznie z rodowitych Polaków.

Henrich z zespołem (2010) zarzucił psychologom, że prowadzą badania głównie na próbach pobranych z D.Z.I.W.Nego społeczeństwa (ang. *W.E.I.R.D. society*), tj. Demokratycznego (ang. *Democratic*), Zachodniego (ang. *Western*), Industrialnego (ang. *Industrialized*), Wykształconego (ang. *Educated*) i Na wysokim poziomie finansowym (ang. *Rich*). Natomiast Kaźmierczak z zespołem (2023) wyraziła niepokój, że do badań zgłaszają się jedynie osoby z zaburzeniami afektywnymi i osobowości. Postulujemy, aby wyjść poza schematy, starannie dobierać respondentów, a następnie rzetelnie raportować: kiedy, kogo, w jaki sposób, w jakich okolicznościach zbadano i kto przeprowadził badanie.

Rekomendacje dla badaczy

Na podstawie doświadczeń zebranych w trakcie przygotowywania niniejszego przeglądu sformułowaliśmy następujące rekomendacje dla publikujących badaczy:

- 1) podając procedurę badania i próbę, należy rozważyć, które informacje są niezbędne do przeprowadzenia prawidłowej replikacji – i je opisać;

- 2) absolutne minimum informacji w opisie próby to: procedura badania umożliwiająca replikację, zastosowana metoda doboru i wielkość próby, forma badania, a z cech socjodemograficznych: płeć i wiek; niemniej jednak informacje te powinny zostać rozszerzone o pozostałe cechy uwzględnione w niniejszym przeglądzie;
- 3) omawiane elementy wskazane byłoby umieścić w przeznaczonym do tego miejscu w artykule – sekcji *Próba czy Procedura badania*;
- 4) w przypadku badań online konieczne jest określenie standardu informacji niezbędnych do oceny metodologicznej; proponujemy, aby były to następujące informacje:
 - a) platforma wykorzystana do zbierania danych (np. Google Forms, Qualtrics, LimeSurvey),
 - b) sposób dystrybucji linku do ankiet (np. wysyłka mailowa, umieszczanie w social mediach – jakich?) i kto był odpowiedzialny za jego rozsyłanie,
 - c) okres zbierania danych,
 - d) procedura badania, czyli jakie kwestionariusze i w jakiej kolejności wypełniano,
 - e) występowanie pytań weryfikujących rzetelność respondentów,
 - f) kryteria włączenia (przynajmniej w postaci pytania weryfikującego przynależność do interesującej badacza populacji) i kryteria wykluczenia z odsetkiem osób wykluczonych.

Ograniczenia

Przeprowadzona przez nas analiza nie jest wolna od ograniczeń. Po pierwsze, skoncentrowaliśmy się na artykułach z arbitralnie wybranych czasopism o różnym prestiżu. Po drugie, badania z artykułów obejmujących ich cykl były wprowadzane do analiz osobno. Mogło to skutkować większym udziałem tekstów, w których mniej miejsca poświęcono opisom prób. Po trzecie, analizom poddane zostały zarówno badania eksperymentalne, jak i dotyczące różnic indywidualnych, prowadzone (głównie) w schematach międzyosobowych, ale także wewnątrzosobowych. W przypadku tych ostatnich sformułowane w niniejszym artykule wnioski mają mniejsze, ale niebagatelne znaczenie (por. Thompson i Campbell, 2004; Wojciszke, 2004). Dlatego wskazane byłoby przeprowadzenie odrębnych analiz skoncentrowanych wyłącznie na badaniach w schematach wewnątrzosobowych. Po czwarte, niejednoznaczne opisy doboru skutkowały zakwalifikowaniem do kategorii *brak informacji*, co zapewne spowodowało jej nadreprezentację ($n = 82$) i niedoreprezentowanie, np. doboru przypadkowego ($n = 10$). Przeprowadzony przegląd miał charakter eksploracyjny, więc warto byłoby go powtórzyć na innej puli czasopism, a także przyrzeć się trendom międzynarodowym.

Bibliografia

Adam, A. M. (2020). Sample Size Determination in Survey Research. *Journal of Scientific Research and Reports*, 26(5), 90–97. <https://doi.org/10.9734/jsrr/2020/v26i530263>

- Babbie, E. (2021). *Badania społeczne w praktyce*. Wydawnictwo Naukowe PWN.
- Baker, R., Brick, J. M., Bates, N. A., Battaglia, M., Couper, M. P., Dever, J. A., Gile, K. J., Tourangeau, R. (2013). Summary Report of the AAPOR Task Force on Non-probability Sampling. *Journal of Survey Statistics and Methodology*, 1(2), 90–143. <https://doi.org/10.1093/jssam/smt008>
- Bayat, D., Mohamadpour, H., Fang, H., Xu, P., Krueger, F. (2023). The Impact of Order Effects on the Framing of Trust and Reciprocity Behaviors. *Games*, 14(2), artykuł 21. <https://doi.org/10.3390/g14020021>
- Bethlehem, J. (2010). Selection bias in web surveys. *International Statistical Review*, 78(2), 161–188. <https://doi.org/10.1111/j.1751-5823.2010.00112.x>
- Brase, G. L. (2009). How different types of participant payments alter task performance. *Judgment and Decision Making*, 4(5), 419–428. <https://doi.org/10.1017/S193-0297500001248>
- Brzeziński, J. (2019). *Metodologia badań psychologicznych. Wydanie nowe*. Wydawnictwo Naukowe PWN.
- Cheung, K. L., ten Klooster, P. M., Smit, C., de Vries, H., Pieterse, M. E. (2017). The impact of non-response bias due to sampling in public health studies: A comparison of voluntary versus mandatory recruitment in a Dutch national survey on adolescent health. *BMC Public Health*, 17, artykuł 276. <https://doi.org/10.1186/s12889-017-4189-8>
- Faul, F., Erdfelder, E., Lang, A.-G., Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Henrich, J., Heine, S. J., Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2–3), 61–83. <https://doi.org/10.1017/S0140525X0999152X>
- Kaźmierczak, I., Zajenkowska, A., Rogoza, R., Jonason, P. K., Ścigała, D. (2023). Self-selection biases in psychological studies: Personality and affective disorders are prevalent among participants. *PLOS ONE*, 18(3), artykuł e0281046. <https://doi.org/10.1371/journal.pone.0281046>
- Komunikat Ministra Edukacji i Nauki z dnia 17 lipca 2023 r. w sprawie wykazu czasopism naukowych i recenzowanych materiałów z konferencji międzynarodowych (2023) (Polska). <https://www.gov.pl/web/edukacja-i-nauka/komunikat-ministra-edukacji-i-nauki-z-dnia-17-lipca-2023-r-w-sprawie-wykazu-czasopism-naukowych-i-recenzowanych-materialow-z-konferencji-miedzynarodowych>
- Lakens, D. (2022). Sample size justification. *Collabra: Psychology*, 8(1), artykuł 33267. <https://doi.org/10.1525/collabra.33267>
- Lusinchi, D. (2018). ‘The Great Fiasco’ of the 1948 presidential election polls: status recognition and norms conflict in social science. *Annals of Science*, 75(2), 120–144. <https://doi.org/10.1080/00033790.2018.1466194>
- Maslow, A. H., Sakoda, J. M. (1952). Volunteer-error in the Kinsey study. *The Journal of Abnormal and Social Psychology*, 47(2), 259–262. <https://doi.org/10.1037/h0054411>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLOS Medicine*, 6(7), artykuł e1000097. <https://doi.org/10.1371/journal.pmed.1000097>

- Nikel, Ł. (2021). Polska adaptacja Skali do badania motywacji uczniów w szkole podstawowej. *Polskie Forum Psychologiczne*, 26(1), 97–112. <https://doi.org/10.34767/PFP.2-021.01.06>
- Salganik, M. J., Heckathorn, D. D. (2004). Sampling and Estimation in Hidden Populations Using Respondent-Driven Sampling. *Sociological Methodology*, 34(1), 193–240. <https://doi.org/10.1111/j.0081-1750.2004.00152.x>
- Sapra, R. L. (2022). How to Calculate an Adequate Sample Size? W: S. Nundy, A. Kakar i Z. A. Bhutta (red.), *How to Practice Academic Medicine and Publish from Developing Countries? A Practical Guide* (s. 81–93). Springer Singapore. <https://doi.org/10.1007/978-981-16-5248-6>
- Sarstedt, M., Bengart, P., Shaltoni, A. M., Lehmann, S. (2018). The use of sampling methods in advertising research: a gap between theory and practice. *International Journal of Advertising*, 37(4), 650–663. <https://doi.org/10.1080/02650487.2017.1348329>
- Singh, A. S., Masuku, M. B. (2014). Sampling techniques and determination of sample size in applied statistics research: An overview. *International Journal of Economics, Commerce and Management*, 2(11), 1–22.
- Shaughnessy, J. J., Zechmeister, E. B., Zechmeister, J. S. (2002). *Metody badawcze w psychologii*. Gdańskie Wydawnictwo Psychologiczne.
- Squire, P. (1988). Why the 1936 Literary Digest Poll Failed. *The Public Opinion Quarterly*, 52(1), 125–133. <http://www.jstor.org/stable/2749114>
- Thompson, V. A., Campbell, J. I. D. (2004). A power struggle: Between- vs. within-subjects designs in deductive reasoning research. *Psychologia: An International Journal of Psychology in the Orient*, 47(4), 277–296. <https://doi.org/10.2117/psysoc.2004.277>
- Ustawa z dnia 19 listopada 2009 r. o grach hazardowych Dz.U. 2009 nr 201, poz. 1540 (2019) (Polska). <https://isap.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU20092011540>
- Wojciszke, B. (2004). Systematycznie modyfikowane autoreplikacje: logika programu badań empirycznych w psychologii. W: J. M. Brzeziński (red.), *Metodologia badań psychologicznych. Wybór tekstów* (s. 44–68). Wydawnictwo Zysk i Ska.
- Zimbardo, P. G., Gerrig, R. J. (2012). *Psychologia i życie*. Wydawnictwo Naukowe PWN.

Podziękowania

Autorzy chcieliby wyrazić podziękowanie dla reszty zespołu, która służyła wsparciem w procesie przygotowywania niniejszego artykułu. Szczególne podziękowania należą się Angelice Rusin i Julii Szymańskiej.