

Samples in Psychological Research – Conclusions and Recommendations from the Literature Review

Joanna Świdowska¹

The Maria Grzegorzewska University, Institute of Psychology

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

Magdalena Puchalska

The Maria Grzegorzewska University, Institute of Psychology

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

Maciej Grzegorzczak

Cardinal Stefan Wyszyński University in Warsaw, Institute of Psychology

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

Zuzanna Szalek

Cardinal Stefan Wyszyński University in Warsaw, Institute of Psychology

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

Natalia Nadolna

Cardinal Stefan Wyszyński University in Warsaw, Institute of Psychology

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

Martyna Andryszkiewicz

University of Lodz, Institute of Psychology

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

Maria Jastrzębska

Cardinal Stefan Wyszyński University in Warsaw, Institute of Psychology

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

Przemysław Szablowski

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

Maja Śmigrodzka

The Maria Grzegorzewska University, Institute of Psychology

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

¹ Correspondence address: jswidowska@aps.edu.pl.

Abstract

Objective: The goal of the review was to identify the elements occurring in the both survey procedure and sample descriptions and to formulate recommendations for the authors on this basis.

Method: The study was conducted following the PRISMA standard. A total of 305 articles by Polish researchers from four journals were analyzed. Critical evaluation have been conducted for 230 articles of the analyzed sample, encompassing 295 research descriptions.

Results: Analysis revealed that researchers most frequently report the research procedure, sampling method, sample size, form of the research and sociodemographic characteristics, such as: gender or sex and age.

Conclusions: The conclusions drawn from the analysis and insights gained from the review led to the formulating recommendations regarding comprehensive methodological description of research procedures and samples.

Keywords: sampling, research procedure description, data quality, generalization, literature review

Psychology is an empirical science that deals with human behavior and cognitive processes (Zimbardo & Gerrig, 2012). It means that psychologists study a certain number of individuals and, through inductive inference, attempt to describe, explain, and predict actions of an individual and their underlying mechanisms. In order maximize the extent, which the conclusion can be generalized to, research should be based on representative samples.

Adequate sampling is, therefore, crucial to conducting a valid survey, and insufficient attention to this matter can lead to invalid results. One of the most famous methodological errors are the *Literary Digest* magazine's pre-election polls of 1936 that predicted Landon's success (Babbie, 2021). According to Squire (1988), the factors that contributed to the wrong prediction were: 1) low response rate (2 out of 10 million surveys received); 2) inadequate sample, which, as Babbie (2021) points out, comprised of representatives of the affluent part of the society.

Nevertheless, the described failure still rendered as beneficial for Gallup, who, using quota sampling, accurately predicted the election outcome. Unfortunately, nine years later, when Gallup together with Crossley and Roper used the same method used to predict Dewey's victory, it was not successful (Lusinchi, 2018). While planning the quota sampling, they did not take into account the transformation of the social structure caused by migrations from rural to urban areas. Lack of due diligence in selecting a representative sample can result in inaccurate findings and adverse consequences for the researchers, which led *Literary Digest* to bankruptcy two years following the regrettable elections (Babbie, 2021). On the other hand, the incidents of 1946 were labeled by *The New Yorker* as "the total collapse of the public opinion polls" (Lusinchi, 2018, p. 3).

Other well-known examples include the Kinsey report on the sexuality of Americans (Maslow & Sakoda, 1952) or Leary and Alpert's research on the effectiveness of LSD therapy in reducing the risk of repeat offending (Brzeziński, 2019).

Taking into account the above-mentioned reports, scientific quantitative articles by psychologists affiliated with Polish research institutions have been analyzed. The aim was to examine the details of the description of the sampling procedure and the sample itself, as these are crucial elements for making replication feasible.

Sampling Methods

In the literature, there are distinguished probabilistic and non-probabilistic sampling methods. Probabilistic methods are based on the premise that every person in the population has equal chance of participating in the study, and was selected randomly, which allows to obtain a representative sample. By contrast, non-probabilistic sampling does not guarantee that obtained sample reflects the target population in terms of the characteristics that interest the researcher. Therefore, generalizing conclusions from such studies should be approached very carefully (Shaughnessy et al., 2020).

The literature (Brzeziński, 2019; Sarstedt et al., 2018) distinguishes the following fundamental random sampling methods: simple, stratified and cluster. They require a sampling frame (a list of all elements in the population) and a random number generator.

There are many non-probabilistic sampling methods. One example of such approach involves researchers selecting respondents based on specific criteria (purposive and quota sampling) or by accidental sampling. The latter method is not recommended as no population variable is controlled, hindering the generalization of findings (Brzeziński, 2019).

Purposive sampling involves selecting individuals that, according to researchers or expert judgment, are typical representatives of the target population. This means that respondents meet pre-defined criteria for inclusion, such as profession, or place of residence. A refinement version of this method is quota sampling, where researchers determine the distributions of population characteristics of interest at the outset, and aim to reproduce the same proportions of these variables in the sample (Sarstedt et al., 2018).

Convenience sampling is frequently utilized in social research, whereby the selection of participants is influenced by the availability and accessibility of the researchers rather than by theoretical considerations. Individuals who are easy to find or recruit are selected as participants. This approach results in the under-representation of certain population segments. Such samples are biased and susceptible to many errors, and, what is more, often they are comprised of volunteers (Baker et al., 2013). Volunteers are individuals who voluntarily participate in a research by responding to published advertisements. They form a biased sample, differing from the general population because of, for instance: higher levels of education, intelligence, and need for social approval (Brzeziński, 2019).

Compared to other selection methods, samples consisting of volunteers are more prone to non-response and self-selection biases (Cheung et al., 2017).

As far as hidden or hard-to-reach populations are concerned, and where the members of the population may be unwilling to share their sensitive information with strangers, such as AIDS patients, the snowball sampling method is very useful. According to this method, researchers select a small number of population representatives and then ask them to recruit more people. This process leverages social networks of the respondents. It is worth noting that the representativeness of the first respondents is crucial to the quality of the sample (Salganik & Heckathorn, 2004).

Sampling in online surveys, e.g., by posting links on portals, is a separate matter. Respondents are individuals with access to the Internet, who visit the relevant website at a given time and decided to participate in the study. Here, the researcher does not have any control over the selection of specific individuals. Such samples may be biased and prone to error due to the use of the Internet user population and self-selection bias (Bethlehem, 2010).

Sample Size Dependency on the Power of the Statistical Test

The selection of an adequate sample size is crucial for substantiating claims about the population from which it was drawn. It is also essential for the external validity of the study. The possible methods of estimating its size depend on the sampling method and the estimated population parameter (Brzeziński, 2019). It is emphasized in the literature that the sample should not be too small, as this reduces the reliability of the study, but it should also not be too big, as it increases the risk of committing a Type I error. Thus, both estimates can be reneedered as biased, leading to under- or overestimation of population parameters (Sapra, 2022).

The sample size is also crucial for achieving adequate statistical power, which is the probability of detecting truly existing differences or correlations (Singh and Masuku, 2014). Two main estimation methods are distinguished: 1) *a priori*, where the sample size is estimated before the study based on test power, significance level, and effect size, 2) *post hoc*, which aims to determine whether the sample size used minimizes the risk of Type II error at a specified significance level and effect size. Listing all available methods, however, out of scope for this paper (see Faul et al., 2007).

A considerable part of the determinants of sample size (e.g., confidence level chosen, hypotheses, statistical tests, type of parameter estimated) is highly dependent on the researcher's decisions. Therefore, determining sample size should be a thoughtful and justified decision, not arbitrary.

Ethical, Legal, and Compensation Aspects in Research

From the sample selection to the conclusion of a study, the psychologist should be primarily guided by the freedom and dignity of each participant and

not by the desire to publish groundbreaking outcomes (Brzeziński, 2019). Participation in the study should be voluntary and anonymous, preceded by granting informed consent. In recent years, the ethical aspects of conducted research have gained importance, resulting in a requirement to obtain the approval from the relevant research ethics committee before commencing the research.

Legal conditions related to the collecting and storing data, especially sensitive information, have also changed. Another noteworthy aspect to address is the respondents compensation. Setting aside controversies associated with its application (see Brase, 2009), it is crucial to be aware of the legal and financial implications. It turns out that a better solution is to provide rewards (even smaller ones) for every participant rather than randomly recognize only some respondents with greater rewards. In the latter case, the Tax Office may deem this method as a promotional lottery, according to the definition in Article 2 of Polish The Act on Gambling Games (2009). Promotional lotteries, as per Article 7 of the mentioned regulation, “may be organized, based on a permit granted,” which, according to Article 32, “is granted by the director of the tax administration chamber.”

The issues raised regarding sample selection in psychological research have led to the formulation of the following research questions:

1. What information about the sample do Polish researchers provide, and what is missing in their descriptions?
2. What information about the research procedure and sample should be included in order to consider the descriptions accurate and comprehensive?
3. What standards should be in place for describing the research procedure and sample?

Method

The literature review adhered to the PRISMA standard (Moher et al., 2009), ensuring scientific rigor. It mandates reporting the number of identified and excluded publications, the adopted inclusion/exclusion criteria, and the article selection process.

Article Identification

Due to the research focus on empirical psychological studies by Polish authors, inspecting the state of databases was considered suboptimal. Therefore, first step was to identify journals publishing such studies. It was crucial to ensure that the topics of the studies published in chosen journals were diverse, to reach national and international articles, and to include journals of different prestige levels to reduce the effect of biased source selection. Through discussions within the team and consultations with a group of five experts (publishing

psychologists), the following journals have been selected: *Personality and Individual Differences* (PAiD; 100 points on the MEiN list (Polish The Communiqué the Minister of Education and Science, 2023), IF 4.3), *Frontiers in Psychology* (*Frontiers*; 70 points, IF 3.8), *Annals of Psychology* (*Annals*; 70 points), *Polish Psychological Forum* (PPF; 40 points).

Inclusion and Exclusion Criteria

Articles eligible for analysis had to meet the following criteria: 1) present quantitative empirical research, 2) be published as peer-reviewed original articles, 3) be published between 2021–2022, 4) have the first author affiliated with a Polish academic institution, 5) be written in Polish or English.

There were no restrictions on the number of authors, the affiliation of the other authors (in case of multi-author publications), topics, the number of studies described, the origin of the participants, or the research model.

The exclusion criterion was the absence of measurement of latent psychological constructs or no access to the full article.

Articles Selection for the Analysis

A total of 305 articles were identified from the journal search. After an initial selection based on titles and abstracts, 63 articles were excluded as they did not meet the inclusion criteria. These included mainly literature reviews/meta-analyses (9), theoretical articles (25), qualitative studies (8) and other non-empirical texts (21). From the remaining 242 articles, seven were not included due to a lack of access to the full text. After a critical appraisal of the content of the articles, a further five were excluded because they lacked psychological variables measurement. Finally, 230 articles were analyzed in the detailed review. A list of these can be found in the table in Appendix 1 in the open repository (see https://osf.io/k4qsw/?view_only=dc40b39f6a9746b9a5781a2d8c5bfa67).

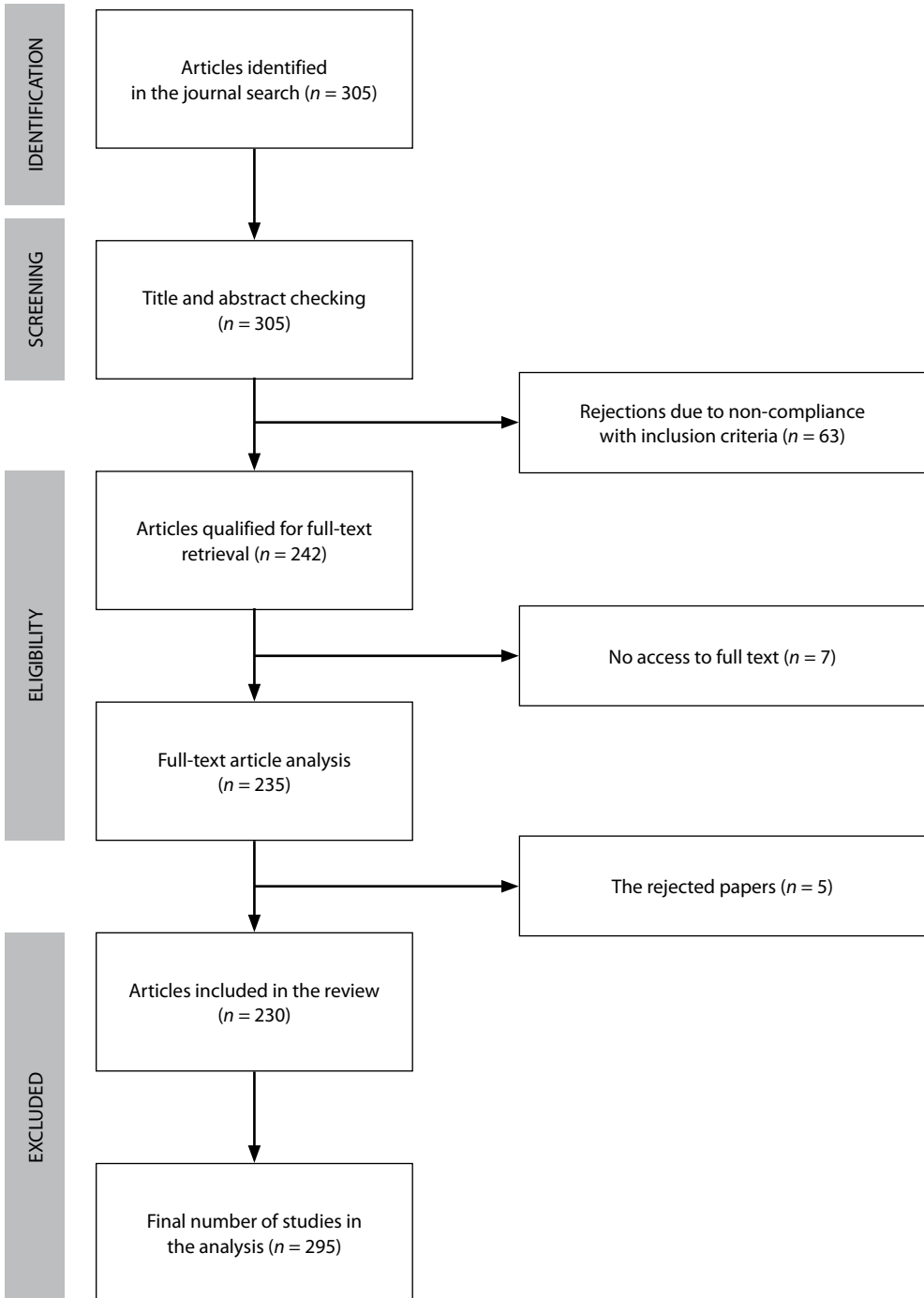
To ensure consistency in the assessment of the articles, the authors underwent special training, which included coding four texts. In addition, after working individually on the assigned part of the articles, the team consulted on any emerged uncertainties.

Modally, one article described a single study ($n = 185$), 30 papers presented sequences of two studies involving different samples, 11 articles described three studies, 3 – four studies, and one was a series of five studies. Since each of the studies included information about the participants, they were analyzed as separate entries. In other words, if an article discussed two studies, each sample description was coded separately. As a result, 295 sample descriptions were identified.

Figure 1 (p. 117) shows the steps in the process of extracting sample descriptions.

Figure 1

The article selection process



Coded Variables

The guidelines for preparing a report on empirical research do not provide clear criteria for characterizing respondents. Therefore, before starting the coding process, preferred elements of sample descriptions were identified based on this paper authors' previous experience and methodological knowledge. These elements should be included to assess the quality of the conclusions presented or to attempt to replicate the study. They cover the following information:

- 1) *a detailed description of the research procedure*, including the activities performed by participants, the number and sequence of questionnaires, and other information specific to the described study;
- 2) *the target population* to which the researchers intend to generalize conclusions;
- 3) *the sampling method* used;
- 4) *form of the study* (face-to-face or online);
- 5) *inclusion criteria* for study participants;
- 6) *exclusion criteria* for respondents;
- 7) *the percentage of people excluded* (whether disclosed or calculable);
- 8) *the sample size*, i.e., the final number of participants;
- 9) *the justification for the sample size*;
- 10) *when study was conducted*;
- 11) *who conducted the study*;
- 12) *the compensation* of the study participants;
- 13) *demographic data*, including:
 - a) *gender or sex*,
 - b) *age* – for quantitative measures presented as mean with standard deviation and for ordinal measures presented as the frequency of individuals in each age group,
 - c) *education* – for specific groups such as children and students, it was considered that it was provided implicitly,
 - d) *nationality* given explicitly, as there may be foreigners fluent in Polish among the residents, students, and those filling out questionnaires in Polish.

In addition, information regarding whether it was an experiment and whether the sample included students (in whole or in part) was also gathered. A detailed summary of the variables mentioned and their coding methods can be found in the table in Appendix 2 in the open repository (see https://osf.io/k4qsw/?view_only=dc40b39f6a9746b9a5781a2d8c5bfa67).

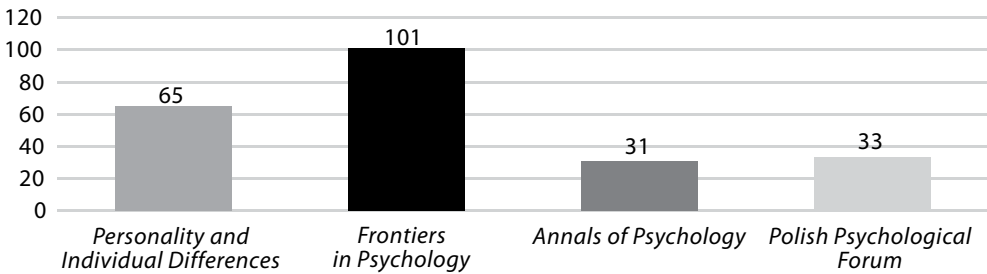
Results

General Characteristics of the Articles

Figure 2 (p. 119) illustrates the contribution of each journal in the analyzed article database. The majority of articles were published in *Frontiers* (43.9%), the least in *Annals* (13.5%). Table 1 shows the number of studies reported in the papers.

Figure 2

Number of articles included in analyses by journal (N = 230)



Of the studies described, 45 were experiments (15.3%) and 9 (3.1%) were longitudinal. The remaining studies described relationships between variables or comparisons of different populations. The topics addressed by the authors were very diverse. A notable percentage of studies focused on COVID-19 ($n = 33$; 11.2%), which was expected for the analyzed period. The second largest group of studies focused on adaptations or proposals for new questionnaires ($n = 17$; 5.8%).

Table 1

Number of studies described in articles by journal (N = 295)

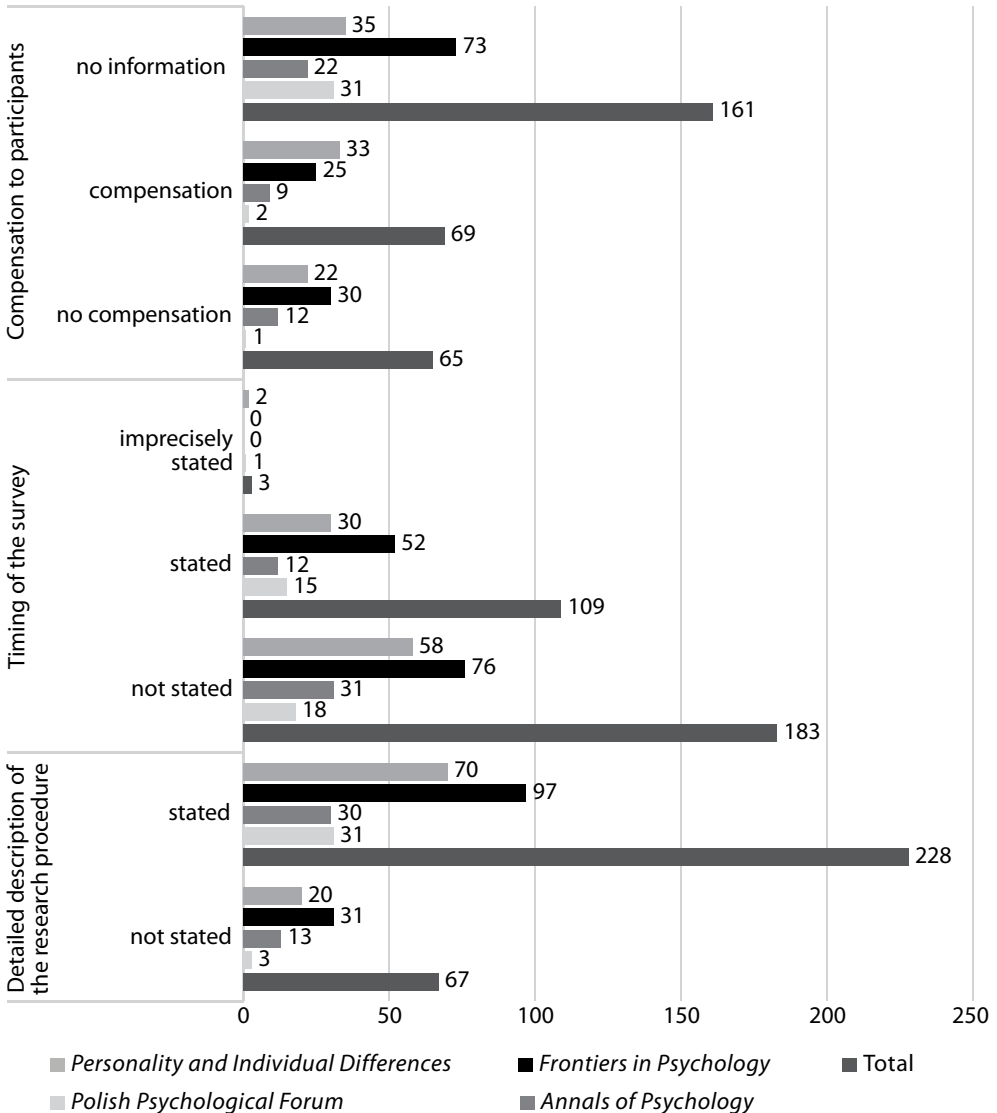
Number of studies in the article	<i>Personality and Individual Differences</i>		<i>Frontiers in Psychology</i>		<i>Annals of Psychology</i>		<i>Polish Psychological Forum</i>		Total	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
1	46	70.8	86	85.1	21	68.8	32	97.0	185	80.5
2	14	21.5	7	6.9	8	25.0	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
<i>N</i>	90	30.5	128	43.4	43	14.6	34	11.5	295	100.0

A Detailed Description of the Research Procedure, Date When Study Was Held, Who Conducted the Study, and Information about How Participants Were Compensated

Figure 3 (p. 120) shows whether the descriptions included a) detailed steps of the research process to allow later replication, b) information on when the research was conducted, and c) information if and how participants were compensated.

Figure 3

Presence of information on the detailed procedure and timing of the survey, as well as on the remuneration of respondents by journal (N = 295)



In 77.3% of studies, researchers provided sufficient information for replication, regardless of the journal ($\chi^2(3) = 5.3, p = .151$). Where information was lacking, this was mainly due to the use of purposive sampling (due to interest in a specific population, e.g., the chronically ill), with only the validation of questionnaires ($n = 16$; 23.9%) or obtaining data from an external company (research

panel; $n = 13$; 19.4%). Regarding the study format, the percentage of detailed descriptions is lower in online studies ($n = 89$; 74.2%) than in face-to-face studies ($n = 122$; 87.8%) and mixed-mode studies ($n = 12$; 92.3%).

Only 38.0% of the papers ($n = 112$) provided information when survey was conducted. It constitutes the second similarity between journals ($\chi^2(6) = 7.9$; $p = .249$). It is likely to be an effect of the publication process – the long wait for journal decisions and uncertainty about whether the article will be published in the first journal.

Information on compensation to participants was found in 45.4% of studies ($n = 134$), although how frequently it occurred varied between journals, $\chi^2(6) = 31.5$; $p < .001$. It was most frequently reported in articles in PAiD ($n = 55$; 61.1%), where participants were usually (36.7%) compensated for their participation (incl. vouchers, points, gifts). On the contrary, PFP papers were the least likely to mention rewarding participants, with only three cases (8.8%).

The occurrence of information about who conducted the study (Table 2) also differed significantly between journals, $\chi^2(6) = 31.5$; $p < .001$. The modal case (59.0%) was the absence of such information. It is a worrying signal, but the likely reason for such a high percentage is the impersonal form of scientific articles and the use of the passive voice in English. It is also noteworthy that research panels (including Ariadne, Prolificm, and Pollster) frequently appeared in PaiD publications (36.7%).

Table 2

Information on who conducted the survey by journal (N = 295)

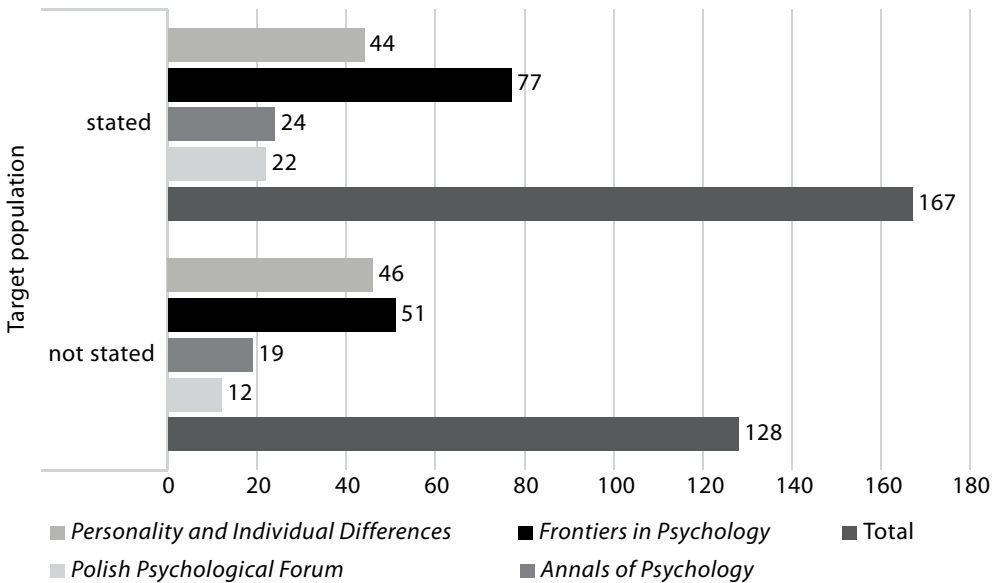
The study conductors	<i>Personality and Individual Differences</i>		<i>Frontiers in Psychology</i>		<i>Annals of Psychology</i>		<i>Polish Psychological Forum</i>		Total	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Researcher/Authors	11	12.2	25	19.5	1	2.3	1	2.9	38	12.9
Research companies	33	36.7	3	2.3	4	9.3	0	0	40	13.6
Students	2	2.2	1	0.8	1	2.3	1	2.9	5	1.7
Institutional staff (e.g. psychologists, nurses, prison staff)	3	3.3	7	5.5	0	0	0	0	10	3.4
Professionals	2	2.2	4	3.1	0	0	4	11.8	10	3.4
Research assistants	5	5.6	10	7.8	2	4.7	1	2.9	18	6.1
Missing information	34	37.8	78	60.9	35	81.4	27	79.4	174	59.0

The Target Population, Sampling Method, and the Form of Study

In 167 studies (56.6%), authors conducted research on a specific population (e.g., adaptation of a questionnaire; see Nickel, 2021) or generalized findings to a general population (Figure 4). There were no statistically significant differences between journals in this regard, $\chi^2(3) = 3.8$; $p = .289$.

Figure 4

Information on target population by journal ($N = 295$)



The information about the sampling method remained at a similar level of about 70% across all journals, $\chi^2(3) = 0.3$; $p = .959$. However, the sampling method that offered the highest chance of obtaining a representative sample appeared in only five cases (1.7%) described in *Frontiers*. The descriptions were dominated by purposive sampling (24.4%), convenience sampling (16.6%), and research panels (13.6%). It is worth noting that in 27.8% of the studies, the sampling method was not mentioned. The types of sampling methods across journals are shown in Table 3 (p. 123).

The vast majority of studies (92.2%), regardless of the journal, reported information on the form of the study: face-to-face, online, and mixed (Table 4, p. 123), $\chi^2(3) = 2.6$; $p = .453$. Stationary studies, accounting for more than half of the cases, were the modal format in all journals except *PaiD*. In the latter, online studies predominated (57.8%), possibly due to the pandemic period.

Table 3

Sampling methods used in studies by journal (N = 295)

Sampling method	<i>Personality and Individual Differences</i>		<i>Frontiers in Psychology</i>		<i>Annals of Psychology</i>		<i>Polish Psychological Forum</i>		Total	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Random sampling	0	0	5	3.9	0	0	0	0	5	1.7
Research Panel	33	36.7	4	3.1	3	7.0	0	0	40	13.6
Purposive sampling	9	10.0	41	32.0	11	25.6	11	32.4	72	24.4
Quota sampling	1	1.1	0	0	0	0	1	2.9	2	0.7
Volunteers	3	3.3	4	3.1	2	4.7	4	11.8	13	4.4
Snowball sampling	3	3.3	6	4.7	0	0	5	14.7	14	4.7
Convenience sampling	13	14.4	25	19.5	8	18.6	3	8.8	49	16.6
Consecutive sampling	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
Accidental sampling	2	2.2	2	1.6	6	14.0	0	0	10	3.4
Missing information	26	28.9	34	26.6	13	30.2	9	26.5	82	27.8

Table 4

Form of study by journal (N = 295)

Form of Study	<i>Personality and Individual Differences</i>		<i>Frontiers in Psychology</i>		<i>Annals of Psychology</i>		<i>Polish Psychological Forum</i>		Total	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Face-to-face	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
Mixed	2	2.2	9	7.0	1	2.3	1	2.9	13	4.4
Missing information	9	10.0	7	5.5	5	11.6	2	5.9	23	7.8

Sample Size and Its Justification, Inclusion and Exclusion Criteria, and Percentage of Exclusions

The sample sizes varied widely, ranging from 10 to 7108 participants. Table 5 presents basic descriptive statistics of the distribution of this variable for the entire dataset and separately for each journal. The distributions obtained were strongly right-skewed, indicating a prevalence of samples with sizes smaller than the mean. Consequently, and taking into account the considerable differences in the frequency of studies in these journals, 1) the median seems to be a better measure to characterize these distributions, and 2) the use of non-parametric tests to compare sample sizes between journals would be biased with a higher risk of error due to the different degrees of distortion of the distribution.

Table 5

Descriptive statistics of sample size distributions by journal (N = 295)

	<i>Personality and Individual Differences</i>	<i>Frontiers in Psychology</i>	<i>Annals of Psychology</i>	<i>Polish Psychological Forum</i>	Total
M	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

To present the distribution of the analyzed variable more clearly and to explore possible sources of differences in the size of the samples compared, the quantitative variable was recoded into an ordinal (Table 6, p. 125). It appears that the type of journal significantly differentiated sample size categories, $\chi^2(12) = 36.3; p < .001$. In the case of *PaiD*, the modal range was from 251 to 500 participants (31.1%), while in *Annals* most studies were conducted with samples ranging from 131 to 250 participants (37.2%). In *Frontiers* and *PFP*, most studies were conducted with small samples of up to 130 people.

In addition, an attempt was made to determine what other factors might differentiate the sample sizes. It was observed that a detailed study description [$\chi^2(4) = 9.8; p < .05$], the individual who conducted it [$\chi^2(4) = 13.0; p < .05$], and the time of the study [$\chi^2(8) = 26.8; p < .001$] were more common in larger samples (between 251 and 1000 individuals) than in smaller ones (up to 250 individuals). Detailed results are available in Appendix 3 in the open repository (link: https://osf.io/k4qsw/?view_only=dc40b39f6a9746b9a5781a2d8c5bfa67).

Table 6

Sample size categories by journal (N = 295)

Sample size	<i>Person- ality and Individual Differences</i>		<i>Frontiers in Psychology</i>		<i>Annals of Psychology</i>		<i>Polish Psy- chological Forum</i>		Total	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Small samples (up to 130 respondents)	11	12.2	46	35.9	14	32.6	15	44.1	86	29.2
Samples of 131 to 250 participants	19	21.1	18	14.1	16	37.2	5	14.7	58	19.7
Samples of 251 to 500 participants	28	31.1	34	26.6	7	16.3	11	32.4	80	27.1
Samples of 501 to 1000 participants	22	24.4	21	16.4	4	9.3	3	8.8	50	16.9
Samples over 1001 participants	10	11.1	9	7.0	2	4.7	0	0	21	7.1

Inclusion and exclusion criteria for participants were reported in about one-third of the studies. Inclusion criteria were most frequently reported in articles published in *Frontiers* (46.1%) and least frequently in *PaiD* (20.0%), $\chi^2(3) = 15.9$; $p < .001$. On the contrary, exclusion criteria were most common in *Annals* (39.5%) and least frequent in *PFP* (11.8%), $\chi^2(3) = 10.1$; $p < .05$. Information on the percentage or number of people excluded appeared in approximately 29.5% of studies, regardless of the journal, $\chi^2(6) = 6.5$; $p = .367$.

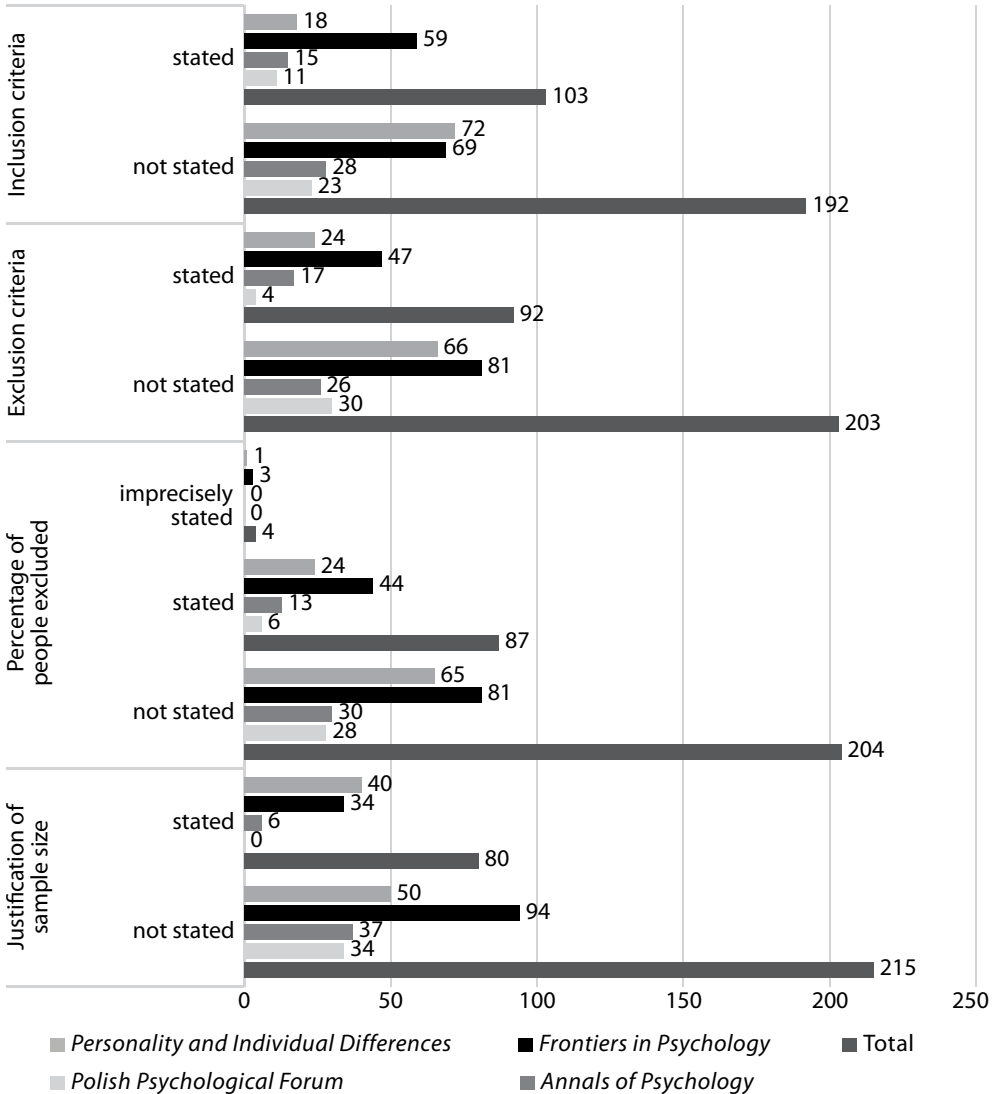
Sample size justification was present in 27.1% of studies (Figure 5, p. 126). Foreign journals outperformed Polish journals in this regard. In *PFP*, authors never mentioned this aspect, and in *Annals* it only appeared six times (14.0%). This information was most frequently found in *Frontiers* (44.4%).

**Demographic Information about the Respondents:
Sex/Gender, Age, Education, Nationality**

When describing participants, the most basic information was considered: sex or gender, age, educational level, and nationality (see Figure 6, p. 127). Although information on sex or gender and age appeared in over 90% of the descriptions, there was considerable variation between journals. For sex or gender, the main reason for variation was *PaiD*, where this information was always provided, $\chi^2(3) = 7.9$; $p < .05$. Similarly, in the reporting of age, the most frequent omission of this element occurred in descriptions in *PFP* (17.6%), $\chi^2(3) = 13.7$; $p < .01$.

Figure 5

Information on inclusion and exclusion criteria, percentage of people excluded and justification of sample size by journal ($N = 295$)

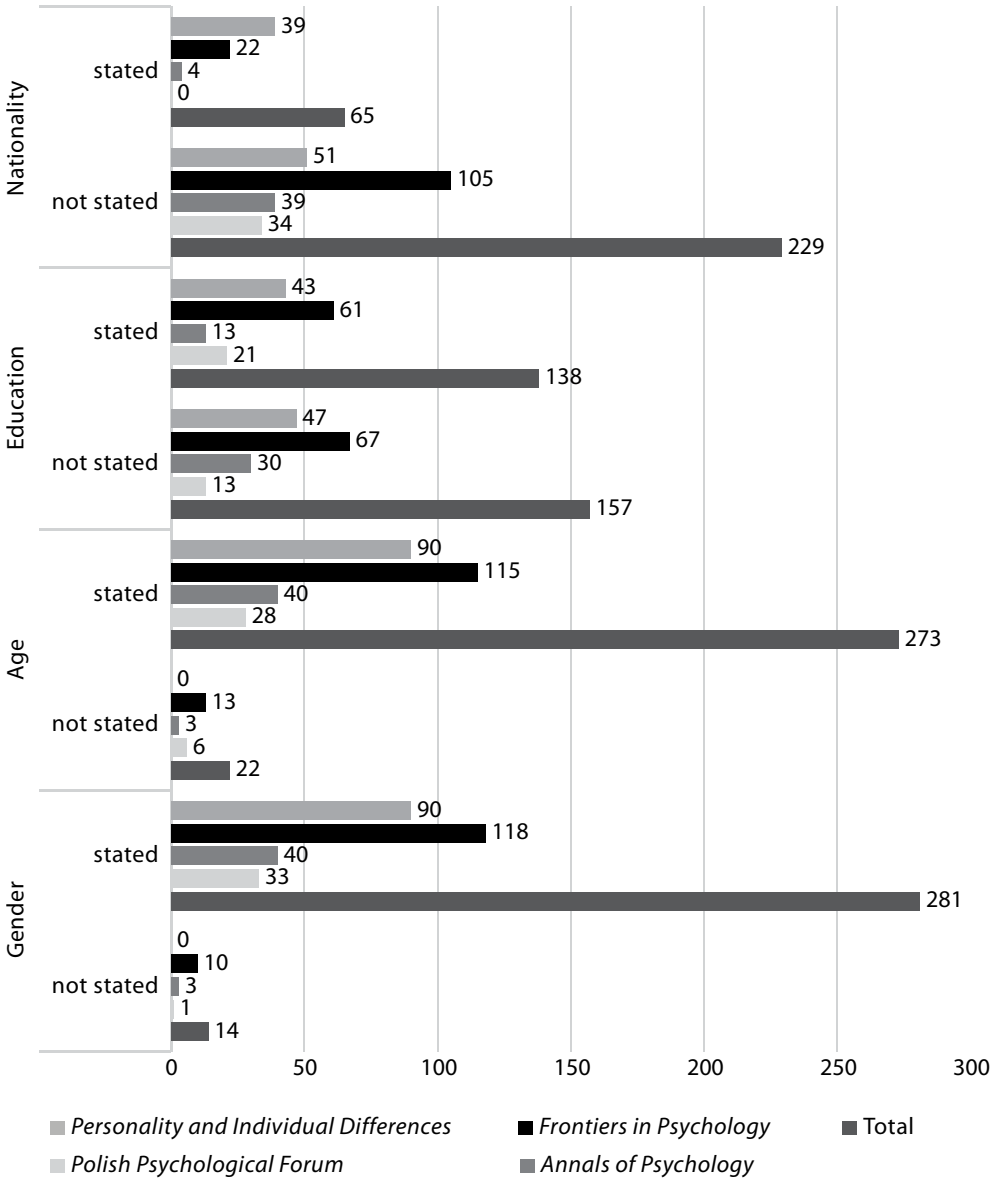


Educational level was most common in PFP articles (61.8%) and least common in *Frontiers* (30.2%), $\chi^2(3) = 7.9$; $p < .05$. Such information appeared in 46.8% of the papers when a liberal criterion was applied (i.e., considering these details present in articles describing infants, preschoolers, and students). The least frequently reported characteristic, regardless of the journal,

was the explicit indication of nationality (22.1%). This information was mainly mentioned in PaiD publications (43.3%), whereas it was not mentioned at all in PFP, $\chi^2(3) = 39.0; p < .001$. It should be noted that foreigners can also use the Polish language.

Figure 6

Information on demographics by journal (N = 295)



Discussion

This review aimed to identify elements present in descriptions of the research process, particularly those regarding the sample itself, and to formulate recommendations based on this analysis.

Among the elements analyzed, the most common ones in the descriptions were: a description of the procedure enabling study replication, the sampling method used and the sample size, the form of the study, and socio-demographic characteristics such as sex or gender and age. It is a modest amount of information that requires comment.

The analyses showed that 77.3% of the study descriptions included information about the procedure. When describing face-to-face studies, it should be specified where and under what circumstances they were conducted, e.g., during work/classes, during a break, or individually/in groups. In addition, it is crucial to list the questionnaires used as well as the order in which they were presented (order effect; Bayat et al., 2023). It is also important to indicate how socio-demographic data was collected, for instance, whether via a self-designed survey or otherwise. This information was particularly rare in online surveys. The increasing popularity of online studies, highlights the need to identify the information that should be provided.

Nevertheless, the research procedure should include a detailed description of the participants' tasks and a summary of what constitutes the subject of the researcher is interested in (population or phenomenon), when the data were collected, for how long, and by whom. Unfortunately, this information was particularly scarce, but it is crucial for the methodological assessment of the quality of the conclusions.

Furthermore, the reader must exhibit patience while seeking information about the study procedure and participants. It turns out that the sections of the articles dedicated to these issues (e.g., Sample, Procedure) are not the only places where this information is enclosed. Often, these details were scattered in other sections, such as Instruments or Limitations. In some cases, specific sample-related data were only included in the Abstract. This makes it difficult to understand the data collection process, which may make it difficult to assess the internal validity and replicability of the study.

Although only 27.8% of studies failed to identify the sampling method, unfortunately, this was not due to the accuracy of the descriptions. Researchers did not always explicitly state the sampling method, but it could be inferred from the descriptions. There were also inconsistencies between the declared method and the subsequent procedure. The situation became even more complicated when several different approaches to reaching participants were described, often reported as snowball sampling. It should be emphasized that this technique requires the respondents to recruit the majority of participants. Simply asking individuals to send information to more people is not sufficient.

In more than half of the studies, the authors referred to a specific population or generalized their findings to it. Such an approach is justified if a given sample is representative of the population, and only a random sampling method

would allow this. Unfortunately, this was the case in only five studies. Among the non-probabilistic sampling methods that come closest to random sampling (Brzeziński, 2019), quota sampling was used only twice.

Descriptions of online surveys often only mentioned placing the link on websites or Facebook without specifying the exact location. The difference between accidental sampling, where links are placed everywhere, and purposive sampling, where *specific groups of people* with particular interests/professions are searched, affects the ability to replicate studies and assess the external validity of the conclusions presented.

There is a growing concern about the impact of the sample size on the ability to detect a specific effect size and control for type I and type II errors (Adam, 2020; Lakens, 2022). This information, combined with explicitly stated inclusion and exclusion criteria and the percentage/number of people excluded, is lacking, especially in studies published in national journals.

The last criterion was socio-demographic data. As expected, almost all studies reported information on sex or gender and age. Information on the educational level appears less frequent. Information on nationality was hardly present, even in international journals. It is worth noting that both samples of students and samples of Polish speakers (without stating that it is their mother tongue) may not consist exclusively of native Poles.

Henrich et al. (2010) criticized psychologists for conducting studies mainly on samples from the W.E.I.R.D. society, i.e., Western, Educated, Industrialised, Rich, and Democratic. Meanwhile, Kaźmierczak et al. (2023) expressed concern that only individuals with affective and personality disorders participate in studies. It is recommended to go beyond these patterns, carefully selecting respondents, and then accurately reporting when, by whom, how, under what circumstances the study was conducted, and who conducted the research.

Recommendations for Researchers

Based on the experience gained in the course of preparing this review, the following recommendations for publishing researchers have been formulated:

- 1) when reporting the research procedure and sample, consider what information is essential for proper replication and describe it comprehensively;
- 2) the absolute minimum information in the sample description should include a replicable research procedure, the sampling method, sample size, form of the study, and sociodemographic characteristics such as sex or gender and age. However, other relevant features highlighted in this review should be added to these details;
- 3) the elements discussed should be placed in the relevant section of the article, such as the *Sample* or *Research Method* section;
- 4) as far as online studies are concerned, it is essential to determine a methodological information requirements. We suggest the following details:
 - a) the platform used for data collection (e.g., Google Forms, Qualtrics, LimeSurvey),

- b) the method of distributing the survey link (e.g., email, social media – specify which) and who is responsible for distributing it,
- c) the data collection period,
- d) the research procedure, including the questionnaires used and their sequence,
- e) validity check questions,
- f) inclusion criteria (at least in the form of a verification question regarding membership of the researcher's target population) and exclusion criteria, with the percentage of excluded participants.

Limitations

There are several limitations to presented literature review. First, only articles from arbitrarily selected journals with various levels of prestige were examined. Second, the research from articles that covered a series of studies was analyzed separately. This may have resulted in higher proportion of papers with less space for study descriptions. Third, both experimental and individual differences studies were included in the analysis. The latter were mainly conducted in interpersonal but also in intrapersonal contexts. Therefore, the conclusions presented in this article are of minor, but not insignificant importance (cf. Thompson & Campbell, 2004; Wojciszke, 2004). Thus, we suggest conducting separate analysis focusing exclusively on studies conducted in intrapersonal settings. Lastly, unclear descriptions of sampling led to some of it being categorized as missing information, which may have resulted in its overrepresentation ($n = 82$) and underrepresentation, such as random sampling ($n = 10$). As the presented review was exploratory, it would be valuable to repeat it with a different set of journals and to examine international trends.

References

- 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 [The Practice of Social Research]*. 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), Article 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/S1930297500001248>
- Brzeźniński, J. (2019). *Metodologia badań psychologicznych. Wydanie nowe [Methodology of the Psychological Research. New Edition]*. 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, Article 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., & Ściagała, D. (2023). Self-selection biases in psychological studies: Personality and affective disorders are prevalent among participants. *PLOS ONE*, 18(3), Article 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 [Polish The Communiqué the Minister of Education and Science on the list of scientific journals and peer-reviewed materials from international conferences of 17 July 2023]. (2023). <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), Article 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), Article e1000097. <https://doi.org/10.1371/journal.pmed.67>
- 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/psychoc.2004.277>
- Nikel, Ł. (2021). Polska adaptacja Skali do badania motywacji uczniów w szkole podstawowej [Polish Adaptation of the Elementary School Motivation Scale]. *Polskie Forum Psychologiczne [Polish Psychological Forum]*, 26(1), 97–112. <https://doi.org/10.34767/PFP.2021.01.06>
- 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? In S. Nundy, A. Kakar, & Z. A. Bhutta (Eds.), *How to Practice Academic Medicine and Publish from Developing Countries? A Practical Guide* (pp. 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 [Research methods in psychology]*. 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 [Polish The Act on Gambling Games of 19 November 2009, Journal of Laws of 2009, No. 201, item 1540]. (2019). <https://isap.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=W-DU20092011540>
- Wojciszke, B. (2004). Systematycznie modyfikowane autoreplikacje: logika programu badań empirycznych w psychologii [Systematically Modified Auto-Replications: the logic of the empirical research program in psychology]. In J. M. Brzeziński (Ed.), *Metodologia badań psychologicznych. Wybór tekstów [Methodology of the Psychological Research. Selected papers]* (pp. 44–68). Wydawnictwo Zysk i Ska.
- Zimbardo, P. G., & Gerrig, R. J. (2012). *Psychologia i życie [Psychology and Life]*. Wydawnictwo Naukowe PWN.

Acknowledgments

The authors would like to express their thanks to the all members of the team who provided support in the preparation of this article. Special thanks go to Angelika Rusin and Julia Szymańska.