

Sampling methods and sample populations in quantitative mass communication research studies: A 15-year census of six journals

Manuscript accepted for publication in:

Communication Research Reports

By:

Joseph Erba, Brock Ternes, Piotr Bobkowski, Tara Logan, & Yuchen Liu

Abstract: This study evaluated the sampling methods and sample populations used in all U.S.-based survey and experimental mass communication studies published between 2000 and 2014 in six major journals ($N = 1,173$). Most studies used nonprobability samples and more than half used student samples. Experiments used more nonprobability and student samples than surveys. Funded studies used more probability and non-student samples than non-funded studies. Implications of results pertaining to population validity and interpretations of findings for mass communication research are discussed.

Keywords: Sampling; sample population; survey; experiment; mass communication research; research funding

Joseph Erba (PhD, University of North Carolina at Chapel Hill, 2013) is an Assistant Professor in the William Allen White School of Journalism and Mass Communications at the University of Kansas. Brock Ternes (PhD, University of Kansas, 2016) is an Instructor in the Sociology/Anthropology Department at SUNY–Cortland. Piotr Bobkowski (PhD, University of North Carolina at Chapel Hill, 2010) is an Associate Professor in the William Allen White School of Journalism and Mass Communications at the University of Kansas. Tara Logan (MS, University of Kansas, 2016) is an Adjunct Professor in the William Allen White School of Journalism and Mass Communications at the University of Kansas. Yuchen Liu (MS, University of Kansas, 2016) is a Doctoral Student in the William Allen White School of Journalism and Mass Communications at the University of Kansas. Correspondence: Joseph Erba: erba@ku.edu

How researchers sample (i.e., probability or nonprobability methods) and whom they sample (i.e., demographic, social, or professional identities of participants) are two critical factors that can affect the overall external validity of a study's results. Sampling methods and sample populations together constitute population validity, a component of external validity that pertains to the prudence of drawing conclusions about a population based on results from a studied sample (Lowry, 1979). Like other disciplines, mass communication studies often rely on convenience samples of undergraduate college students (Meltzer, Naab, & Daschmann, 2012; Rotfeld, 2003). Findings from such research have influenced our understanding of the role media play in the everyday lives of audiences. Given the ubiquity of media in society and the diversity of U.S. society, it is particularly important for mass communication researchers to reflect on the sampling methods and sample populations used in their studies, as results from these studies have been used to build theories and inform practice.

Thus far, there has not been a systematic appraisal of the use of nonprobability and student samples in mass communication research. To address this limitation, the present study contributes an assessment of sampling methods and sample populations in quantitative mass communication studies conducted in the United States and published over 15 years (2000-2014) in six leading peer-reviewed communication journals. The analysis considers how these sample characteristics relate to research methods (survey and experiment) and to the presence of research funding, and assesses the population validity of quantitative mass communication research results. The following research questions guided the analysis:

RQ1: How frequently are probability and nonprobability sampling methods used in survey and experimental mass communication research studies?

RQ2: How frequently are college student and non-student samples used in survey and experimental mass communication research studies?

RQ3: What is the relationship between research funding and the use of nonprobability and student samples in survey and experimental mass communication research studies?

Methods

The sampling frame consisted of quantitative mass communication research articles with U.S. samples, published between 2000 and 2014 in six peer-reviewed communication journals (*Communication Research* [CR], *Journal of Broadcasting & Electronic Media* [JOBEM], *Journal of Communication* [JC], *Journal of Computer-Mediated Communication* [JCMC], *Journalism and Mass Communication Quarterly* [JMCQ], *Mass Communication and Society* [MCS]). These journals were selected based on their impact factor scores and affiliations with professional organizations for mass communication research. Prior analyses of mass communication research also sampled from these journals (e.g., Potter & Riddle, 2007).

To be included in the analysis, an article had to: (1) contain a mass communication study; (2) include at least one participant; (3) present only quantitative data; (4) present data collected only in the United States (our codebook is accessible via the Open Science Framework [OSF]: <https://osf.io/kfyaz/>). Within an article, each study that used a unique sample was coded individually. Of the 2,797 articles published in the six journals, 1,173 studies met the inclusion criteria.

Four variables were coded in each study: method used (survey, experiment), sampling method (probability, nonprobability), sample population (college students, non-students), research funding (present, absent) (see codebook for variable definitions).

Five coders underwent training on coding definitions and instrument. Intercoder reliability was assessed with a randomly selected subsample of 180 articles. Krippendorff's *alpha* coefficients and overall agreement were calculated. There was satisfactory agreement on study inclusion ($\alpha = .84$, 92% agreement). Inclusion disagreements were reconciled before variable coding proceeded. The remaining reliability measures also were satisfactory: research method, $\alpha = .81$, 93% agreement; sampling method, $\alpha = .78$, 90% agreement; sample population, $\alpha = .79$, 94% agreement; research funding, $\alpha = .92$, 98% agreement. Each coder coded approximately a fifth of the remaining studies. A further quality check was performed by two coders who both coded every 10th article. Cohen's *kappa* coefficients were calculated between the coders' coding and the original coding, which were also satisfactory: research method, $\kappa = .95$, 97.5% agreement; sampling method, $\kappa = .83$, 91.5% agreement; sample population, $\kappa = .97$, 98.3% agreement; research funding, $\kappa = .81$, 90.7% agreement.

Results

Studies were distributed in the six journals as follows: *JC*, 21.4%; *CR*, 17.8%; *JMCQ*, 17.6%; *JOBEM*, 17.2%; *MCS*, 13.8%; *JCMC*, 12.2%. Results presented here are aggregated across the six journals. Difference-of-proportions tests were conducted to assess statistical significance when appropriate. Data used for this project can be viewed in OSF (<https://osf.io/tcqk5/>).

Nonprobability sampling was used much more often (82.6%, $n = 969$) than probability sampling (17.4%, $n = 204$). Slightly more than half of the studies (51.1%, $n = 599$) used college student samples; 48.9% ($n = 574$) used non-college student samples. Surveys were more frequent (55.1%, $n = 646$) than experiments (44.9%, $n = 527$). A little less than a fifth of the studies received funding (19.8%, $n = 232$).

RQ1 asked about the association between sampling methods and research methods. Less than one-third of surveys used probability sampling (29.6%, $n = 191$), compared to only 2.5% of experiments ($n = 13$) ($z = 12.18, p < .001$) (see Figure 1). It is important to note, however, that the majority of surveys did *not* use probability sampling (70.4%, $n = 455$).

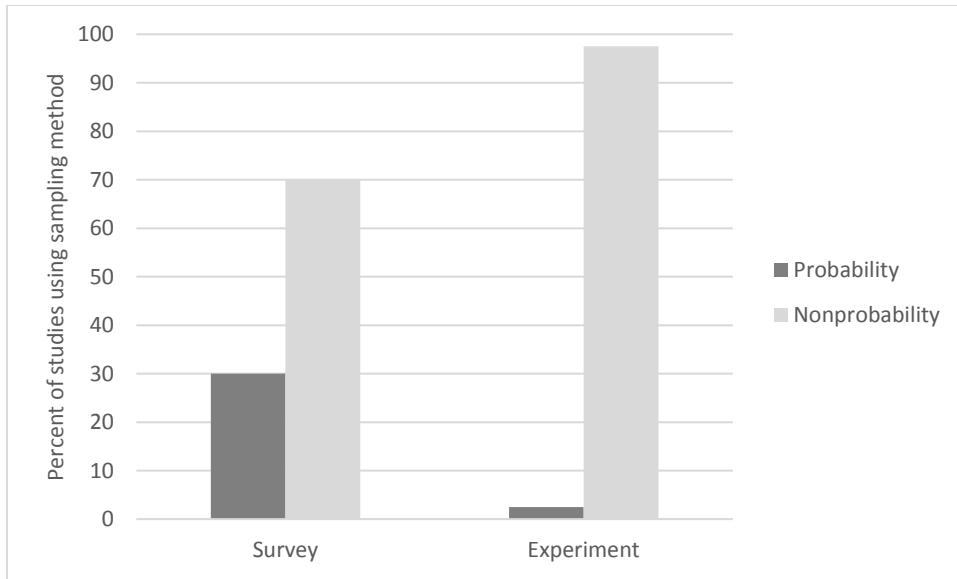


Figure 1: Sampling methods used in survey and experimental studies

RQ2 asked about the association between population samples and research methods. Most experiments (78.7%) used college student samples ($n = 415$), while 28.5% of surveys used college student samples ($n = 184$) ($z = -17.13, p < .001$) (see Figure 2).

With respect to study funding (RQ3), 29.7% of funded studies used probability sampling methods ($n = 69$), compared to 14.3% of studies that did not receive any funding ($n = 135$) ($z = -5.54, p < .001$). This pattern was repeated in studies using each research method. A greater proportion of funded surveys used probability sampling (43.8%, $n = 60$), than unfunded surveys

(25.7%, $n = 131$) ($z = -4.11, p < .001$). Similarly, 9.5% of funded experiments used probability sampling ($n = 9$), compared to 0.9% ($n = 4$) of unfunded experiments ($z = -4.86, p < .001$).

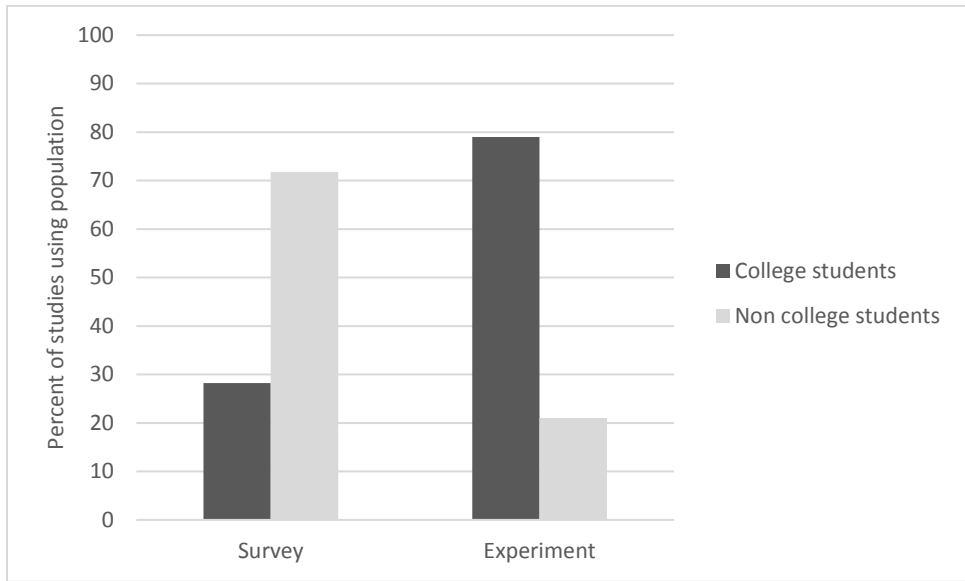


Figure 2: Sample population used in survey and experimental studies

Examining the relationship between funding and sample populations, 72.8% of funded studies used non-college student samples ($n = 169$), compared to 43.0% of unfunded studies ($n = 405$) ($z = -8.13, p < .001$). This pattern held for each research method. For funded surveys, 89.8% used non-college student samples ($n = 123$), compared to 66.6% of unfunded surveys ($n = 339$) ($z = -5.33, p < .001$). Similarly, 48.4% of funded experiments used non-college student samples ($n = 46$), compared to 15.3% of unfunded experiments ($n = 66$) ($z = -7.15, p < .001$).

Conclusion

Results reveal that these mass communication studies have primarily used nonprobability sampling methods, with more than eight of every ten studies relying on such methods, and, to a lesser extent, student samples, which were used in approximately half of the studies. Survey

studies were more likely to use probability sampling methods and non-student samples than experimental studies. These results may encourage researchers to reflect on the relationship between their sampling practices and populations, and external validity of their findings. Indeed, research has shown significant differences between results from college student and non-college student samples, both in mass communication research and in social science more generally (Basil, Brown, & Bocarnea, 2002; Hooghe, Stolle, Mahéo, & Vissers, 2010; Peterson, 2001).

Moreover, data obtained from nonprobability samples may not be representative of the population of interest (Macias, Springston, Weaver, & Neustifter, 2008; Moy & Murphy, 2016). Ha et al. (2015), who examined mass communication surveys published between 2001 and 2010, and who likewise found that most used nonprobability sampling, concluded that “the high use of non-probability samples especially among the college students means the results not only cannot be generalized to the general population, but also not to the student population” (p. 54). While not all researchers may aim to generalize their results, as some have argued that the main purpose of experiments is to explore causal relationship and not to generalize to a larger population (Courtright, 1996; Lang, 1996), it nevertheless may be prudent for researchers to use caution when interpreting findings and drawing conclusions based on their sampling methods and sample populations.

The results also show that funding is related to increased use of probability and non-student research, as these research elements tend to be more costly than nonprobability and student samples. This relationship also may be due to funding agencies often requiring applicants to maximize the external validity of proposed research, which is one of the most important factors in evaluating study proposals (Dearing & Larson, 2002).

Our findings only represent U.S.-specific studies published in the six journals we selected during a particular timeframe, and may not be representative of all published quantitative mass communication research articles. In addition, our analysis did not investigate variables related to participants' characteristics (e.g., race, income) that could also be taken into account to evaluate a study's external validity.

While it is costly to assemble samples that are representative of a general population, the American National Election Studies (ANES)—which is a collaboration between leading public opinion research centers—may provide a valuable model for researchers to emulate. Leading mass communication research programs could pool their resources and seek additional public and private support to establish a representative, longitudinal media study that becomes the field's gold standard while also motivating complementary research.

References

- Basil, M. D., Brown, W. J., & Bocarnea, M. C. (2002). Differences in univariate values versus multivariate relationship: Findings from a study of Diana, Princess of Wales. *Human Communication Research, 28*(4), 501-514.
- Courtright (1996). Rationally thinking about nonprobability. *Journal of Broadcasting & Electronic Media, 40*(3), 414-421.
- Dearing, J., & Larson, S. (2002). Private foundation funding of applied communication research. *Journal of Applied Communication Research, 30*(4), 358-368.
- Ha, L., Hu, X., Fang, L., Henize, S., Park, S., Stana, A., & Zhang, X. (2015). Use of survey research in top mass communication journals 2001–2010 and the total survey error Paradigm. *Review of Communication, 15*(1), 39-59.

- Hooghe, M., Stolle, D., Mahéo, V. A., & Vissers, S. (2010). Why can't a student be more like an average person?: Sampling and effects in social science field and laboratory experiments. *The Annals of the American Academy of Political and Social Science*, 628(1), 85-96.
- Lang, A. (1996). The logic of using inferential statistics with experimental data from nonprobability samples: Inspired by Cooper, Dupagne, Potter, and Sparks. *Journal of Broadcasting & Electronic Media*, 40(3), 422-430.
- Lowry, D. T. (1979). Population validity of communication research: Sampling the samples. *Journalism Quarterly*, 56(1), 62-68, 76.
- Macias, W., Springston, J. K., Weaver, R. A. L., & Neustifter, B. (2008). A 13-year content analysis of survey methodology in communication related journals. *Journal of Current Issues & Research in Advertising*, 30(1), 79-94.
- Meltzer, C. E., Naab, T., & Daschmann, G. (2012). All student samples differ: On participant selection in communication science. *Communication Methods and Measures*, 6(4), 251-262.
- Moy, P., & Murphy, J. (2016). Problems and Prospects in Survey Research. *Journalism & Mass Communication Quarterly*, 93(1), 16-37.
- Peterson, R. A. (2001). On the use of college students in social science research: Insights from a second-order meta-analysis. *Journal of Consumer Research*, 28(3), 450-461.
- Potter, W. J., & Riddle, K. (2007). A content analysis of the media effects literature. *Journalism & Mass Communication Quarterly*, 89(1), 90-104.
- Rotfeld H. J. (2003). Convenient abusive research. *The Journal of Consumer Affairs*, 37(1), 191-194.