

- Montgomery, K. C. (2007). *Generation digital: Politics, commerce, and childhood in the age of the Internet*. Cambridge, MA: MIT Press.
- Rogelberg, S. G., Fisher, G. G., Maynard, D. C., Hakel, M. D., & Horvath, M. (2001). Attitudes toward surveys: Development of a measure and its relationship to respondent behavior. *Organizational Research Methods, 4*, 3–25.
- Rogelberg, S. G., Luong, A., Sederburg, M. E., & Cristol, D. S. (2000). Employee attitude surveys: Examining the attitudes of noncompliant employees. *Journal of Applied Psychology, 85*, 284–293.
- Schmitt, N., & Stults, D. M. (1985). Factors defined by negatively keyed items: The result of careless respondents? *Applied Psychological Measurement, 9*, 367–373.
- Sliter, K. A., & Zickar, M. J. (2014). An IRT examination of the psychometric functioning of negatively worded personality items. *Educational and Psychological Measurement, 74*, 214–226.
- Thompson, B. (1994, January). *It is incorrect to say “the test is reliable”*: Bad language habits can contribute to incorrect or meaningless research conclusions. Paper presented at the Annual Meeting of the Southwest Educational Research Association, San Antonio, TX. ERIC Number: ED367707.
- Woods, C. M. (2006). Careless responding to reverse-worded items: Implications for confirmatory factor analysis. *Journal of Psychopathology and Behavioral Assessment, 28*, 186–191.
- Wright, K. B. (2005). Researching Internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. *Journal of Computer-Mediated Communication, 10*(3), 00. doi:[10.1111/j.1083-6101.2005.tb00259.x](https://doi.org/10.1111/j.1083-6101.2005.tb00259.x)

Difference in Response Effort Across Sample Types: Perception or Reality?

Shan Ran, Mengqiao Liu, Lisa A. Marchiondo, and Jason L. Huang
Wayne State University

Landers and Behrend (2015) question organizational researchers' stubborn reliance on sample source to infer the validity of research findings, and they challenge the arbitrary distinctions researchers often make between sample sources. Unconditional favoritism toward particular sampling strategies (e.g., organizational samples) can restrict choices in methodology, which in turn may limit opportunities to answer certain research questions. Landers and Behrend (2015) contend that no sampling strategy is inherently superior (or inferior), and therefore, all types of samples warrant careful consid-

Shan Ran, Mengqiao Liu, Lisa A. Marchiondo, and Jason L. Huang, Department of Psychology, Wayne State University.

Correspondence concerning this article should be addressed to Shan Ran, 5057 Woodward Avenue, Department of Psychology, Wayne State University, Detroit, MI 48202. E-mail: shan.ran@wayne.edu

eration before any validity-related conclusions can be made. Despite sound arguments, the focal article focuses its consideration on external validity and deemphasizes the potential influence of sample source on internal validity. Agreeing with the position that no samples are the “gold standard” in organizational research and practice, we focus on insufficient effort responding (IER; Huang, Curran, Keeney, Poposki, & DeShon, 2012) as a threat to internal validity across sample sources.

IER, which ranges from intentional random endorsement of response options throughout a survey to occasional lapses in attention to several items, can affect the internal validity of study conclusions (Huang, Liu, & Bowling, 2015). IER can result in random measurement error that attenuates correlations between variables (McGrath, Mitchell, Kim, & Hough, 2010), thus inflating Type II error rate. IER can also produce systematic bias that inflates correlations between variables, thus increasing Type I error rate in particular instances (Huang, Liu, & Bowling, 2015). Given IER’s threat to internal validity, greater attention should be paid to whether sample sources differ in their presence and detection of IER. We begin examining these issues by (a) providing a cursory examination of IER detection in the organizational literature and (b) investigating common practices in IER detection between convenience sampling strategies in recently published journal articles. To promote evidence-based guidelines for sample selection and evaluation in organizational research and practice, we also issue a call and provide recommendations for more IER research.

A Cursory Examination of IER Detection

Organizational researchers often inquire about the extent to which respondents provide quality data, especially when they are recruited from online sources (Barger, Behrend, Sharek, & Sinar, 2011). Although some evidence suggests no differences in data quality between crowdsourcing and student samples (Behrend, Sharek, Meade, & Wiebe, 2011), no studies have directly compared organizational samples with other sources on IER. Perhaps researchers perceive organizational samples as being of higher quality and free of IER. Consistent with this speculation, Liu, Bowling, Huang, and Kent’s (2013) survey of Society for Industrial and Organizational Psychology members showed that IER screening was adopted less frequently in organizational samples compared with student and paid online samples.

However, IER is not uncommon in the organizational literature among both organizational and student samples. In organizational samples, Green and Stutzman (1986) found that 57% of 292 mental health workers reported spending time on tasks they could not have performed at work (i.e., infrequency tasks/items). Green and Veres (1990) found that 18% of 101

clerical workers and 61% of 118 police corporals indicated they had performed infrequency tasks that were not part of their jobs. Stetz, Button, and Quist (2012) reported that endorsement of six infrequency items ranged from 3% to 15% among 72 police officers. Hough, Eaton, Dunnette, Kamp, and McCloy (1990) chose a more stringent cutoff for carelessness (six out of eight endorsements on infrequency items) and identified 7.4% careless respondents in 9,152 military personnel. Meanwhile, in undergraduate student samples, Meade and Craig (2012) reported careless responding among 10%–12% of students in a lengthy survey. Chiaburu, Huang, Hutchins, and Gardner (2014) found that 15% of 133 undergraduate students devoted only 2 seconds per item in a low-stakes survey. Unfortunately, we could not find any published reports that directly address IER occurrence in crowdsourcing or other online samples.

Therefore, there is reason to suspect that respondents directly recruited from organizations may not always be attentive to survey measures. IER is likely present in datasets from a wide range of sample sources. Given the cursory nature of the review presented above, we attempted a more systematic examination of current IER screening practices in highly impactful journals in organizational research.

A Survey of IER Detection in Three Major Journals

Evidence-based strategies to detect and screen IER (e.g., infrequency, inconsistency, response pattern, response time, and multivariate outliers) have recently been advanced (Huang et al., 2012; Huang, Bowling, Liu, & Li, 2015; Meade & Craig, 2012). To compare quality control practices between sample sources, we examined data screening procedures pertaining to IER detection in recent articles published in the *Journal of Applied Psychology*, *Journal of Management*, and *Academy of Management Journal*. Our goals were to document common practices on quality control against IER and investigate whether more screening occurred for nonorganizational samples relative to organizational samples. We selected articles published between 2012 and 2014, following the release of two IER-related empirical articles in 2012 (Huang et al., 2012; Meade & Craig, 2012). The search resulted in 252 samples from *Journal of Applied Psychology*, 104 samples from *Journal of Management*, and 107 samples from *Academy of Management Journal*.

Based on our analysis, reporting practices to prevent or screen IER are very rare in these top-tier journal articles. Most studies utilize completion of focal measures (i.e., missing data) as a proxy for data quality to determine inclusion of cases in analyses. However, we were surprised to find that only three studies from the targeted journals from 2012 to 2014 reported screening for IER. Johnson and Allen (2013) and Pieterse,

Van Knippenberg, and Van Dierendonck (2012) utilized multivariate outliers to remove individuals whose responses constituted extremely divergent combinations of scores. Bagger and Li (2014) examined the consistency of responses to demographic items to ensure data quality. These three studies utilized an online panel (Johnson & Allen, 2013), a student sample (Pieterse et al., 2012), and a snowball sample of working adults (Bagger & Li, 2014). No studies using organizational samples reported IER detection or screening. Overall, very few studies reported detecting or screening IER, regardless of sample source. An exception is experimental studies, where manipulation checks are usually applied to screen participants who did not attend to the experimental stimuli. In our survey of articles, experimental manipulation checks were reported as a proxy of IER screening with both organizational and student samples (King et al., 2012).

Two explanations may be offered for the limited application of IER detection and screening to the organizational literature. First, IER may not occur frequently. However, the existence of IER within organizational and student sample sources has been documented in the literature (as reviewed earlier) and highlighted in recent Society for Industrial and Organizational Psychology presentations (e.g., Nieminen, Kotrba, Denison, & Carter, 2014). To capture the presence of IER across sample sources, we examined patterns of IER in our own datasets using samples of undergraduate students and Amazon Mechanical Turk (MTurk) workers. Responses to three extreme infrequency items (“I have never used a computer,” “I eat cement occasionally,” and “I can teleport across time and space”) were examined, where failure to disagree with these items indicated severe IER. Among three college student samples ($Ns = 402, 1,046, \text{ and } 1,250$), IER frequencies for individual items ranged from 3.7% to 17.5%. Among four samples from MTurk ($Ns = 325, 389, 415, \text{ and } 428$), item-level IER frequencies ranged from 2.5% to 11.2%. We conclude that most samples, regardless of source, are subject to the threat of IER. However, because different objectives, instructions, IER items, and survey measures were adopted across our datasets and in the publications reported earlier, we caution readers against making inferential IER comparisons between sample sources at this time.

This leads to a second possible explanation for the lack of IER detection and screening. Researchers, reviewers, and/or editors may overlook response effort as a critical determinant of validity or believe that it is a less important methodological factor compared with sample source, response rate, and other sample characteristics. Subsequently, insufficient attention to participant motivation may undermine the internal validity of results. It is important that researchers adopt an a priori IER screening method based on empirical recommendations (e.g., Huang et al., 2012; Meade & Craig, 2012)

and that reviewers and editors exert collective effort to create a standard for IER screening and reporting in the literature. This leads to our call for the advancement of evidence-based IER detection, screening, and prevention.

Call for Action

Organizational researchers have not paid sufficient attention to participant response effort nor consistently documented their treatment of IER. To better inform organizational researchers about IER, we call for more rigorous investigations to determine whether participant motivation and effort during surveys differ between convenience sampling sources (e.g., organizational, online, student samples). We provide three directions for future research on survey data quality assessment including the detection, screening, and prevention of IER. We recommend that these approaches be applied across sample sources.

Detection

The use of infrequency items, consistency indices, and multivariate outliers are promising approaches to detecting IER in surveys (Huang et al., 2012; Meade & Craig, 2012). However, these particular approaches may not be applicable to all surveys or other forms of measurement (e.g., behavioral measurement). Opportunities exist for researchers to generate and examine the effectiveness of alternative IER detection techniques. For instance, response latency at the item level can be used as a proxy for response effort. Behaviors associated with lack of effort (e.g., excessive number of breaks, noncompliance with instructions) may also indicate IER.

Screening

IER is a continuous construct. Identifying and removing cases based on IER require adopting an appropriate IER cutoff score. The lack of empirically derived cutoffs (for an exception, see Huang et al., 2012) may result in different sample sizes and power across studies, which may in turn affect study results. To address this issue, simulation studies may be employed to generate cutoff “standards” and provide evidence-based recommendations for researchers and practitioners.

Prevention/Intervention

Because IER detection and screening are reactive approaches to securing data quality, proactive techniques should be increasingly used to decrease IER and improve data quality. More research is needed to develop and examine the effectiveness of proactive approaches (e.g., monetary incentives, instructions, etc.) for promoting respondent effort. In addition, researchers should examine the extent to which different contextual factors (e.g., organizational support, survey design, etc.) influence response effort across sample

sources. Sample source may also moderate the effectiveness of these interventions.

Conclusion

In sum, we echo Landers and Behrend (2015) that data source does not necessarily equate to data quality and extend the discussion to consider internal validity as a function of IER. We found that IER appeared to exist across convenience sample sources, yet very few studies in top-tier organizational journals assessed IER. Therefore, we contend that evidence-based strategies should be used with greater frequency to detect and screen IER regardless of sampling strategy. We encourage more research on methods for detecting and reducing IER between sampling sources and provide recommendations for scholars to do so.

References

- Bagger, J., & Li, A. (2014). How does supervisory family support influence employees' attitudes and behaviors? A social exchange perspective. *Journal of Management*, *40*, 1123–1150. doi:10.1177/0149206311413922
- Barger, P., Behrend, T. S., Sharek, D. J., & Sinar, E. F. (2011). I-O and the crowd: Frequently asked questions about using Mechanical Turk for research. *The Industrial-Organizational Psychologist*, *49*(2), 11–17.
- Behrend, T. S., Sharek, D. J., Meade, A. W., & Wiebe, E. N. (2011). The viability of crowdsourcing for survey research. *Behavior Research Methods*, *43*, 800–813. doi:10.3758/s13428-011-0081-0
- Chiaburu, D. S., Huang, J. L., Hutchins, H. M., & Gardner, R. G. (2014). Trainees' perceived knowledge gain unrelated to the training domain: The joint action of impression management and motives. *International Journal of Training and Development*, *18*, 37–52. doi:10.1111/ijtd.12021
- Green, S. B., & Stutzman, T. (1986). An evaluation of methods to select respondents to structured job-analysis questionnaires. *Personnel Psychology*, *39*, 543–564. doi:10.1111/j.1744-6570.1986.tb00952.x
- Green, S. B., & Veres, J. G. (1990). Evaluation of an index to detect inaccurate respondents to a task analysis inventory. *Journal of Business and Psychology*, *5*, 47–61. doi:10.1007/bf01013945
- Hough, L. M., Eaton, N. K., Dunnette, M. D., Kamp, J. D., & McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. *Journal of Applied Psychology*, *75*, 581–595. doi:10.1037/0021-9010.75.5.581
- Huang, J. L., Bowling, N. A., Liu, M., & Li, Y. (2015). Detecting insufficient effort responding with an infrequency scale: Evaluating validity and participant reactions. *Journal of Business and Psychology*, *30*, 299–311. doi:10.1007/s10869-014-9357-6
- Huang, J. L., Curran, P. G., Keeney, J., Poposki, E. M., & DeShon, R. P. (2012). Detecting and deterring insufficient effort respond to surveys. *Journal of Business and Psychology*, *27*, 99–114. doi:10.1007/s10869-011-9231-8
- Huang, J. L., Liu, M., & Bowling, N. A. (2015). Insufficient effort responding: examining an insidious confound in survey data. *Journal of Applied Psychology*, *100*, 828–845. doi:10.1037/a0038510

- Johnson, R. C., & Allen, T. D. (2013). Examining the links between employed mothers' work characteristics, physical activity, and child health. *Journal of Applied Psychology, 98*, 148–157. doi:10.1037/a0030460
- King, E. B., Botsford, W., Hebl, M. R., Kazama, S., Dawson, J. F., & Perkins, A. (2012). Benevolent sexism at work gender differences in the distribution of challenging developmental experiences. *Journal of Management, 38*, 1835–1866. doi:10.1177/0149206310365902
- Landers, R. N., & Behrend, T. S. (2015). An inconvenient truth: Arbitrary distinctions between organizational, Mechanical Turk, and other convenience samples. *Industrial and Organizational Psychology: Perspectives on Science and Practice*. Advance online publication. Retrieved from <http://www.siop.org/journal/8.2/landers.pdf>
- Liu, M., Bowling, N. A., Huang, J. L., & Kent, T. (2013). Insufficient effort responding to surveys as a threat to validity: The perceptions and practices of SIOP members. *The Industrial–Organizational Psychologist, 51*(1), 32–38.
- McGrath, R. E., Mitchell, M., Kim, B. H., & Hough, L. (2010). Evidence for response bias as a source of error variance in applied assessment. *Psychological Bulletin, 136*, 450–470. doi:10.1037/a0019216
- Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods, 17*, 437–455. doi:10.1037/a0028085
- Nieminen, L., Kotrba, L., Denison, D., & Carter, N. (2014, May). *Insufficient effort responding: Uncovering an insidious threat to data quality*. In J. L. Huang & M. Liu (Co-chairs), *Insufficient effort responding to surveys: From impact to solutions*. Symposium presented at the annual conference of Society for Industrial and Organizational Psychology, Honolulu, HI.
- Pieterse, A. N., Van Knippenberg, D., & Van Dierendonck, D. (2012). Cultural diversity and team performance: The role of team member goal orientation. *Academy of Management Journal, 56*, 782–804. doi:10.5465/AMJ.2010.0992
- Stetz, T. A., Button, S. B., & Quist, J. (2012). Rethinking carelessness on job analysis surveys: Not all questions are created equal. *Journal of Personnel Psychology, 11*, 103–106. doi:10.1027/1866-5888/a000061

Participant Motivation: A Critical Consideration

Alyssa K. McGonagle
Wayne State University

Landers and Behrend (2015) make a good argument that more consideration should be given to sampling strategies in light of the specific research question prior to data collection and that nonorganizational samples should not be automatically dismissed by journal editors and reviewers. Yet, the authors

Alyssa K. McGonagle, Department of Psychology, Wayne State University.
Correspondence concerning this article should be addressed to Alyssa K. McGonagle, Department of Psychology, Wayne State University, 5057 Woodward Avenue, 7th Floor, Detroit, MI 48202. E-mail: alyssa.mcgonagle@wayne.edu