Mechanical Turk versus Student Samples Comparisons and Recommendations

Stephen De Lurgio
University of Arkansas, SDeLurgio@walton.uark.edu

Amber Grace Young
University of Arkansas, ayoung@walton.uark.edu

Zach Steelman
zsteelman@walton.uark.edu

Mechanical Turk versus Student Samples
Comparisons and Recommendations

Stephen De Lurgio SdeLurgio@walton.uark.edu; Zach Steelman PhD
ZSteelman@walton.uark.edu; Amber Young PhD AYoung@walton.uark.edu

Mechanical Turk and other online crowdsourcing markets have become a go-to source of data for researchers across scientific disciplines. Indeed, use of crowd-sourced samples is accepted in many reputable journals including *MIS Quarterly*, *Journal of MIS*, *Journal of the AIS*, and *Journal of Strategic Information Systems*. Steelman et al. (2014) investigated how Mechanical Turk data compared with data from student samples and consumer panels and found the data to be comparable and reliable for academic research. The use of Mechanical Turk has grown markedly since the original data collection and analysis approximately ten years ago. Research using Mechanical Turk in management increased 2,117% from 2012 to 2019 (Aguinis et al., 2021). To see if the findings of Steelman et al. (2014) still hold, we conduct a partial replication to compare Mechanical Turk workers with students using UTAUT 2 (Venkatesh et al., 2012) as our theoretical model and virtual reality headsets as the focal IT artifact. We collect student responses and three samples from Mechanical Turk: one composed of US workers, one of non-US workers, and one of workers from anywhere in the world. Like Steelman et al. (2014) we compare the demographic differences, psychometric properties, and measurement invariance across these samples. Our findings confirm that Mechanical Turk offers a viable alternative to student samples supporting Steelman et al. (2014). Yet, we did find evidence that Mechanical Turk samples drawn from US only, non-US only, and worldwide populations respond more similarly to each other than to students. Student samples were less likely to confirm well-established theoretical relationships with their responses than were Mechanical Turk workers. This may be due to significant demographic differences across samples. Student samples were least diverse in terms of age, race, education level, and income. Thus, Mechanical Turk samples may be a better representation of the general population than students. We call for IS scholars to embrace replication research as a vital part of scientific inquiry. Future research that replicates published studies using a variety of samples will contribute to understanding of how sampling bias affects research findings. Comparing findings across samples can reveal boundaries of generalizability. Further, replication of high-impact studies as technology evolves is crucial for keeping the IS field up to date. Future research comparing Mechanical Turk with other online crowdsourcing markets like Prolific Academic is needed.

References
Aguinis, H., Villamor, I., & Ramani, R. S. (2021). MTurk research: review and recommendations. *Journal of Management*, *47*(4), 823–837.
Steelman, Z. R., Hammer, B. I., and Moez Limayem. (2014). Data collection in the digital age. *MIS Quarterly*, 38(2), 355–378.
Venkatesh, V., Thong, J. Y. L., and Xu, X. (2012).Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, 36(01), 157–178.

Presentation at TREO Talks in conjunction with the 42nd International Conference on Information Systems, ICIS 2021
TREO Talks are not peer-reviewed and not a formal part of the ICIS 2021 Proceedings
All TREO Talks are available in the TREO Talks section of the AIS e-Library