Agreement Between Online and in-Person Auditory Causal Property Ratings

CMU Kilthub, May 27, 2022. DOI: https://dx.doi.org/10.1184/R1/19914853
Laurie M. Heller, Department of Psychology, Carnegie Mellon University, Pittsburgh PA 15213

In 2020, Auditory Lab at Carnegie Mellon conducted an online replication of a 2019 in-person study. The purpose was to compare the accuracy and quality of auditory perception data from the two data collection methods.

In 2019, we collected in-person ratings of causal properties of a set of environmental sounds. Listeners with self-reported normal hearing listened to a set of 50 everyday sounds from the Dataset of Environmental Sound Classification (Piczak, 2015). In each trial, listeners made a set of ratings of one sound by indicating how likely it was that the sound event was caused by a certain action. For example, on a given trial, a listener might hear the sound of clapping and subsequently be asked to rate how likely the sound was to have been produced by each of 22 different actions: dripping, splashing, tapping, scraping, crushing, crumpling, breaking, sawing, rubbing, rolling, blowing, pouring, exploding, exhaling, rotating, vibrating, wailing, groaning, calling, gasping, ringing, or singing. The verbs were chosen to pertain to at least two items and not uniquely identify any one sound (i.e. they were causal properties of some sounds but were not sound labels). Each of the 37 participants judge all actions for all sounds with the presentation; the presentation order of sounds was randomized. This task was performed in the Auditory Lab at CMU using a traditional psychoacoustic approach. Listeners sat in a sound attenuating booth and listened over high-quality headphones. The results of this study were reported by Heller & Sheikh (2019).

In 2020, we closely replicated that study online using the Qualtrics survey platform. Eight of the 50 sounds were identical to the 2019 study and 19 of the 22 verbs were the same. All the procedures were the same except that we instructed 49 participants to wear headphones in a quiet location of their own choosing. After informed consent, the experiment began with a trial in which they adjusted their headphone volume to a comfortable level.

We compared the average results between these two studies for the 8 overlapping sounds and found excellent agreement between them, with an average correlation of 0.95. The table showing this result is included below. Each entry shows the correlation between the lab and online study of the average participant ratings for the set of 19 overlapping actions for a given sound.

We propose that our online results are reliable for a few reasons, as follows. Our participants completed all ratings on each of the stimuli. This within-subject design took almost an hour and therefore would not have been feasbile on some other online platforms that allow each participant to perform a task for only a few minutes. In addition, our student population understands the value of behavioral data because they are undergraduates enrolled in an introductory course in a Psychology department. This might have made our participants especially compliant with the headphone setup task and the instructions. In addition, we do not need to weed out responses from bots and our undergraduate populations were very similar between the studies.

	Car horn	Clapping	Clock tick	Crickets	Engine	Glass breaking	Mouse click	Rain
Car horn	**0.966**	0.300	0.385	0.494	0.259	-0.139	0.359	-0.302
Clapping	-0.040	**0.842**	0.651	-0.055	0.259	0.095	0.739	0.424
Clock tick	0.167	0.961	**0.961**	-0.045	0.231	-0.116	0.989	0.034
Crickets	0.373	0.087	0.262	**0.926**	0.716	-0.167	0.174	-0.211
Engine	0.119	0.278	0.454	0.703	**0.960**	0.153	0.347	-0.173
Glass breaking	-0.159	-0.003	-0.141	-0.225	-0.068	**0.986**	-0.085	-0.088
Mouse click	0.209	0.980	0.929	-0.052	0.193	-0.085	**0.993**	0.044
Rain	-0.346	0.196	-0.024	-0.253	-0.158	-0.047	0.058	**0.963**
							Correlation Avg.	**0.950**

References

Heller, L., & Sheikh, A. (2019). *Acoustic features of environmental sounds that convey actions*. Auditory Perception, Cognition, and Action Meeting, November 2019. https://apcsociety.org/apcam/2019/#heller

Piczak, K. J. (2015). ESC: Dataset for Environmental Sound Classification. *Proceedings of the 23rd ACM International Conference on Multimedia*, 1015–1018. https://doi.org/10.1145/2733373.2806390

Acknowledgements:

Thank you to Asadali Sheikh for collecting some of the data, Jessica Smith for formatting the table, Seojun Jang for verifying the data, and Urszula Oszczapinksa for helpful comments.