Massive Online Experiments in Cognitive Science

Joshua Hartshorne (joshua.hartshorne@bc.edu)
Department of Psychology, Boston College

Joshua R. de Leeuw (jdeleeuw@vassar.edu)
Department of Cognitive Science, Vassar College

Laura Germine (laura@manybrains.net)
Department of Psychiatry, Harvard Medical School
Institute for Technology in Psychiatry, McLean Hospital

Katharina Reinecke (reinecke@cs.washington.edu)
Paul G. Allen School of Computer Science and Engineering, University of Washington

Mariela Jennings (marielajennings@gmail.com)
Department of Psychology, Boston College

Keywords: online experiments, experimental design, methodology

Overview

This full-day workshop focuses on *Massive Online Experiments* (MOEs). MOEs have transformative potential, as they effectively allow researchers to run hundreds of experiments simultaneously (cf. Hartshorne & Germine, 2015; Reinecke & Gajos, 2014). The goal of the workshop is to help a broad cross-section of cognitive scientists begin to incorporate MOEs into their research.

What are Massive Online Experiments? Massive Online Experiments use samples of tens or hundreds of thousands of volunteer subjects. Unlike standard online experiments that recruit subjects through online labor markets like Amazon Mechanical Turk or Prolific, MOEs are usually run through researcher-developed websites like gameswithwords.org, testmybrain.org, and labinthewild.org. Subjects are unpaid volunteers who choose to participate because the experiments are intrinsically motivating. Experiments often take the form of a game or of a ”brain test” that allows subjects to learn about themselves. Durations are typically short (under 15 minutes), though longer and/or longitudinal studies are possible.

Why use Massive Online Experiments? The large, demographically diverse set of volunteers that participate in MOEs allow researchers to design experiments that are not feasible in the traditional laboratory or online labor market model. MOEs enable statistically powerful comparisons across demographic variables like age, culture, and education level. Previous MOEs have simultaneously probed cross-cultural differences in domains such as moral reasoning, personality, and aesthetic preferences (Bleidorn et al., 2016; Gebauer et al., 2014; Reinecke & Gajos, 2014). Moreover, large samples allow researchers to test the generalizability of findings across different theoretically-relevant parameters by running dozens of experimental variations simultaneously. When MOEs are combined with modern statistical analyses such as multi-level modeling, researchers can study generalization across large numbers of subjects and items simultaneously (Baribault et al., 2017).

The Challenges of Massive Online Experiments. There are two major challenges to MOE adoption. First, researchers are trained to think about experiments under the constraints of the traditional laboratory approach. The benefits of MOEs are only realized in full if the experimental design is tailored to the format. Considerations include both constraints — such as designing studies that subjects are intrinsically motivated to complete — and also the extraordinary opportunities that MOEs provide.

The second set of challenges are technical. MOEs present an additional layer of technical demands beyond a standard online experiment. Researchers must host an entire virtual laboratory to attract subjects. Studies may go viral, causing substantial traffic spikes. Assignment of items or conditions to subjects must be carefully done when there are tens of thousands of items, since random assignment at that volume will inevitably over- and under-sample some items.

Workshop Format

The workshop will be divided into four sessions. The first session will provide a motivational overview: What are MOEs, how do they work, and what are they good for? This will consist of three talks:

Massive Online Experiments: Challenges, history, and prospects, Presenter: Joshua Hartshorne

This talk will present the motivation for conducting MOEs and how they can be used to address common limitations and frustrations of in-lab experiments. I review and address common concerns, such as limitations in available methods and data quality. I also address the barriers to a more widespread implementation of MOEs and preview some of the emerging solutions, which will be covered in more detail in the afternoon.

Massive Online Experiments towards Population-level Cognitive Science, Presenter: Laura Germine

For this talk, I approach the applicability of MOEs for addressing some of the biggest challenges in public health research. First, MOEs help address issues such as power, generalizability, and rapid translation that are critical bottlenecks for health research. Second, MOEs open the door to a new population-level science of cognition, which seeks to understand and characterize cognitive functioning from the perspective of population-level individual differences. Finally, I will argue that MOEs are not a convenience for researchers, but rather an ethical imperative towards building a robust and representative science of cognition.

Massive Online Experiments for Research Across Demographic and Geographic Boundaries, Presenter: Katharina Reinecke

In this talk, I will show how MOEs can help overcome the replicability crisis in cognitive science and related fields by studying larger and more diverse populations than possible in laboratory studies or in studies conducted on online labor markets. I will discuss how offering participants personalized feedback at the end of the study attracts people with varying motivations for participating in such experiments. The talk will also show how MOEs can be mutually beneficial to both participants and researchers by providing learning opportunities for participants as well as reliable data quality for researchers.

The second session will explore recent examples of MOEs across a range of disciplines. This will include presentations of work by Drs. Hartshorne, Germine, and Reinecke, covering the fields of psycholinguistics, clinical psychology, and human-computer interaction, respectively. Additional presentations will be selected from an open call. Depending on the number of presentations, they will take the form of either a series of rapid oral presentations or a poster session. The second session will conclude with an open panel discussion and Q&A focused on the uses of MOEs.

The two afternoon sessions will introduce participants to technology for MOEs. The third session will focus on jsPsych, a JavaScript framework for creating online studies (de Leeuw, 2015). It can be used for both traditional online and laboratory studies, as well as MOEs. Josh de Leeuw will provide a tutorial overview of how to use jsPsych for online studies broadly.

The fourth session will cover Pushkin. While jsPsych addresses many basic needs for running experiments online, Pushkin address the additional challenges of MOEs, such as handling massive traffic fluctuations and efficiently distributing large numbers of items across large numbers of subjects. Pushkin provides many additional resources for maintaining an Internet laboratory: a website template, an interactive forum, social media integration, and the ability for subjects to create an account (useful for longitudinal studies, etc.). Mariela Jennings will provide a demonstration and tutorial.

The workshop will conclude with open discussion of the challenges and benefits of adopting MOEs. The aim of this discussion is to help attendees consider ways in which MOEs could augment their own research, and help the presenters better understand the needs of the cognitive science community to guide future development efforts on MOE technology and training.

References

Baribault, B., Donkin, C., Little, D. R., Trueblood, J., Oravecz, Z., van Ravenzwaaij, D., … Vandekerckhove, J. (2017). Meta-studies for robust tests of theory. Retrieved from https://osf.io/preprints/g84py/

Bleidorn, W., Schönbrodt, F., Gebauer, J. E., Rentfrow, P. J., Potter, J., & Gosling, S. D. (2016). To Live Among Like-Minded Others: Exploring the Links Between Person-City Personality Fit and Self-Esteem. *Psychological Science*, *27*(3), 419–427.

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12.

Gebauer, J. E., Bleidorn, W., Gosling, S. D., Rentfrow, P. J., Lamb, M. E., & Potter, J. (2014). Cross-cultural variations in big five relationships with religiosity: a sociocultural motives perspective. *Journal of Personality and Social Psychology*, *107*(6), 1064–1091.

Hartshorne, J. K., & Germine, L. T. (2015). When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span. *Psychological Science*, *26*(4), 433–443.

Reinecke, K., & Gajos, K. Z. (2014). Quantifying Visual Preferences Around the World. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 11–20). New York, NY, USA: ACM.