A Data Revolution in the Cognitive Sciences

John V. McDonnell (john.mcdonnell@nyu.edu) (organizer)
New York University Department of Psychology, 6 Washington Pl. 8th Floor
New York, NY 10003 USA

Winter Mason (m@winteram.com)
Howe School of Technology Managemeent, Stevens Institute
Hoboken, NJ 07030, USA

John Myles White (jmw@johnmyleswhite.com)
Princeton University Department of Psychology
Princeton, NJ 08540-1010 USA

Keywords: Machine Learning; Cognitive Neuroscience; Online Experiments; Social Network Analysis; Text Analysis; Methods

In the 1980s and 90s, the broad availability of PCs power fueled a renaissance in cognitive science, changing the way we run experiments, the tools we use for analysis, and lowering computational barriers to the development of sophisticated computational models of cognition. As technology continues to advance, the advent of big data is now leading to methodological advances on various fronts, from online experimentation to the mining of large datasets. These new opportunities have also come with new challenges, as scientists must overcome the challenges these methods bring. The goal of this symposium would be to highlight how researchers are making use of these new tools and the techniques they have used to solve the challenges they represent. Speakers will discuss online laboratory experiments, machine learning techniques applied to neuroscientific imaging, and online social network analysis.

The emergence of social conventions in social networks

Winter Mason, Meeyoung Cha, Krishna Gummadi, Farshad Kooti, and Haeryun Yang Social conventions and norms are a powerful and ubiquitous guide for behavior. However, the way in which conventions emerge in communities is not very well understood, largely due to a paucity of available data. In this study, we leverage the widespread use of the micro-blogging platform Twitter and focus on competing conventions for attributing reposts to the original source. We analyze over 1.7 billion "tweets" from 54 million users, from the very first tweet in 2006 up to September 2009, and observe how the conventions emerged and spread through the network of Twitter users. We observe that initially the most successful conventions were borrowed from natural language ("via" and "retweeting"), but eventually a community-specific convention came to dominate ("RT"). We see that some conventions are used by divergent groups of people, while others are abandoned in favor of more efficient expressions. We also observe the failure of some conventions despite higher efficiency (i.e., fewer characters) and explicit endorsement of their adoption. Our results suggest that there are some features that encourage the adoption of one convention over another, but that there is still significant inherent unpredictability in what convention will come to dominate.

Context and decision making in a massive online experiment

John Myles White Online labor markets, like Amazon's Mechanical Turk (AMT), offer psychologists many opportunities. The convenience of the virtual lab provided by AMT has already won over many psychologists, but transitioning research from the lab to the web browser offers other benefits. AMT allows psychologists to deploy experiments that are fully automated: The recruitment, instruction, core experiment and debriefing periods can be identical for all subjects, which largely removes the possibility that undocumented components of an experiment might exert substantive influences on the final results (Doyen et al., 2012). Moreover, the ease of recruiting large numbers of subjects for tasks provides important increases in statistical power, which can allay concerns about false positive psychology (Simmons et al., 2012). But I will argue that the primary value of AMT is not purely methodological: the strength of the virtual lab is that it allows psychologists to pursue large-scale between-subjects designs in which a large number of subjects perform a very small number of trials. This type of research is often eschewed because of the prohibitive cost of recruiting hundreds or thousands of subjects in the lab, but previous research (e.g. Gneezy et al. 2006) suggests that studies of decision-making can be powerfully influenced by contextual cues. Our recent work has found evidence that the effects of context on decisions may be even more problematic than previously believed: in our studies of decision-making, we find that the effects of ostensibly innocuous local context can shift the basic qualitative results of experiments.

Can online data be trusted? Learning tasks on Amazon's Mechanical Turk.

John V. McDonnell, Todd M. Gureckis Amazon Mechanical Turk (AMT) has attracted attention from experimental psychologists interested in gathering human subject data more efficiently. However, relative to traditional laboratory studies, many aspects of the testing environment are not under the experimenter's control. We have empirically evaluated the fidelity of AMT for use in cognitive behavioral experiments. These types of experiment differ from simple surveys in that they require multiple trials with sustained attention from participants. Our initial attempts to replicate the classic Shepard, Hovland, and Jenkins (1961) task were only partially successful. However, after systematically studying the effects of compensation and validation we were able to more closely match previous in-lab findings. Specifically, we found that compensation altered the rate of sign-ups, but not the quality of the data. Conversely, we found that using a strict measure to ensure that participants had understood instructions resulted in a considerable improvement in performance and convergence with in-lab findings.

References

Doyen, S., Klein, O., Pichon, C. L., & Cleermans, A. (2012). Behavioral priming: It's all in the mind, but whose mind? *PloS one*, *7*, e29081.

Gneezy, U., List, J. A., & Wu, G. (2006). The uncertainty effect: When a risky prospect is valued less than its worst possible outcome. *The Quarterly Journal of Economics*, *121*, 1283–1309.

Shepard, R., Hovland, C., & Jenkins, H. (1961). Learning and memorization of classifications. *Psychological Monographs*.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*, 1359–1366.