

Leveraging the Wisdom of Crowds in a Data-Rich Utopia

Ravi Iyer & Jesse Graham

University of Southern California

Commentary on “Scientific Utopia: I. Opening scientific communication”

by Brian A. Nosek & Yoav Bar-Anan, *Psychological Inquiry*

Correspondence concerning this article should be addressed to Ravi Iyer, Department of Psychology, University of Southern California, SGM 501, 3620 South McClintock Ave., Los Angeles, CA 90089-1061. E-mail: raviyer@usc.edu

Nosek and Bar-Anan suggest that psychologists should embrace digital communication in order to create a system that accomplishes a number of worthy goals. Their proposed system would improve access to research, reduce the burden on reviewers, and reduce the financial burden on universities. It would undoubtedly decrease the excessive amount of time between the conduct of research and the publication of results. However, for many, speed and efficiency are far less important than the core purpose of science: the production of accurate knowledge. Currently the burden of identifying truth falls to editors and reviewers who altruistically donate their efforts toward the production of better knowledge. There is no pretense of efficiency, but the hope is that expert peer review leads to the dissemination of articles with well-

The Wisdom of Crowds in a Data Rich Utopia

supported results. The recommendations that Nosek and Bar-Anan make are likely to also be evaluated on this singular dimension by many social scientists. Specifically, can the “wisdom of crowds” approach that they advocate – where any member of a professional society can review any article – truly lead to better scientific knowledge? In this commentary we suggest that it can, in part by reducing the systematic biases inherent in a closed review system limited to psychologists. Given the diversity of fields and organizations using data to answer questions of psychological interest, open peer review can not only improve the efficiency of our science, but the quality of our science as well.

When does the Wisdom of Crowds work?

The “wisdom of crowds” effect (Galton, 1907; Surowiecki, 2004) is based on a mathematical principle that psychologists use routinely. Specifically, just as averaging across questions that measure some psychological construct reduces the error of the overall measurement, averaging across the opinions of individuals reduces impact of the error contributed by each individual opinion. We naturally take advantage of this principle when we ask several friends for their opinions about a restaurant before forming an opinion or when we ask for several professional opinions about a medical diagnosis. In psychology, we take advantage of this principle when we do meta-analyses aggregating across studies, or when we ask for the opinions of several expert reviewers of a journal article as opposed to just one. If error in the review process were not a concern, there would be no need to ask more than one reviewer for an opinion.

A great deal of research has been done on the “wisdom of crowds” effect to determine the circumstances under which crowds outperform individuals. Specifically,

The Wisdom of Crowds in a Data Rich Utopia

the effectiveness of crowds depends upon two factors: expertise and diversity (Larrick, Mannes, & Soll, 2011). Members of the crowd must have some expertise about the question being asked; there must be some signal from which the noise can be separated. Members of the crowd must also make independent, largely uncorrelated errors, in order for aggregation to cause the error to cancel out. “Wise” crowds, which meet these criteria, have been shown to be better than political scientists at predicting election outcomes (Sjoberg, 2009) and as good as FIFA at predicting soccer outcomes (Herzog & Hertwig, 2011).

Do participants in an open review process who hold “membership in a professional society” have the requisite expertise to evaluate psychological research? The example from the target article, in which errors in the analysis of the Implicit Association Test were not caught, is instructive. Expertise is becoming more specialized and our experience converges with Nosek and Bar-Anan’s assertion that reviewers within psychology do not always have the comprehensive expertise to simultaneously “evaluate the theory, design, instrumentation, analysis strategy, and interpretation of every component” of submitted articles. For example, we sometimes receive reviews containing reservations about potential issues in conducting internet research, but no reviewer has ever suggested using server logs, IP addresses, or browser data to examine these issues, likely because the specialized knowledge of internet instrumentation is outside of most psychologists’ expertise. As instrumentation, statistical methods, and domains of inquiry become more specialized, membership in a discipline (or inclusion in any journal’s list of potential reviewers) is becoming a weaker predictor of expertise regarding any individual aspect of a paper. Some of the most instructive criticisms we

The Wisdom of Crowds in a Data Rich Utopia

have received in our work have been from political scientists, internet programmers, political pundits, journalists, philosophers, and biologists, who all have domain expertise in questions we address. Participants themselves, once debriefed about our hypotheses, often have some of the most insightful comments. While none of these non-psychologists may individually have a better overall perspective than the average psychology researcher, in aggregate, their collective expertise is far more comprehensive.

The question of independence is even more problematic for the current review process, as the kinds of errors that psychologists make are likely to be positively correlated. Cooperative groups (Lee & Shi, 2010) in which social influence can influence evaluations (Lorenz, Rauhut, Schweitzer, & Helbing, 2011) have been found to make worse aggregate judgments, and as such, peer review conducted by members of a single discipline, who socialize and have collective norms, is likely to contain positively correlated error amongst members. The result is likely to be the introduction of systematic error of the type identified by Paul Rozin (2001), whereby the study of specific phenomena is collectively dismissed as unimportant and specific methodologies are collectively seen to be unacceptable. The novel perspectives that an open review process could bring could mitigate the systematic error that naturally occurs in closed review systems.

Data Abundance vs. Data Scarcity

A closed review system may have been adaptive in an environment where the use of data to answer questions about human thought and behavior was uncommon. In such an environment, most of the expertise regarding how to collect, analyze, and interpret

The Wisdom of Crowds in a Data Rich Utopia

such data may have been concentrated within a single academic discipline. However, the use of data to answer questions of interest to psychologists is increasingly widespread, with teams of data scientists at companies like Facebook, Target, and Google, data journalists like Nate Silver at *The New York Times*, and even individuals interested in self-discovery (e.g. www.quantifiedself.com; Swan, 2009) increasingly collecting and analyzing data about human thought and behavior. The abundance of individuals with the interest and expertise to comment on empirical psychological data is a relatively new phenomenon, and statistical discussions are increasingly common in online discussions of political polling (e.g. fivethirtyeight.blogs.nytimes.com) and baseball outcomes (e.g. www.fangraphs.com), indicating that an informed and interested public – again, in the aggregate – could provide meaningful input in the review process.

Further, the abundance of data that exists and the ease with which data can be collected through programmer APIs (application programming interfaces), open protocols, and services like Mechanical Turk, means that collective open peer review can include the ultimate test of any study – namely, replication. At our current research website, yourmorals.org, we currently have data that replicates a great deal of psychological research, but no incentive or natural method for publicizing these replications. Other individuals, organizations, and research labs likely have data that replicates, or fails to replicate, current research. In the ongoing open peer review process that Nosek and Bar-Anan propose, we could connect our replications to current published research within this process, creating greater confidence in published findings. In such an environment, research practices such as “p-hacking” (Simmons, Nelson, & Simonsohn, 2011) or outright data fabrication (Vogel, 2011) would be disincentivized as

The Wisdom of Crowds in a Data Rich Utopia

researchers could count on any notable finding being tested publicly. Such practices would effectively leverage the mathematical principle behind the wisdom of crowds to reduce the error associated with data collection, in addition to the error associated with the review process, further improving the pursuit of truth.

Avoiding the Fate of Record Labels and Newspapers

Psychologists currently inhabit a uniquely valuable space in the world of ideas, at the intersection of two societal trends. Increasingly, the world is concerned with improving psychological outcomes (e.g. Di Tella & MacCulloch, 2008) as physical needs become less of a concern for many citizens of relatively wealthy countries (Inglehart & Abramson, 1994). Increasingly, the world seeks to use empirical data to achieve any important goal, as evidenced by demand for data scientists and data journalists at a wide range of companies. The unique expertise of psychologists is the use of data to predict psychological outcomes and, as such, some of the world's most prominent and forward thinking companies can be seen at our annual meetings (e.g. Kramer, 2012).

However, being at the intersection of societal trends can be both a blessing and a curse. For example, technological advances in how we share and distribute information have vastly improved the accessibility and consumption of digitally distributable goods, such as news and music. Yet rather than benefiting from these technological trends, newspapers and record labels, which previously controlled the distribution of this information, have been marginalized in favor of more forward-thinking organizations. Businesses like StumbleUpon, Digg, Pandora, and Spotify leverage aggregated data provided by their users to perform the editorial functions that previously were provided

The Wisdom of Crowds in a Data Rich Utopia

by a select few. Nosek and Bar-Anan's suggestions that psychologists embrace technological change provide an opportunity for academic psychologists to be one of these forward thinking groups, leveraging technology to provide both efficient and high quality information, and retaining our unique place in the marketplace of ideas.

References

- Di Tella, R., & MacCulloch, R. (2008). Gross national happiness as an answer to the Easterlin Paradox? *Journal of Development Economics*, *86*(1), 22–42.
- Galton, F. (1907). Vox populi. *Nature*, *75*, 450–451.
- Herzog, S. M., & Hertwig, R. (2011). The wisdom of ignorant crowds: Predicting sport outcomes by mere recognition. *Judgment and Decision Making*, *6*(1), 58–72.
- Inglehart, R., & Abramson, P. R. (1994). Economic security and value change. *American Political Science Review*, 336–354.
- Kramer, A. (2012). Ripples in the Ocean: Emotional Contagion on Facebook. Presentation at the Society for the Study of Personality and Social Psychology annual meeting.
- Lee, M. D., & Shi, J. (2010). The accuracy of small-group estimation and the wisdom of crowds. *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 1124–1129).
- Larrick, R, Mannes, A., & Soll, J. (2011). The social psychology of the wisdom of crowds. In J. Krueger (Ed.), *Frontiers of Social Psychology: Social Psychology and Decision Making*. Philadelphia, PA: Psychology Press.
- Lorenz, J., Rauhut, H., Schweitzer, F., & Helbing, D. (2011). How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences*, *108*(22), 9020.
- Nosek., B., & Bar-Anan, Y. (2012). Scientific Utopia: I. Opening scientific communication. *Psychological Inquiry*
- Rozin, P. (2001). Social psychology and science: Some lessons from Solomon Asch. *Personality and Social Psychology Review*, *5*(1), 2–14.

The Wisdom of Crowds in a Data Rich Utopia

- 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(11), 1359–1366.
- Sjöberg, L. (2009). Are all crowds equally wise? A comparison of political election forecasts by experts and the public. *Journal of Forecasting*, 28(1), 1–18.
- Soll, J. B. (n.d.). *The Social Psychology of the Wisdom of Crowds* Richard P. Larrick Duke University Albert E. Mannes Carnegie Mellon University.
- Surowiecki, J. (2004). *The wisdom of crowds*. New York: Random House.
- Swan, M. (2009). Emerging patient-driven health care models: an examination of health social networks, consumer personalized medicine and quantified self-tracking. *International journal of environmental research and public health*, 6(2), 492–525.
- Vogel, G. (2011). Psychologist Accused of Fraud on 'Astonishing Scale'. *Science*, 334(6056), 579–579.
- Vul, E., & Pashler, H. (2008). Measuring the Crowd Within. *Psychological Science*, 19(7), 645.