

Evidence of Fraud in an Influential Field Experiment About Dishonesty – A Personal Reply

Max H. Bazerman

[This post is singly authored, without review by the other authors involved in the 2012 publication described. I want to thank Uri Simonsohn, Joe Simmons, Leif Nelson and the anonymous researchers that they mention for helping to correct the scientific record.]

I am completely convinced by the analyses provided by Simonsohn, Simmons, and Nelson and their conclusion that the field experiment (Study 3) in Shu, Mazar, Gino, Ariely, and Bazerman (2012) contains fraudulent data; as a result, Shu, Gino, and I contacted PNAS to request retraction of the paper on July 22, 2021. I was not directly involved in the collection or statistical analysis of the data in Study 3. Nonetheless, I was a coauthor of the study. It is my hope that laying out my perspective on how events unfolded will help others avoid problems in the future.

Shu, Mazar, Gino, Ariely, and Bazerman (2012) came together in a merger of two prior non-published empirical efforts. Mazar-Ariely independently provided the data for Study 3, while Shu-Gino-Bazerman had written a paper containing two laboratory experiments (Studies 1 and 2). The Shu-Gino-Bazerman group knew of the Mazar-Ariely data from multiple public presentations by Ariely. Each project appeared to respond to limitations of the other, and both projects focused on the prediction that signing before filling out a form leads to greater honesty than the traditional process of signing after.

There were indications of problems from the start (2011).

The first time I saw the combined three-study paper was on February 23, 2011. On this initial reading, I thought I saw a problem with implausible data in Study 3. I raised the issue with a coauthor and was assured the data was accurate. I continued to ask questions because I was not convinced by the initial responses. When I eventually met another coauthor responsible for this portion of the work at a conference, I was provided more plausible explanations and felt more confidence in the underlying data. I would note that this coauthor quickly showed me the data file on a laptop; I did not nor did I have others examine the data more carefully.

We published the 2012 paper and it received a great deal of attention. I then believed the core result – that signing first leads to greater honesty than signing after. I presented our work in academic contexts and taught the finding to MBA and executive audiences. Multiple organizations implemented our idea of moving the signature.

In 2017-2019, when we tried to replicate and extend the finding from the 2012 paper, we were unable to do so. We reported this lack of replication in two 2020 papers and backed away from the conclusions of the 2012 paper.

When I began a new project using mturk experiments with Ariella Kristal and Ashley Whillans on how to induce honesty online, we started by using the “signing first” strategy as a demonstration that the existing literature provided hints on how to induce honesty. I believed that this was going to be a straightforward extension study, replicating the 2012 result in an online context.

Our first attempt did not find a difference between study participants who signed before or after, failing to provide a replication of the 2012 paper. Four additional experiments yielded the same result. We decided to do a pure, large sample size replication of the first lab experiment (Study 1). Kristal and Whillans advocated for inviting the other four authors from the original paper into the project, to make it more collaborative rather than adversarial. All four (Shu-Mazar-Gino-Ariely) agreed to join the efforts in running the large-scale pure replication, and all seven authors eventually published our paper documenting our inability to replicate the 2012 finding in [PNAS](#) as well as an article in [Scientific American](#) both explicitly rejecting the conclusion of the original study. Based on this massive failed replication project, I did my best to express my lack of confidence in any of the results in the 2012 paper in our 2020 PNAS publication.

Our work in 2019 uncovered red flags with the field experiment, but we did not retract the paper.

In the process of working on the 2020 paper, Kristal uncovered an unexplainably large difference in the pre-measure odometer reading (pre-treatment) between conditions, raising questions about whether a randomized experiment took place. I expressed these concerns to my co-authors and was told that the randomization failure had been discussed collectively in 2012. I have no memory of such a conversation and can find no evidence of such a conversation having included me. I would not have agreed to include the study had I known of these issues.

After the publication of the 2020 PNAS paper, PNAS raised the issue of whether the 2012 paper should be retracted. I emailed my coauthors arguing vigorously for retraction, and followed up explicitly noting the lack of randomization and the lack of transparency about the work.

Shu and I were the only two of the original five authors explicitly in favor of retraction, and lacking a majority we did not retract the 2012 paper. I now believe I should have independently taken action to push for retraction even without a majority of co-authors in favor of such action.

In sum, I wish I had worked harder to identify the data were fraudulent, to ensure rigorous research in a collaborative context, and to promptly retract the 2012 paper. While I had doubts and raised questions, I believed the responses I received. We reported our failure to replicate the 2012 finding, but I should have argued more forcefully to retract the paper sooner.

I wish to thank Ariella Kristal and Ashley Whillans for their excellent work in setting the research record straight on the signing first phenomenon, as well as Uri Simonsohn, Joe Simmons Leif Nelson, and the anonymous researchers for all they do to help us create social science in a manner that the world can trust.

References

Shu, L., Mazar, N., Gino, F., Ariely, D., & Bazerman, M. Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end. *Proceedings of the National Academy of Sciences*, 2012, 109: 38, 15197-15200.

Kristal, A.S., Whillans, A.V., Bazerman, M.H., Gino, F., Shu, L.L., Mazar, N., & Ariely, D. Signing at the beginning versus at the end does not decrease dishonesty. *Proceedings of the National Academy of Sciences*. 2020, 117, 7103- 7107.

Kristal, A.S., Whillans, A.V., Bazerman, M.H., Gino, F., Shu, L.L., Mazar, N., and Ariely, D. When we're wrong, it's our responsibility as scientists to say so. *Scientific American*, March 21, 2020.

Sincerely,

A handwritten signature in black ink, appearing to read 'M. H. Bazerman', with a long horizontal flourish extending to the right.

Max H. Bazerman