A Note on the Provenance of My Data – A "Data Chain of Custody"

Recent cases of fraud in the behavioral sciences showed that replicating results based on posted data is a different problem than validating whether posted data are authentic in the first place. The field is still struggling with establishing mechanisms to prevent and detect fraud. However, many solutions being proposed involve some version of a "data chain of custody" (DCOC) – a document that details how data were collected, who was involved, and who produced the published results.

The DCOC document linked here includes all published studies in all projects for which I began data collection after the launch of the Open Science Framework. Having closely examined the origins of my data, I can confidently state that they were collected as reported in the published manuscripts. In the interest of complete transparency, my co-authors and I are happy to make the raw data available to any interested party, contingent on IRB approval.

A DCOC document can help to both prevent and detect fraud. Most obviously it lays the responsibility for the data collection process at the foot of clearly identified individuals. In cases of retrospectively trying to establish the credibility of a body of work, it tells anyone interested who to ask for access to the raw data. Examining the path the data took from generation to publication makes it easier to consider when and how data could have been tampered with and how credible these risks may be.

Most of my data were collected via Qualtrics surveys that were administered either to online participants, to behavioral lab participants, or occasionally to special populations such as Executive Education learners. Qualtrics surveys are often created by one author and shared among the authorship team and research assistants or are created by a research assistant and then shared with the authors. Qualtrics keeps a record of who each survey is shared with (although only the original creator of the survey can see that record and occasionally this record can be lost when researchers move schools). In the wake of the Gino scandal, one of my co-authors, M. Yeomans also undertook an external audit required by his university of three of our joint papers which entailed sharing the relevant Qualtrics files with external auditors.

Qualtrics also keeps a record of any changes made to data from "the back end." If participants enter data and a researcher then decided to change the data for any reason, the cell that has been changed is marked. For this reason, knowing where and how access to the original survey can be obtained assuages many concerns about data integrity. This of course is not perfect proof– presumably there are as many ways to commit data fraud as there are ways to rob banks – however, because many of those ways seem highly improbable, risk is minimized when raw Qualtrics data can be inspected.

More people having access to a Qualtrics survey is both a risk and a benefit to data integrity. It is a risk because every collaborator on a survey could accidentally or intentionally change data. It is a benefit because more people being involved increases the chances of somebody noticing something amiss. More collaborators also means that more people can download and share the raw data with interested parties, whether journal editors, other researchers, auditors, or administrators.

In addition to Qualtrics studies, several of my papers feature datasets that were collected in behavioral labs on paper or using bespoke software, shared with my team by other scholars who collected them for other projects, or exist in publicly available datasets. For those datasets (which were handled by more people), we worked hard to establish the exact provenance of the data and who had access along the way. Although some of these have a more complex history, the fact that multiple teams analyzed the same dataset makes it less likely that any one team fabricated data and "got away with it."

Finally, the exercise of tracing back the origins of ten years' worth of data has led my co-authors and I to produce a set of forward-looking recommendations for data management and authentication. Our recommendations, based on a detailed external audit of three analytically complex mixed-method projects are outlined in a working paper in preparation for *Management Science* (Yeomans, et al. 2024).

Study	Who had access to data and where can raw data be currently obtained?	Who produced analyses included in manuscript?	Are data posted on online?	Notes on the Data Source
Tulan, D., Dorsion, C.A., Gibbs, N., & Minson, J.A. (2024). Can conversational receptiveness build trust in the media. ***Policy Insights from Behavioral and Brain Sciences, in press.***			https://osf.io/2wskr/	
Study 1	Tulan, Dorison, Minson	Tulan, Dorison	Yes	Qualtrics/ online participants
Study 2	Tulan, Dorison, Minson	Tulan, Dorison	Yes	Qualtrics/ online participants
Hagmann, D., Minson, J. A., & Tinsley, C. H. (2024). Personal Narratives Build Trust in Ideological Conflict. ***Journal of Applied Psychology, in press.***			https://osf.io/sb7mj/	
Study 1	Hagman, Minson, Tinsley	Hagmann	Yes	Qualtrics/ online participants
Study 2	Hagman, Minson, Tinsley	Hagmann	Yes	Qualtrics/ online participants
Study 3	Hagman, Minson, Tinsley	Hagmann	Yes	Qualtrics/ online participants
Study 4	Hagman, Minson, Tinsley (RA: Dilan Tulan)	Hagmann	Yes	Qualtrics/ online participants
Study 5	Hagman, Minson, Tinsley	Hagmann	Yes	Qualtrics/ online participants
Collins, H. K., Minson, J. A., Kristal, A, Brooks, A.W. Conveying and detecting listening in live conversation, ***Journal of Experiment Psychology: General, 153(2),*** **473–494.**			https://osf.io/w4nf9/?view_only=db83cc05213c4964951ab88b1f9e5de0	
Exploratory Study	Collins	Collins	Yes	Qualtrics / behavioral lab participants
Experiment 1	Collins, Kristal	Collins	Yes	Qualtrics / behavioral lab participants
Experiment 2	Collins, Minson, Brooks	Collins	Yes	Qualtrics / online participants
Experiment 3	Collins, Minson, Brooks	Collins	Yes	Qualtrics / behavioral lab participants
Experiment 4	Collins, Minson, Brooks	Collins	Yes	Qualtrics / behavioral lab participants
Moore, M., Dorison, C. A., & Minson, J. A. (2023). The contingent reputational benefits of selective exposure to information. ***Journal of Experimental Psychology: General, 152*** **(12), 3490–3525.**			https://osf.io/eu56v/	
Experiment 1	Moore	Moore	Yes	Qualtrics / online participants
Experiment 2	Moore, Dorison	Moore	Yes	Qualtrics / online participants
Experiment 3	Moore, Dorison	Moore	Yes	Qualtrics / online participants
Experiment 4	Moore, Dorison	Moore	Yes	Qualtrics / online participants
Experiment 5	Moore	Moore	Yes	Qualtrics / online participants

Levin, J. M., Bukowski, L. A., Minson, J. A., & Kahn, J. M. (2023). The political polarization of COVID-19 treatments among physicians and laypeople in the United States. *Proceedings of the National Academy of Sciences, 120* (7), e2216179120.			https://github.com/pitt-healthsciences/covid_polarization/	
Study	Joel Levin, Jeremy Kahn, Leigh Bukowski (Grant Personnel: Sher Shah Amin & Billie Davis)	Levin	Yes	Data was collected from a physician panel and a sample of online participants via Qualtrics. The physician data was gathered under an NIH-funded grant and was accessible to the relevant grant personnel (listed in column B). The lay person data and various merged co-variates were only handled by Levin.
Minson, J. A. & Dorison, C. (2022). Why is exposure to opposing views aversive? Reconciling three theoretical perspectives. *Current Opinion in Psychology,* 101435.			https://researchbox.org/679	
Argument study mentioned in text	Minson, Dorison	Minson, Dorison	Yes	Qualtrics / online participants. The surveys for this paper are available in Julia Minson's account although Dorison was the original creator.
Collins, H. K., Dorison, C. A., Gino, F. & Minson, J. A. (2022) Underestimating Counterparts' Learning Goals Impairs Conflictual Conversations. *Psychological Science, 33(10),* 1732-1752.			https://researchbox.org/372	
Study 1A	Collins, Minson	Collins	Yes	Qualtrics / online participants
Study 1B	Collins	Collins	Yes	Qualtrics / online participants
Study 1C	Collins, Dorison, Minson	Collins	Yes	Qualtrics / online participants
Study 2	Collins, Dorison, Minson	Collins	Yes	Qualtrics / paid panel through marketing firm
Study 3	Collins, Minson	Collins	Yes	Qualtrics / online participants
Study 4	Collins, Dorison, Minson	Collins	Yes	Qualtrics / online participants
Study 5	Collins	Collins	Yes	Qualtrics / "snowball" Israeli sample
Dorison, C. & Minson, J. A. (2022). You can't handle the truth! Conflict counterparts over-estimate each other's feelings of self-threat. *Organizational Behavior and Human Decision Processes, 170,* 104-147.			https://researchbox.org/577	
Study 1	Collins, Dorison, Minson	Dorison	Yes	Qualtrics/ participants recruited through national market-research firm
Study 2	Dorison (Survey belongs to expired university account)	Dorison	Yes	Qualtrics / online participants
Study 3	Dorison	Dorison	Yes	Qualtrics / online participants
Study 4	Dorison (Survey belongs to expired university account)	Dorison	Yes	Qualtrics / online participants

Minson, J. A. & Umphres, C. (2020). Confidence in context: Perceived accuracy of quantitative estimates decreases with repeated trials. *Psychological Science, 31* (8), 927-943.			https://osf.io/r3z7p/	
Study 1	Minson, Umphres	Umphres	Yes	Qualtrics / online participants
Study 2	Minson, Umphres	Umphres	Yes	Qualtrics / online participants
Study 3	Minson, Umphres	Umphres	Yes	Qualtrics / online participants
Study 4	Minson, Umphres	Umphres	Yes	Qualtrics / online participants
Study 5	Minson, Umphres	Umphres	Yes	Qualtrics / online participants
Study 6	Minson, Umphres	Umphres	Yes	Qualtrics / online participants
Study 7	Minson, Umphres	Umphres	Yes	Qualtrics / online participants
Yeomans, M., Minson, J., Collins, H.K., Chen, F. & Gino, F. (2020). Conversational Receptiveness: Improving engagement with opposing views. *Organizational Behavior and Human Decision Processes, 160,* 131-148.			https://osf.io/2n59b/	
Study 1A	Yeomans, Collins, Minson, Gino (Auditors: Nick Beasley-Long, James Zhou)	Yeomans	Yes	Qualtrics / online participants
Study 1B	Yeomans, Collins, Minson (Auditors: Nick Beasley-Long, James Zhou)	Yeomans	Yes	Qualtrics / online participants
Study 2	Yeomans, Collins, Minson (Auditors: Nick Beasley-Long, James Zhou)	Yeomans	Yes	Data were collected from lab participants via the iDecision Games platform using Minson iDG account.
Study 3	Yeomans (Auditors: Nick Beasley-Long, James Zhou)	Yeomans	Yes	This study was based on a reanalysis of the wikipedia editors dataset from convoKit, which is publicly available (Zhang et al., 2017; Chang et al,. 2020)
Study 4A	Yeomans, Collins (Auditors: Nick Beasley-Long, James Zhou)	Yeomans	Yes	Qualtrics / online participants
Study 4B	Yeomans, Collins, Gino, Moore (Auditors: Nick Beasley-Long, James Zhou)	Yeomans	Yes	Qualtrics / online participants
Jeong, M., Minson, J. A., & Gino, F. (2020). In high offers I trust: The effect of first offer value on economically vulnerable behaviors. *Psychological Science, 31,* 644–653.			https://osf.io/uzncm/	
Study 1	Jeong, Minson	Jeong	Yes	Field study with data collected in Gmail
Study 2	Jeong, Minson	Jeong	Yes	Qualtrics / online participants
Study 3a	Jeong, Minson	Jeong	Yes	Qualtrics / online participants
Study 3b	Jeong, Minson	Jeong	Yes	Qualtrics / online participants
Study 4	Jeong, Minson	Jeong	Yes	Qualtrics / online participants

Minson, J. A., Chen, F. S. & Tinsley, C. H. (2020). Why won't you listen to me? Measuring receptiveness to opposing views. *Management Science, 66,* 3069-3094.			https://osf.io/xcg6e/	
Study 1A	Minson	Minson	Yes	Qualtrics / online participants
Study 1B	Minson (RA: Logan Berg)	Minson	Yes	Qualtrics / online participants
Study 2	Minson (RAs: Logan Berg, Hanne Collins, Chris Hydock)	Minson	Yes	Qualtrics / online participants
Study 3	Minson (RAs: Hanne Collins, Chris Hydock)	Minson	Yes	Qualtrics / online participants
Study 4	Minson (RAs: Hanne Collins, Chris Hydock)	Minson	Yes	Qualtrics / online participants
Study 5	Minson, Tinsley (RA: Chris Hydock)	Minson	Yes	Qualtrics / online participants
Dorison, C., Minson, J. A., Rogers, T. (2019). Selective exposure partly relies on faulty affective forecasts. *Cognition, 188* , 98-107.			https://osf.io/7mqej/	
Study 1	Minson, Dorison	Dorison	Yes	Qualtrics / online participants. The survey is available in Julia Minson's account although Dorison was the original creator. Dorison's copy is not availabel since the move from Harvard to Georgetown.
Study 2	Minson, Dorison	Dorison	Yes	Qualtrics / online participants. The survey is available in Julia Minson's account although Dorison was the original creator. Dorison's copy is not availabel since the move from Harvard to Georgetown.
Study 3	Minson, Dorison	Dorison	Yes	Qualtrics / online participants. The survey is available in Julia Minson's account although Dorison was the original creator. Dorison's copy is not availabel since the move from Harvard to Georgetown.
Study 4	Minson, Dorison	Dorison	Yes	Qualtrics / online participants. The survey is available in Julia Minson's account although Dorison was the original creator. Dorison's copy is not availabel since the move from Harvard to Georgetown.
Study 5	Minson, Dorison	Dorison	Yes	Qualtrics / online participants. The survey is available in Julia Minson's account although Dorison was the original creator. Dorison's copy is not availabel since the move from Harvard to Georgetown.

Logg, J., Minson, J. A. & Moore, D. A. (2019). Algorithm Appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes, 151*, 90-103.			https://osf.io/b4mk5/	
Experiment 1A	Logg	Logg	Yes	Qualtrics / online participants. Survey available in Logg's qualtrics account. List of collaborators unavailable because of university move.
Experiment 1B	Logg	Logg	Yes	Qualtrics / online participants
Experiment 1C	Logg	Logg	Yes	Qualtrics / online participants. Survey available in Logg's qualtrics account. List of collaborators unavailable because of university move.
Experiment 1D	Logg	Logg	Yes	Qualtrics / pool of academic researchers. Survey available in Logg's qualtrics account. List of collaborators unavailable because of university move.
Experiment 2	Logg	Logg	Yes	Qualtrics/ behavioral lab participants. Survey available in Logg's qualtrics account. List of collaborators unavailable because of university move.
Experiment 3	Logg	Logg	Yes	Qualtrics/ online participants. Survey available in Logg's qualtrics account. List of collaborators unavailable because of university move.
Experiment 4	Logg	Logg	Yes	Qualtrics/ mix of online partitipants and a pool of national security experts. Survey available in Logg's qualtrics account. List of collaborators unavailable because of university move.
Jeong, M., Minson, J. A., Yeomans, M. & Gino, F. (2019). Communicating with warmth in distributive negotiations is surprisingly counterproductive. *Management Science, 65,* 5813-5837.			https://osf.io/t7sd6/	
Study 1	Jeong, Yeomans	Jeong, Yeomans	Yes	Qualtrics/ online participants
Study 2	Jeong, Yeomans (RAs: Vanessa Ruales & Shibeal O'Flaherty)	Jeong, Yeomans	Yes	Field study with data collected in Gmail
Study 3	Minson, Jeong, Yeomans	Jeong, Yeomans	Yes	iDecision Games/ behavioral lab participants
Study 4	Jeong, Yeomans	Jeong, Yeomans	Yes	Qualtrics/ online participants

Minson, J. A., Van Epps, E., Yip, J., & Schweitzer, M. (2018). Eliciting the truth, the whole truth, and nothing but the truth: The effect of question type on deception. *Organizational Behavior and Human Decision Processes, 147*, 76-93.			https://osf.io/v56k4/	
Study 1	Van Epps	Van Epps	Yes	Qualtrics/ online participants
Study 2	Van Epps and Research Assistants at Wharton Behavioral Lab	Van Epps	Yes	Qualtrics/ behavioral lab participants
Study 3	Minson, Van Epps	Minson	Yes	Qualtrics/ behavioral lab participants
Study 4	Van Epps	Van Epps	Yes	Qualtrics/ online participants
Minson, J. A., Mueller, J. S., & Larrick, R. P. (2018). The contingent wisdom of dyads: When discussion enhances vs. undermines the accuracy of collaborative judgments. *Management Science, 64,* 4177-4192.			https://pubsonline.informs.org/doi/suppl/10.1287/mnsc.2017.2823	
Study 1	Minson	Minson	Yes	Qualtrics/ behavioral lab participants
Study 2	Minson	Minson	Yes	Qualtrics/ behavioral lab participants
Huang, K., Yeomans, M., Brooks, A.W., Minson, J. A. & Gino, F. (2017). It doesn't hurt to ask: Question-asking increases liking. *Journal of Personality and Social Psychology, 113,* 430-452.			https://osf.io/8k7rf/	
Study 1	Huang, Yeomans (Auditors: Nick Beasley-Long, James Zhou)	Huang, Yeomans	Yes	Qualtrics/ behavioral lab participants. Additional collaborators on these surveys may have been deleted during Huang's move from Harvard to Georgetown.
Study 2A	Huang, Yeomans (Auditors: Nick Beasley-Long, James Zhou)	Huang, Yeomans	Yes	Qualtrics/ online participants. Additional collaborators on these surveys may have been deleted during Huang's move from Harvard to Georgetown.
Study 2B	Huang, Yeomans (Auditors: Nick Beasley-Long, James Zhou)	Huang, Yeomans	Yes	Qualtrics/ online participants. Additional collaborators on these surveys may have been deleted during Huang's move from Harvard to Georgetown.
Study 3	Huang, Yeomans (Auditors: Nick Beasley-Long, James Zhou)	Huang, Yeomans	Yes	This study used the SpeedDate corpus, which was shared with us via email by one of the original authors, Dan Jurafsky.