

Past is future for the era of COVID-19 research in the social sciences

Dalton Conley^{a,1} and Tim Johnson^{b,1}

Over the last few decades, social scientists have experienced the causal revolution, the replication crisis, and, now in just a matter of months, another epoch: the era of coronavirus disease 2019 (COVID-19) research. According to Google Scholar, approximately 142,000 COVID-19–related articles have appeared since 2020. That amounts to about 389 articles per day, or, roughly, one article every 4 minutes. Many of these articles are in the social sciences—that is, concerned not directly with medical outcomes but rather with COVID-19’s impact on social, behavioral, and economic outcomes.

Thus far, most of this research has had a direct focus on managing COVID-19, yet a growing number of articles enlist the pandemic to study basic questions about financial investment, education, politics, learning, crime, and other aspects of social life. As COVID-19 research in the social sciences moves toward basic science, we anticipate that it will increasingly intersect with the recent scholarly trends in the social sciences: the “causal revolution,” which shifted social scientists toward research designs that could establish causal relations between study variables instead of mere correlations, and the “replication crisis,” which focused



Social scientists investigating COVID-related topics are not the first to use a historical moment to advance general knowledge about human behavior. Other one-time events, from the U.S. Vietnam-era draft lotteries to Hurricane Katrina to the 2008 financial crisis, have spawned important trends in social science research. Image credit: Dave Cutler (artist).

^aDepartment of Sociology, Princeton University, Princeton, NJ 08544; and ^bAtkinson Graduate School of Management, Willamette University, Salem, OR 97301

The authors declare no competing interest.

Published under the [PNAS license](#).

Any opinions, findings, conclusions, or recommendations expressed in this work are those of the authors and have not been endorsed by the National Academy of Sciences.

¹To whom correspondence may be addressed. Email: dconley@princeton.edu or tjohnson@willamette.edu.

Published March 24, 2021.

social scientists on ensuring the reproducibility of published findings.

Researchers who incorporate the pandemic into causal inference strategies will encounter questions about the generalizability of findings produced by rigorously studying a historically distinct event. And others may question the usefulness of conclusions that rest on analyses that do not appear to be replicable in the traditional sense of rerunning an experiment. Fortunately, past research using “natural experiments” offers answers to such questions.

Causation and Generalization

Social scientists investigating COVID-related topics are not the first to use a particular historical moment to advance general knowledge about human behavior. Previous one-time events, from the U.S. Vietnam-era draft lotteries (1–3) to Hurricane Katrina (4, 5) to the 2008 financial crisis (6, 7), have spawned important research literatures. Researchers investigate such

Careful research designs that, say, compare individuals who were subject to lock-down orders to similar individuals who were not have the potential to answer major questions about how absences from work, social isolation, and other more general phenomena induced by the pandemic affect our social, economic, and political behavior.

phenomena not merely to make their work relevant to current events but also to study potentially exogenous influences on individuals’ day-to-day lives—influences that tend to be hard to pin down.

That is, social scientists typically rely on observational data—free-range humans, so to speak. But those observational data come from a world in which everything seems to cause everything else: an endogenous system in which one cannot confidently claim that a change in one variable’s value alone influences the value of a different variable (8, 9). Randomized experiments conducted “in the field”—that is, in the non-laboratory environment in which the behavior occurs—might help. However, for many of the subjects we seek to study, social scientists cannot intervene in their research subjects’ lives with arbitrary experimental treatments (10). We cannot, for example, randomly assign some couples to remain married and some to divorce just to study the impact of marital dissolution on children. Thus, major events that create an exogenous source of variation in an otherwise endogenous system become an attractive substitute for controlled experiments.

For instance, about 50 years ago, Leslie A. Zebrowitz examined whether numbers assigned in the U.S. Vietnam-era draft lotteries influenced draft-eligible men’s sense of agency in the world (1). At the time, social psychology had gravitated toward a theory of learning that predicted the strength of behavioral

reinforcement based on whether individuals believed that their own actions, as opposed to outside forces or luck, influenced the outcomes in their lives (11). Those with a stronger sense of personal agency—or “internal” locus of control (LOC), as psychologists labeled it—would perceive causal connections between their actions and outcomes, thus bolstering learning (11). But where did this sense of control come from? One theory posited that past experiences, such as exposure to chance events, shaped locus of control. The draft lotteries provided a compelling opportunity to test that possibility.

On December 1, 1969, the first Vietnam-era draft lottery had used a chance drawing of birthdates to assign numbers that determined the rank order by which men, born from 1944 to 1950, would be assessed for military service in 1970. By comparing lottery numbers with draft-eligible men’s beliefs about their personal volition, Zebrowitz showed that chance events did appear to shape individuals’ sense of agency over their lives, at least in the short run. In the ensuing 50 years, scores of studies have examined the economic, family, health, educational, career, and intergenerational consequences of draft lottery numbers.

However, in what may be a harbinger for the era of COVID-19 research, Zebrowitz’s inaugural study received much less attention and fewer citations than later studies using the lottery design. The intuitive nature of her research design allowed it to race through the review process, but its tailoring to a specific application overshadowed the general usefulness of its novel methodology. Only when researchers highlighted the draft lottery design as a general-purpose “natural experiment” years later (2) and used it to study long-standing questions unrelated to the draft lottery itself, such as the effect of labor market absence on earnings (3), did it become a multipurpose tool for answering many questions across the social sciences because of its high-stakes, real-world implementation of a quasi-random experiment.

The era of COVID-19 research will need to make this transition, too. And, indeed, such a transition appears on the near horizon for COVID-related research. Although not as pure of an experiment as the draft lottery, the multifaceted and locally disjointed responses that occurred at different times, in different places, across different age groups (school closures, shelter-in-place orders, testing availability, and so on) provide variation that can help researchers isolate many aspects of social life that can then be studied as independent variables. In their quest to conduct basic science using the pandemic as a natural experiment, researchers have, for example, examined the “interrupted time series”—that is, break in data trends—attributable to the pandemic, noting that the pandemic changed support for institutions and disrupted patterns of crime cyclicity. Further work along these lines will ensure that COVID-related research contributes to longstanding questions in the social sciences and, in so doing, extends its impact beyond the current crisis.

Careful research designs that, say, compare individuals who were subject to lock-down orders to similar individuals who were not have the potential to answer major questions about how absences from work, social isolation, and other more general phenomena induced by the pandemic affect our social, economic, and political behavior. Moreover, the unfortunate, yet unmistakable, disparities in education and childcare brought by the pandemic ought to result in comparisons of the outcomes obtained by adjacent birth cohorts, thus answering big questions about child development and learning while also assessing the potential damage that has resulted from certain communities being unable to help children through the pandemic thanks to constraints of work, resources, and so on. If social scientists can design studies that satisfy the exogeneity (“as if random”) assumption, then the present pandemic and its myriad social responses may provide researchers much more grist for the mill of science than previous natural experiments ever did—especially because the pandemic has affected so many aspects of life and operated across so many local contexts.

A One-Time Event

However, if the COVID-19 pandemic yields mountains of basic science, can that research be replicated? That is, how can the study of a one-time event offer a solid foundation for future scholarship? Debates about the validity and reproducibility of COVID-19 research have already begun (12), and these questions, we anticipate, will figure into future discussions in the social sciences. The answer to these questions is “yes,” but perhaps not in the way we would normally imagine.

Again, early research using the Vietnam-era draft lotteries is instructive here. One potential reason why Zebrowitz’s inaugural draft lottery study might not have gained much traction is that it was subject to a near-term replication that failed to obtain the same results (13). That is, a follow-up study appeared to show no effect of the lottery results on LOC. However, such results did not necessarily refute Zebrowitz’s findings. Although the initial study might have uncovered the reported patterns by random chance (i.e., have been a false positive), it could have also been the case that the first lottery carried different effects than subsequent ones and initial effects changed across time. Perhaps once people got used to the lottery system, it no longer had the same effect. People became accustomed to the “new normal”—just as they might during a pandemic.

This is the peculiar nature of replication for the study of a one-time event such as the COVID-19 pandemic: One cannot repeat the same experiment because of the historical uniqueness of the events under study, but one may be able to find additional datasets that contain the same variables, or one may be able to study the possible effect over longer time periods. After its earliest studies on LOC, the scholarly literature on the Vietnam draft lottery took each of these approaches.

One way of replicating a study is by gauging the robustness of a finding across time. For example, one of the authors of the present work (D.C.) replicated an earlier study of the lottery’s effect on mortality (2). D.C. did not merely check the analysis of the previous study by performing the same statistical procedures on the same data; instead, he and his coauthor, Jennifer Heerwig, collected data that covered a later period and found that the original effect became less pronounced with time. Others have taken a similar tack to show that early differences in earnings attributable to draft-induced military service dissipated in later years (14), as did the effects of lottery numbers on political behavior (15) when studied with more comprehensive data (16).

Thus, replication, when studying a one-time event, takes a different form. It does not involve carefully repeating the procedures of a previous experiment or redoing analyses on the same data as a past study. Instead, it can involve finding data comparable with those of an earlier study and then analyzing them to see whether they lead to similar conclusions as did earlier analysis. Or replication can entail studying the robustness of a finding in newer data to gauge whether the effect has changed with time. This conceptualization comports with recent characterizations of replication that regard it broadly as the collection and presentation of any evidence bearing on earlier research (17).

The era of COVID-19 research will need to adopt this more general vision of replication. One cannot expect that any future event will ever be sufficiently similar to the pandemic that one could perform a “pure” replication. Instead, researchers studying the pandemic will need to consistently seek new datasets from the same period as the original data to validate the accuracy of initial research on COVID-19.

Those studies will be fascinating. For instance, will the habits of social distancing recede quickly or persist long after the pandemic wanes? Will leisure habits adopted during lockdowns influence people’s health and well-being for the long-term? Will household divisions of labor induced by school closures affect

The potential for the era of COVID-19 research to inform basic questions in the social sciences is great, but steps must be taken now to ensure that durable knowledge emerges.

gender inequality in the decades to come? Only persistent data collection, repeated analyses, and the passage of time—ingredients for the replication of a one-time event—will answer these questions.

Best Yet to Come?

The potential for the era of COVID-19 research to inform basic questions in the social sciences is great, but steps must be taken now to ensure that durable knowledge emerges. Consider, for instance, the lack of follow up on another historical event: The Negative

Income Tax Experiment (NIT). The NIT was an actual experiment conducted to see how labor supply (and family dynamics) change (or not) when people receive a guaranteed minimum income. Not only would follow-up on the NIT be theoretically interesting, results of such a follow-up would also directly inform current policy debates around Universal Basic Income. Unfortunately, NIT follow-up has been limited.

Because the NIT used social security numbers to index participants, it should be possible to follow up. However, the original records were stored in facilities subjected to flood damage, making the records unreadable. Thus, when one of the current authors (D.C.) attempted to study the NIT's long-term effects, it did not appear possible.

In contrast, researchers exploring the impact of Hurricane Katrina and resulting floods on political phenomena have benefitted from the fact that archived weather data for localities can be merged with sociopolitical data that contain geographical information (4, 5). The Vietnam lotteries used birthdates to assign draft numbers, thus any records of what happened to each individual draftee or nondraftee did not need to be kept; rather, future researchers could

merely index their new dataset to the draft via the birthdates of those born in the relevant cohorts. In either case, it was easy to create novel datasets to answer new research questions—as we ourselves have done with federal employment records (18)—as long as the salient index variables were present.

The lesson for the era of COVID-19 research is perhaps obvious: We do not yet know what long term follow-ups will be of interest, so we should preserve detailed records of what is going on now. With only modest coordination, government agencies and universities could curate the data currently being used to track cases and manage the pandemic so that future generations of researchers can use them, in ethically sound ways, to understand the pandemic's wider implications and to replicate relevant work.

Studying a one-time event in the social sciences inevitably commits a researcher's focus on the past. But if previous studies of major one-time events offer any guidance for the era of COVID-19 research in the social sciences, that commitment can generate robust, basic research insights that help us understand the present and gain insight into the future.

- 1 L. A. McArthur, Luck is alive and well in New Haven: Serendipitous finding on perceived control of reinforcement after the draft lottery. *J. Pers. Soc. Psychol.* **16**, 316–318 (1970).
- 2 N. Hearst, T. B. Newman, S. B. Hulley, Delayed effects of the military draft on mortality: A randomized natural experiment. *N. Engl. J. Med.* **314**, 620–624 (1986).
- 3 J. D. Angrist, Lifetime earnings and the Vietnam era draft lottery: Evidence from Social Security Administrative records. *Am. Econ. Rev.* **80**, 313–336 (1990).
- 4 J. Chen, Voter partisanship and the effect of distributive spending on political participation. *Am. J. Pol. Sci.* **57**, 200–217 (2013).
- 5 B. Sinclair, T. E. Hall, R. M. Alvarez, Flooding the vote: Hurricane Katrina and voter participation in New Orleans. *Am. Polit. Res.* **39**, 921–957 (2011).
- 6 Y. M. Margalit, Explaining social policy preferences: Evidence from the Great Recession. *Am. Polit. Sci. Rev.* **107**, 80–103 (2013).
- 7 Y. Margalit, Political responses to economic shocks. *Annu. Rev. Polit. Sci.* **22**, 277–295 (2019).
- 8 S. L. Morgan, C. Winship, *Counterfactuals and Causal Inference* (Cambridge University Press, Cambridge, 2nd ed., 2015).
- 9 G. W. Imbens, D. B. Rubin, *Causal Inference for Statistics, Social, and Biomedical Sciences* (Cambridge University Press, Cambridge, 2015).
- 10 R. McDermott, P. K. Hatemi, Ethics in field experimentation: A call to establish new standards to protect the public from unwanted manipulation and real harms. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 30014–30021 (2020).
- 11 J. B. Rotter, Generalized expectancies for internal versus external control of reinforcement. *Psychol. Monogr.* **80**, 1–28 (1966).
- 12 T. Gustot, Quality and reproducibility during the COVID-19 pandemic. *JHEP Reports* **2**, 100141 (2020).
- 13 K. J. Kaplan, M. Moore, Loss of freedom versus luck as determinants of externality: Failure to replicate McArthur draft lottery findings. *Represent. Res. Soc. Psychol.* **3**, 39–45 (1972).
- 14 J. D. Angrist, S. Chen, J. Song, Long-term consequences of Vietnam-era conscription: New estimates using Social Security data. *Am. Econ. Rev.* **101**, 334–338 (2011).
- 15 R. S. Erickson, L. Stoker, Caught in the draft: The effects of Vietnam draft lottery status on political attitudes. *Am. Polit. Sci. Rev.* **105**, 221–225 (2011).
- 16 T. C. Davenport, D. P. Green, K. Hanson, Are there long-term effects of the Vietnam draft on political attitudes or behavior? Apparently not. *J. Exp. Political Sci.* **6**, 71–80 (2019).
- 17 B. A. Nosek, T. M. Errington, What is replication? *PLoS Biol.* **18**, e3000691 (2020).
- 18 T. Johnson, D. Conley, Civilian public sector employment as a long-run outcome of military conscription. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 21456–21462 (2019).