

Navigating Scholarly Exchange in Today's Media Environment

D. Sunshine Hillygus, Duke University

Growing concerns about the reliability and validity of published empirical research has helped to fuel an increasing number of scholarly exchanges about research replications and reproductions, which often play out in social media, anonymous blog posts, and the media. I argue that we undermine our collective efforts to promote transparent and rigorous scientific practice if we fail to pay attention to language and communication—in our exchanges with both the media and each other.

Grimmer et al.'s (2018, in this issue) critical comment on Hajnal, Lajevardi, and Nielson (2017) and the response back by Hajnal, Kuk, and Lajevardi (2018, in this issue), is just one example in a recent spate of scholarly exchanges involving research reproductions and replications. These *Journal of Politics* publications capture the official record of the exchange, but discussion of these works also played out more widely on Twitter, Facebook, blog posts, online anonymous forums, and various media outlets—often with far more hostile and antagonistic comments than what we see in this official record. For example, in an article headlined “Media Fooled by Bunk Voter ID Study,” the *Washington Examiner* called the original study “nearly as bad” as “fake news” (Adams 2017).¹ This seems to be just the latest case in which a replication study—a seemingly standard part of the scholarly process—has turned into a firestorm of incivility, toxic commentary, and media incitement. My intent is not to wade into this particular exchange; rather, I hope it might prompt a broader discussion about the challenges of navi-

gating scientific discourse in today's media environment.² In this brief note, I argue that civil scientific critique is more difficult, but also more important, than ever before.

An essential part of the scientific process is to extend and replicate earlier findings—scientific knowledge is inherently cumulative.³ Across the social and medical sciences, there have been growing concerns about the replicability and robustness of knowledge claims made on the basis of statistically significant findings. Publication bias toward novel and significant findings is well documented, with null results typically relegated to the “file drawer” of academia (Franco, Malhotra, Simonovits 2014; Ioannidis 2005). The research decisions of individual scholars—selective subgroup analysis, underpowered studies, incomplete reporting, and so on—increase the likelihood of finding a significant result (Simmons, Nelson, and Simonsohn 2011).⁴ While the field of psychology has been ground zero for these discussions, the issues at hand are transforming research practices and norms in political science as well. For instance, we are seeing a new commitment to re-

D. Sunshine Hillygus is professor of political science and public policy at Duke University, 203 Gross Hall, Box 90204, Durham, NC 27708.

This article was solicited by the editors and went through editorial review.

1. The article posted on March 15, 2017; the working draft of comments by Grimmer, Hersh, Meredith, Mummolo, and Nall was circulated on Twitter on March 13, 2017.

2. To be clear, my comments in this note reflect a broad range of recent exchanges. The *New York Times* write-up of Amy Cuddy's research offers the most prominent example (Dominus 2017). Within political science, other examples include Fowler and Hall (forthcoming), Imai, King, and Velasco Rivera (2017), Milkis et al. (2017), and Wood and Porter (2018).

3. It is useful to distinguish *reproduction*, in which independent scholars analyze the same data as the original work, from *replication*, in which independent scholars analyze different data collected from the same data-generating process. *Generalizability* refers to the ability for the same findings to be re-created using other conditions, populations, or settings.

4. One psychology blogger creatively summarized the issues as follows: “Results sections are kind of like Instagram posts—beautiful, clear, but not necessarily accurate. Researchers can cherry-pick the best angle, filter out the splotches, and make an ordinary hot dog look scrumptious” (*Sometimes I'm Wrong*, <http://sometimesimwrong.typepad.com/wrong/2017/10/results-blind.html>; accessed December 5, 2017).

search transparency and data sharing in the field.⁵ If we can all agree that scientific knowledge depends on replication and extension, why do these exchanges seem so fraught for the original authors and the replicators?

We are in the midst of evolving standards of proof for empirical research, but it is happening without an established code of conduct for scientific discourse.⁶ Unfortunately, today's media environment encourages hostile and combative scientific exchange, just like what we see in other computer-mediated interpersonal communications.⁷ In this note, I contend that we undermine our collective efforts to promote transparent and rigorous scientific practice if we fail to pay attention to language and communication—in our exchanges with both the media and each other.

In the wake of congressional efforts to eliminate National Science Foundation funding for political science research, many in the field have made efforts to promote the societal relevance of our research by communicating to general audiences through blog posts and media engagement. Yet there are inherent tensions between media incentives and scientific practice. Research is full of uncertainty, assumptions, and caveats, whereas scientific media coverage privileges simplicity—the content needs to be easily digestible and topped with a click-bait headline. Media coverage is driven not by the methodological rigor of a study but by the novelty, drama, or political implications of the conclusions. Media-grabbing findings are more likely to be subject to replication, and failed replications (at times) offer the conflict and controversy worthy of media coverage.⁸ More generally, scientific communication is complicated by varying standards for judging research across different audiences, whether the media, legal environments, industry, or academia (including by field and subfield).

5. Admittedly, one of my frustrations with the DA-RT (data access and research transparency) initiative is that many journals have declared support for the principles of transparency but have not actually established detailed reporting standards. Despite widespread consensus about the general principles of transparency, it is more difficult (and contentious) to define exactly what that means. *Public Opinion Quarterly* and *Journal of Experimental Political Science* are rare examples of journals with explicit checklists of information that must be reported.

6. To be sure, there are other research conduct issues worthy of discussion in the field. For example, we have not established clear disclosure/conflict of interest guidelines regarding involvement in legal cases, industry partnerships, or consulting relationships. These connections can create research opportunities and strengthen research, but they can also create a potential for conflict.

7. For an overview of this literature, see Sundar (2015).

8. To be sure, there are examples of terrific science journalists, but popular media reporting often references a “debunked” study without getting into the weedy details of the actual scientific critique, including the extent to which and under what conditions or assumptions the conclusions are challenged.

What is often lost in media coverage of scientific findings is that one study—whether the original research or a replication—does not offer the final word on the topic. We rarely have definitive tests or indisputable findings in social science: research is always imbued with uncertainty. If we were to limit ourselves to research questions for which we can give an unequivocal answer, we might as well close up shop and go home. As one science writer put it, “every result is a temporary truth, one that is subject to change when someone else comes along to build, test, and analyze anew” (Aschwanden 2015). In other words, science is difficult and messy. Even following best research practices, data are still imperfect, research designs are fallible, and analyses necessitate questionable assumptions about modeling and measurement.⁹ As psychologist Sanjay Srivastava explained in the discussion of a high-profile replication failure, “a failure to replicate does not need to impeach the honesty or scientific skills of either the original researcher or the replicator. It does not even mean the effect is not real.”¹⁰

Scholars, editors, and reviewers need to do more to create a “tolerance for imperfection” (Asendorpf et al. 2013) in scientific work—even if that is at odds with media attention (or professional incentives). I would contend that the language and tone of our communications with each other are critical to creating a culture in which researchers more readily acknowledge the warts in their research. I recognize that it is easy to be dismissive of complaints about the “tone” of scientific critique. Evidence and truth are paramount in science, so it should not matter how you say something, so the argument goes.¹¹ *Chronicle of Higher Education* journalist Tom Bartlett (2014) suggests that “no amount of politeness is going to soften the revelation that a published, publicized finding is bogus. Feelings may very well get bruised, reputations tarnished, careers trashed. That is a shame, but while being nice is important, so is being right.” However, this stark perspective on right and wrong glosses over uncertainty, assumptions, and unknowns in scientific research. I am sympathetic to the perspective that social sanctioning is necessary to induce changes in research practices and norms, but we should also make sure that the professional and reputational damage actually matches the scientific missteps. There is a fundamental distinction between malfeasance and naive assumptions, misconduct and honest errors, fabrication and overgeneralizations. When failed

9. The ignorability assumption—that is, the analysis controls for all potential confounders—is a favorite to debate.

10. See *The Hardest Science*, <https://hardsci.wordpress.com/2012/03/12/some-reflections-on-the-bargh-doyen-elderly-walking-priming-brouhaha/> (accessed December 6, 2017). It is easy to point fingers at the media, but this point too often gets lost in scholarly discussions as well.

11. See, e.g., <https://www.talyarkoni.org/blog/2016/10/01/there-is-no-tone-problem-in-psychology/> (accessed December 6, 2017).

replications are mentioned in the same breath as cases of scientific fraud, we blur the lines and raise the stakes of being wrong (see, e.g., Aschwanden 2015). I worry that we are not only discouraging scholars from reconsidering their past research, we are also deterring future research agendas.

Of course, it is easy to say that scientific critique should be professional, courteous, and collaborative. Less clear is exactly how scholars should navigate that process—as either a replicator or the original author subject to criticism. As much as I would like to say that scientific critique should focus on the research question at hand rather than the author, we have to recognize that scientific criticism is inherently personal because our reputations are fundamentally linked to our research.¹² Criticism can still be civil. Indeed, incivility is more likely to trigger self-protective behaviors and responses (Antoci et al. 2016), which could ultimately be counterproductive to advancing open science.

Incivility is especially apparent in online discussions. At its best, online scientific discourse offers a unique and gratifying exchange of ideas—bringing in distinguished scholars, eager students, science journalists, and the general public—so that it simultaneously serves as public dissemination of science, actual scientific dialogue among scholars, and scientific education. Online discussions can connect scholars across universities and disciplinary boundaries, develop and strengthen networks, and communicate research norms. At the same time, the fast-paced environment of online exchange creates opportunities for snap comments, misinterpretation, and escalating insults.¹³ Such discussions can become cesspools of incivility, especially in anonymous comments, where both original authors and replicators are subject to ad hominem attacks.¹⁴ Even when not anonymous, comments can be seen as condescending and self-righteous. Psychologist Will Gervais spoofs the typical online academic exchange:¹⁵

**a purely hypothetical scenario. never happens
[weekly++]**

- Tweeter 1: Look I found a paper. Small Ns, unlikely hypothesis, $p = .04$

12. See *Approaching Significance*, <https://approachingblog.wordpress.com/2017/10/25/scientific-criticism-personal-by-nature-civil-by-choice/> (accessed December 6, 2017).

13. Social researchers attribute this to several different factors: miscommunication, reduced awareness of other's feelings, conformance to perceived norms, lack of social cues, mental defense mechanisms, etc. (see Suler 2004; Walther 2011).

14. If anyone doubts this claim, simply search any of the replication controversies on *Political Science Rumors*, <http://www.poliscirumors.com>.

15. Will Gervais's blog post on August 3, 2017: <http://willgervais.com/blog/2017/8/3/toner-vs-tactics> (accessed March 23, 2018).

- Tweeter 2: Lol, how dumb is this. Pun [funny gif]
- Tweeter 3: Did anyone even teach this people stats? [another funny gif about silly paper]
- Tweeter 4: Stuck in 2010 [facepalm gif]
- Tweeter 5: Joke about entire subfield being a failure
- Sockpuppet: Here's my patented tinfoil hat! Guaranteed to work.

...[time elapses as time inevitably does]...

all day pile-on of mockery

...[time elapses as we all march slowly to our eventual deaths]...

details into separate twitter threads about why tone really matters/is really stupid

...

See y'all same time next week!

Gervais contends that the “tone” of scholarly criticism can drive away the very scholars most needed to broadly shift research norms and practices in the field. He writes, “of course criticism is of paramount importance. It's science. But it can also occur without mockery. There are a great many people in our world who absolutely excel at leveling criticism—even withering criticism—at problematic papers and practices. Many of the most-listened-to voices in the reform movement have crushed poor papers and practices while rarely if ever getting into the mockery business.”¹⁶ The general point is that scientific critique is not just about the scholar on the other end but also about the audience. Just as when talking politics, provocation and vilification are not likely to be effective at winning over anyone other than those who already agree with you.

In a post on the *Data Colada* blog, social psychologist Uri Simonsohn outlines some specific strategies for promoting civil academic exchange (reflecting on his own communication missteps):¹⁷

- “Don't label, describe.” Simonsohn urges scholars to avoid value-laden descriptive phrases—“unwarranted conclusions” or “problematic analyses”—and to in-

16. Ibid.

17. See *Data Colada*, <http://datacolada.org/52> (accessed December 5, 2017).

stead describe the specific issues in a way that allows readers to reach their own conclusion. Evidence should be enough.

- “Don’t speculate about motives.” We cannot know why authors made particular decisions about their analyses. Our own biases can lead us to assume bad intentions, but intentions should be irrelevant to our own analyses, and these sorts of assumptions are likely to steer us toward inflammatory accusations.
- “Reach out.” Simonsohn recommends contacting the authors whose work you criticize before going public. This should be an actual invitation for feedback—not a courtesy notice—so provide the time and materials for the authors to respond. Beyond the substantive feedback, authors are the best equipped to identify unintentionally incendiary conclusions or language.¹⁸

More generally, it is important to remember that what matters is not only the intention of our words but also how they will be perceived by our counterparts and by public audiences. The interpretation of what is civil and what is uncivil is in the eyes of the receiver, not the speaker (Barash 2004). Many academic exchanges sour not over the results themselves but over the interpretation of the results and the conclusions drawn.¹⁹ Indeed, this seems to be one of the key points of contention in the exchange over the effect of voter ID laws. As a heuristic check on tone, Simonsohn suggests imagining going to dinner with the authors and their parents.

We want a disciplinary culture that encourages replication without creating an impression that only the replicators are “unassailably noble” in the pursuit of science (see Dominus 2017). Replications are pieces of scientific work—here again, mistakes can be made, assumptions are required, and the same incentives as original research are in play. This is especially the case because replications are often initially motivated by a scholar’s belief that a finding is wrong. A failed replication is also far more likely to be published than a successful replication. Quality control is just as important for replications as for original research. It is easy to criticize the peer-review process, but gatekeeping by journalists, academic blogs, and Twitter users seems to magnify incivility in a way that undermines productive scientific exchange. We also have to worry about asymmetrical participation from

scholars in such a process—communications researchers commonly refer to this as the 90-9-1 rule of social media engagement: in a given online community, 1% are heavy contributors, 9% are intermittent contributors, 90% are lurkers.²⁰ Perhaps peer review is the worst form of scientific dissemination except for all the others. Certainly, the peer review process has resulted in meaningful revisions in the exchange between Grimmer et al. (2018) and Hajnal et al. (2018).

It is exciting to see the open science movement take hold in political science. Scientific critique and online academic exchanges are fundamental to this evolution. At the same time, we have not yet sorted out the norms and expectations for engaging in the popular press and social media. I hope that journal editors, academic bloggers, and prominent scholars at the forefront of open science efforts will pay attention to the language and tone of scientific critique and will start a conversation about the code of conduct for scientific exchange.

ACKNOWLEDGMENTS

I am grateful for helpful conversations with Jason Roberts, Chris Johnston, Megan Mullin, John Patty, Maggie Penn, Irfan Nooruddin, David Siegel, Jack Zhou, and Jesse Lopez.

REFERENCES

- Adams, Becket. 2017. “Media Fooled by Bunk Voter ID Study.” *Washington Examiner*, March 15. <https://www.washingtonexaminer.com/media-fooled-by-bunk-voter-id-study/article/2617511>.
- Antoci, Angelo, Alexia Delfino, Fabio Paglieri, Fabrizio Panebianco, and Fabio Sabatini. 2016. “Civility vs. Incivility in Online Social Interactions: An Evolutionary Approach.” *PLoS ONE* 11 (11): e0164286.
- Aschwanden, Christie. 2015. “Science Isn’t Broken: It’s Just a Hell of a Lot Harder than We Give It Credit For.” *FiveThirtyEight*, August 19. <https://fivethirtyeight.com/features/science-isnt-broken/> (accessed December 6, 2017).
- Asendorpf, Jens B., et al. 2013. “Recommendations for Increasing Replicability in Psychology.” *European Journal of Personality* 27 (2): 108–19.
- Barash, David P. 2004. “Birds Do It, Bees Do It: Should Professors, Too, Strive to Communicate Sincerity?” *Chronicle of Higher Education* 50 (45): B9.
- Bartlett, Tom. 2014. “Replication Crisis in Psychology Research Turns Ugly and Odd.” *Chronicle of Higher Education*. <https://www.chronicle.com/article/Replication-Crisis-in/147301> (accessed December 7, 2017).
- Dominus, Susan. 2017. “When the Revolution Came for Amy Cuddy.” *New York Times*, October 18. <https://www.nytimes.com/2017/10/18/magazine/when-the-revolution-came-for-amy-cuddy.html> (accessed December 5, 2017).
- Flaherty, Colleen. 2017. “Peer Review’s Give-and-Take.” *Inside Higher Ed*, October 24. <https://www.insidehighered.com/news/2017/10/24/maybe-there-isnt-peer-review-crisis-least-terms-quantity>.

18. To be sure, these are not hard-and-fast rules. A nonresponsive original author should not hold up a project.

19. Thanks to Megan Mullin for this point.

20. If this rule of thumb applies, it implies greater asymmetry—and fewer gatekeepers—than what we see with the current peer review system, according to a recent survey (Flaherty 2017). There has been a hearty debate about the advantages and disadvantages of postpublication and open review.

- Fowler, A., and A. B. Hall. Forthcoming. "Do Shark Attacks Influence Presidential Elections? Reassessing a Prominent Finding on Voter Competence." *Journal of Politics*.
- Franco, A., N. Malhotra, and G. Simonovits. 2014. "Publication Bias in the Social Sciences: Unlocking the File Drawer." *Science* 345 (6203): 1502–5.
- Grimmer, Justin, Eitan Hersh, Marc Meredith, Jonathan Mummolo, and Clayton Nall. 2018. "Obstacles to Estimating Voter ID Laws' Effect on Turnout." *Journal of Politics* 80 (3): 1045–51.
- Hajnal, Zoltan, John Kuk, and Nazita Lajevardi. 2018. "We All Agree: Strict Voter ID Laws Disproportionately Burden Minorities." *Journal of Politics* 80 (3): 1052–59.
- Hajnal, Zoltan, Nazita Lajevardi, and Lindsay Nielson. 2017. "Voter Identification Laws and the Suppression of Minority Votes." *Journal of Politics* 79 (2): 363–79.
- Imai, Kosuke, Gary King, and Carlos Velasco Rivera. 2017. "Do Non-partisan Programmatic Policies Have Partisan Electoral Effects? Evidence from Two Large Scale Randomized Experiments." Working paper. <http://j.mp/2o3NZgO>.
- Ioannidis, John P. A. 2005. "Why Most Published Research Findings Are False." *PLoS Medicine* 2 (8): e124.
- Milkis, Sidney M., Julia Azari, Douglas L. Kriner, Frances E. Lee, Stephen Skowronek, William G. Howell, and Terry M. Moe. 2017. Review of *Relic: How Our Constitution Undermines Effective Government and Why We Need a More Powerful Presidency*, by William G. Howell and Terry M. Moe. *Journal of Politics* 79 (4): e78–e88.
- Simmons, Joseph P., Leif D. Nelson, and Uri Simonsohn. 2011. "False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant." *Psychological Science* 22 (11): 1359–66.
- Suler, John. 2004. "The Online Disinhibition Effect." *CyberPsychology and Behavior* 7 (3): 321–26.
- Sundar, S. Shyam. 2015. *The Handbook of the Psychology of Communication Technology*. Chichester: Wiley.
- Walther, Joseph B. 2011. "Theories of Computer-Mediated Communication and Interpersonal Relations." In Mark L. Knapp and John A. Daly, eds., *The Handbook of Interpersonal Communication*, 4th ed. London: Sage, 443–79.
- Wood, Thomas, and Ethan Porter. 2018. "The Elusive Backfire Effect: Mass Attitudes' Steadfast Factual Adherence." *Political Behavior*. doi:10.1007/s11109-018-9443-y.