


How (and Whether) to Teach Undergraduates About the Replication Crisis in Psychological Science

Teaching of Psychology
2018, Vol. 45(2) 158-163
© The Author(s) 2018
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0098628318762900
journals.sagepub.com/home/top


**William J. Chopik¹, Ryan H. Bremner², Andrew M. Defever¹,
and Victor N. Keller¹**

Abstract

Over the past 10 years, crises surrounding replication, fraud, and best practices in research methods have dominated discussions in the field of psychology. However, no research exists examining how to communicate these issues to undergraduates and what effect this has on their attitudes toward the field. We developed and validated a 1-hr lecture communicating issues surrounding the replication crisis and current recommendations to increase reproducibility. Pre- and post-lecture surveys suggest that the lecture serves as an excellent pedagogical tool. Following the lecture, students trusted psychological studies slightly less but saw greater similarities between psychology and natural science fields. We discuss challenges for instructors taking the initiative to communicate these issues to undergraduates in an evenhanded way.

Keywords

replication, undergraduate pedagogy, open-source materials, best practices, questionable research practices

During the past 10 years, crises surrounding replication, fraud, and best practices for data collection, analysis, and reporting have dominated discussions in the field of psychology. Special issues on replication, new journal policies on transparency and study design, and new methods for detecting questionable research practices have all emerged in an effort to quantify and/or alleviate concern over the reliability of psychological research (e.g., Asendorpf et al., 2013; Diener & Biswas-Diener, 2016; John, Loewenstein, & Prelec, 2012; Jonas & Cesario, 2015; Klein et al., 2014; Nosek & Bar-Anan, 2012; Open Science Collaboration, 2015; Pashler & Wagenmakers, 2012; Roberts, 2015; Schnall, Johnson, Cheung, & Donnellan, 2014; Simmons, Nelson, & Simonsohn, 2011, 2013; Spellman, 2015; Wagenmakers, 2015). Given the centrality of this debate for the future of psychological science, the attention it has received in some of our field's most prestigious outlets, and the fact that it is actively being introduced to the public through various media outlets (Achenbach, 2015; Carey, 2015; Yong, 2016), the question of why this material is not being taught to undergraduate students seemed important, puzzling, and problematic.

At national and regional conferences, in our home departments, and over e-mail, we corresponded with undergraduate instructors from both small liberal arts colleges and large research-oriented institutions about the content of their courses. We quickly discovered that not only had instructors struggled integrating these topics into their lectures, but they also were not sure where to start in trying to teach about these topics.

These reactions are not surprising given that the only suggestions to date about conveying this material to undergraduates have been to encourage instructors to involve undergraduates in replication attempts (Frank & Saxe, 2012; Grahe et al., 2012). However, the question of whether students can fully comprehend the nature of the replication crisis must be dealt with prior to addressing the practical question of whether to involve undergraduates in replication attempts. Thus, we set out to fill this void by developing an effective, up-to-date, and balanced treatment of the replication crisis with a full lecture script and prepackaged 1-hr PowerPoint presentation. We then tested its efficacy at both a large research university with very large class sizes and a mid-sized liberal arts college with very small class sizes.

Our approach was primarily to consolidate the recent history of the replication crisis as well as the arguments for increasing reproducibility suggested by experts. This approach allows instructors unfamiliar with the details of this debate to leverage their colleagues' expertise in a similar way that a textbook might. The lecture attempts a balanced treatment of experts'

¹ Department of Psychology, Michigan State University, East Lansing, MI, USA

² Department of Psychology, University of St. Thomas, St. Paul, MN, USA

Corresponding Author:

William J. Chopik, Department of Psychology, Michigan State University,
316 Physics Rd., East Lansing, MI 48824, USA.
Email: chopikwi@msu.edu

opinions as well as a distillation of statistical discussions down to a level that undergraduates can understand.

Prior to formulating the lecture, we engaged in extended conversations with undergraduates who were considering going on to PhD programs in psychology. The vast majority of them had heard of this crisis, although sometimes only filtered by journalists. They were very curious about the debate. However, because of their haphazard exposure to these discussions, their opinions tended to be quite radical, superficial, and even emotional. Some recounted that they were glad that psychology was “cleaning house” and felt that even “most” psychological results were unreliable. Others expressed anger at the originators of this discussion for even bringing up the issue and casting psychology in a negative light. Our discussions with these students led us to empirically test how learning about the replication crisis affected students’ perceptions of the field.

Selection of Lecture Material

Lecture slides, script, and a justification for topic inclusion are available (<https://osf.io/mh9pe/>). The broad, 1-hr lecture was designed to give students an introduction to the causes and consequences of the replication problem in the field of psychology, the recent attempts to quantify replication and reproducibility, and possible solutions to increasing replication and reproducibility (e.g., Asendorpf et al., 2013; Diener & Biswas-Diener, 2016; John et al., 2012; Jonas & Cesario, 2015; Klein et al., 2014; Nosek & Bar-Anan, 2012; Open Science Collaboration, 2015; Pashler & Wagenmakers, 2012; Roberts, 2015; Schnall et al., 2014; Simmons et al., 2011, 2013; Spellman, 2015; Wagenmakers, 2015).

Method

Participants

Participants were 194 undergraduate students ($M_{\text{age}} = 20.26$, $SD = 1.99$; 77.8% female) enrolled in either an Introduction to Personality course (at a large Midwestern university) or a Psychological Methods course (at a small Midwestern liberal arts college) in Spring 2016. Students reported their ethnicity as White (72.2%), Asian (12.3%), African American (6.4%), multiracial and other ethnicities (5.9%), and Hispanic/Latino (3.2%). The median year in school was 2 (a college sophomore). A large proportion of students (68%) had taken a statistics class in the past or was currently enrolled in one, whereas only 31% had heard of the replication crisis in psychology prior to the lecture. An additional 60 students did not complete one of the three parts of the study (Time 1 survey, Time 2 survey, and attending class the day of the lecture) or failed an attention check in the survey (i.e., “Please choose the third option on the scale below [i.e., #3]”). Those excluded (for whom we had demographic information) were comparable to the main analytic sample with respect to age, year in school, whether they had taken a statistics course, whether they were familiar with the replication crisis, and number of psychology classes taken.

However, these students self-reported having a lower grade point average ($d = 0.82$) were more likely to be men (45% vs. 22.8% of the analyzed sample; $\chi^2 = 5.14$, $p = .02$).

Materials and Procedure

Participants completed one survey in the week prior to the lecture and one follow-up survey in the week after. At both time points, students completed attitudinal questions about the field of psychology, read a news article for a study that had generated a lot of attention but had difficulty replicating, and answered questions about the study described in the news article. At Time 1, participants were randomly assigned to read a news article describing either the interpersonal effects of physical warmth (Lynott et al., 2014; Williams & Bargh, 2008) or how a single exposure to the American flag shifts political views 8 months later (Carter, Ferguson, & Hassin, 2011; Klein et al., 2014).

At Time 2, participants read the *other* news article (i.e., the one they had not previously seen) and answered the same questions that were asked at Time 1. Additionally, students answered general questions about the content and the clarity of the lecture.

Below, we provide an overview of the attitude and comprehension questions, but additional open-ended questions and sample examination items are available on the Open Science Framework site.

Measures

Attitudes about psychology. At both time points, six questions assessed attitudes toward the field of psychology. These items were answered on a scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). These questions can be found in the first six rows of Table 1.

News articles and study evaluation questions. At both time points, participants read a real news article about a study that had garnered media attention but fostered doubts about replicability (Carter et al., 2011; Klein et al., 2014; Lynott et al., 2014; Williams & Bargh, 2008). News articles were copied verbatim from news sources (Charles, 2008; Hochmuth, 2011). Sample sizes for each study were surreptitiously inserted into the news articles by the study investigators.

Following the news article, five multiple-choice questions about the study were presented to participants: “How believable did you find the results of this study?” “If another group of researchers were to run a similar study to the studies described on the previous page, how likely is it that this additional study would reach similar conclusions?” “How surprising were the results of this study?” “How interested are you in reading the empirical article described in the media account?” and “How would you rate the quality of the research described on the previous page?” Each question was asked on a 7-point scale, and scale anchors varied across questions. Finally, participants were asked an open-ended question, “What were your

Table 1. Evaluations of Psychology and Press Psychological Studies.

Item/Code	Before Lecture		After Lecture		d/χ^2
	<i>M</i> (%)	<i>SD</i>	<i>M</i> (%)	<i>SD</i>	
I like psychology	6.36	0.98	6.23	0.95	-0.14
I think psychological research is similar to research in fields like chemistry, physics, or biology	4.49	1.41	4.79	1.41	0.21**
I think psychological research is similar to research in fields like philosophy, literature, or modern languages	4.49	1.21	4.63	1.27	0.11
I think psychology is a “soft” science	3.63	1.52	3.65	1.56	0.01
I am interested in pursuing graduate school in psychology	4.44	2.19	4.40	2.14	-0.02
I trust the results of studies done by psychologists	5.31	0.96	4.94	1.10	-0.36***
Believability of Williams and Bargh (2008)	4.59	1.15	4.35	1.31	-0.20
Replicability of Williams and Bargh (2008)	4.59	1.07	3.96	1.38	-0.51***
Surprise of findings in Williams and Bargh (2008)	4.08	1.30	4.23	1.37	0.11
Interest in learning more about Williams and Bargh (2008)	4.51	1.56	4.64	1.49	0.09
Quality of Williams and Bargh (2008)	3.93	1.20	3.88	1.31	-0.04
Williams and Bargh (2008) was unique/interesting	45.3		47.4		0.07
Williams and Bargh (2008) was a strong study	5.8		6.4		0.03
Williams and Bargh (2008) was a weak study	16.3		14.1		0.15
Williams and Bargh (2008) was believable and made sense	33.7		30.8		0.16
Williams and Bargh (2008) was surprising/unbelievable/confusing	31.4		24.4		1.00
Williams and Bargh (2008) should be replicated or researched more	10.5		17.9		1.90
Williams and Bargh (2008) had too small a sample	7.0		19.2		5.50*
I would like to learn more about Williams and Bargh (2008)	8.1		11.5		0.54
Believability of Carter, Ferguson, and Hassin (2011)	3.85	1.22	3.43	1.29	-0.34*
Replicability of Carter et al. (2011)	3.90	1.34	3.48	1.18	-0.33*
Surprise of findings in Carter et al. (2011)	4.54	1.46	4.73	1.24	0.14
Interest in learning more about Carter et al. (2011)	3.67	1.65	3.89	1.57	0.14
Quality of Carter et al. (2011)	3.65	1.23	3.51	1.23	-0.11
Carter et al. (2011) was unique/interesting	26.0		30.5		0.40
Carter et al. (2011) was a strong study	10.4		2.4		4.26*
Carter et al. (2011) was a weak study	11.7		14.6		0.30
Carter et al. (2011) was believable and made sense	18.2		17.1		0.03
Carter et al. (2011) was surprising/unbelievable/confusing	49.4		46.3		0.14
Carter et al. (2011) should be replicated or researched more	15.6		15.9		0.002
Carter et al. (2011) had too small a sample	5.2		6.1		0.06
I would like to learn more about Carter et al. (2011)	3.9		8.5		1.45

Note. $N = 194$. Paired-sample t tests for attitudinal measures pre- and postlecture. M = mean; SD = standard deviation.

* $p < .05$. ** $p < .01$. *** $p < .001$.

impressions and thoughts about this study?,” which was later coded by the first and third authors ($\kappa = .95$).

Comprehension questions. At Time 2, 12 multiple-choice questions about the lecture were administered. Seven items asked about participants’ agreement on a 7-point scale ranging from 1 (*disagree strongly*) to 7 (*agree strongly*) and can be seen in the upper panel of Table 2. Five additional multiple-choice questions about the importance of various best practices for reproducibility were asked on a scale ranging from 1 (*not important at all*) to 5 (*extremely important*) and can be seen in the lower panel of Table 2.

Results

Attitudes About Psychology

For the most part, student attitudes toward psychology were relatively stable (see top of Table 1; a Bonferroni correction

was applied to the six comparisons made [$\alpha = .008$]). However, after the lecture, students trusted the results of studies done by psychologists less but considered psychology to be more similar to the “natural” or “hard” sciences.

Study Evaluation

We compared the evaluations of students who read a particular news article (e.g., physical warmth) before the lecture to students who read that news article after the lecture (see Table 1). Across both news articles, students thought that the findings of the studies would not replicate after hearing the lecture. Further, students were less likely to believe the findings of the flag priming study post-lecture. This pattern was also found for the believability of the physical warmth study ($d = -0.20$); however, this decline was not significant. Students did not change their other evaluations over time (i.e., the surprising nature of the findings, interest in reading the empirical paper, and the quality of the research).

Table 2. Comprehension Questions.

	Percentage at Least Slightly Agreeing (%)	M	SD
The field of psychology has problems replicating results	68.5	4.85	1.18
Replication of research is only a problem in the field of psychology	1.5	2.17	0.99
The incentive structure in psychological research can undermine the broader goals of science	63.9	4.94	1.49
The results from studies with low statistical power are by definition incorrect	22.7	3.51	1.28
Researchers who perform replication studies are not qualified to conduct psychological research	2.5	2.28	1.05
Studies that receive a lot of media attention are often reliable	2.6	2.56	1.10
It is important for a researcher to report all measures and experimental conditions that were included in a study	96.4	6.42	0.87
	Percentage at Least Very Important (%)	M	SD
For a researcher, how important is choosing a sample size before running a study?	93.8	4.45	0.61
How important is it to make data publicly available so that results can be verified by other researchers?	92.7	4.45	0.64
How important are decisions in data collection, analysis, and reporting in affecting how likely a researcher will find a significant effect?	85.6	4.24	0.72
How important is it to report studies that “don’t work out?”	82.5	4.20	0.80
How important is it that results from a psychology study are counterintuitive (e.g., different from what you would expect)?	34.5	2.93	1.26

Note. $N = 193$ – 194 . Agreement questions were answered on a 7-point scale. Importance questions were answered on a 5-point scale. Percentages reflect proportion of respondents indicating agreement (agree slightly, agree, and agree strongly) or importance (very important, extremely important) above the midpoint of each scale. A duplicate question asking “It is important for a researcher to disclose all measures and experimental conditions that were included in a study” was accidentally included in the survey. The descriptives for this item closely match the item reported above (94.3%; $M = 6.27$, $SD = 1.10$). M = mean; SD = standard deviation.

Examining the open-ended impressions of the study revealed that participants who read the physical warmth news article after the lecture mentioned that the studies had small sample sizes. Participants who read the flag priming news article after the lecture were less likely to mention that it was a strong study. Incidence of other spontaneous evaluations in the open-ended responses did not differ pre- to post-lecture.

Comprehension Questions

As seen in Table 2, the majority of students thought that psychology had problems replicating results (69%), that the incentive structure was problematic (64%), and that it is important to report all the details about a study (96%). Importantly, almost no students thought that this was a problem limited to the field of psychology (2%). Encouragingly, very few students thought that scientists performing replications were less capable of doing research (3%). Students also did not think that media attention was a sign of reliability (3%). In retrospect, the question about low-powered studies may have been confusing for students: About 23% of students thought that studies with low statistical power were by definition incorrect. We meant to convey that low-powered studies provide ambiguous information. The vast majority of students (all %s >82%) thought that it was very important to choose a sample size a priori, make data publicly accessible; that decisions in design, analysis, and reporting affect the likelihood of obtaining significant results; and that it is important to report studies that do not work out. About a third of students agreed that counterintuitive findings

are very important, reflecting the ambiguity about whether they make a meaningful contribution to the literature.

Discussion

We sought to develop a lecture that would allow instructors with only a cursory knowledge of issues related to rigor and reproducibility to convey the most important themes of these topics to their students. Developing and empirically validating a lecture about these issues should help alleviate many of the concerns instructors may have, particularly with respect to keeping up with the debate, how to translate esoteric arguments into intelligible positions, and what effects learning about these issues may have on students.

We attempted to present arguments about replication and reproducibility from an evenhanded perspective. Nevertheless, this debate is ongoing, and prominent researchers disagree about the extent of the problem, solutions to the problem, and many other issues. Thus, the lecture will continue to be updated and empirically tested. Adopting the lecture would be a way for instructors who do not follow this debate closely to stay up to date and keep their classes apprised of any developments.

The lecture was effective in conveying the most important issues about this crisis as evidenced by the comprehension questions. Of particular relevance to students being informed consumers (and future funders) of psychological research, 97% of students responded that media attention was not an accurate indicator of the reliability of a study. In addition, students demonstrated high levels of agreement with current

suggestions about transparency and reproducibility, such as determining a sample size before running a study, making data publicly available, and reporting studies that “don’t work out.” They also seemed to understand that flexible statistical decision-making can lead to questionable and significant findings. Students also correctly stated that the studies presented in the press releases may have trouble replicating, given the issues communicated in class.

It may be worrisome for some instructors to see that after the lecture, students reported trusting the studies of psychologists slightly less. However, one cannot both (a) admit that psychology has problems replicating some findings and (b) expect the trust level of individual findings to remain exactly the same. These decreases were counteracted by students’ greater appreciation for study design elements that increase reproducibility. In other words, it could be that students trust individual studies slightly less, but they can also now identify reproducible research when they see it.

Although students’ attitudes about the field experienced small changes pre- to post-lecture, learning about these issues did not change individuals’ intentions to pursue graduate study. This observation suggests that the recent discussions about replication will not cause a shortage in graduate school applicants anytime soon—at least as long as students learn about it in an evenhanded way from their instructor. Whether exposure to media reports about replication issues, be they biased or not, affects interest in attending graduate school is an open question.

Given the limited time available (~1 hr), some topics were not covered by the lecture. However, methodological advances for evaluating psychological research and conceptual approaches for enhancing reproducibility could also be integrated into other psychology courses (Finkel, Eastwick, & Reis, 2015; Gelman & Loken, 2013; Nuijten, Hartgerink, van Assen, Epskamp, & Wicherts, 2015; Simonsohn, Nelson, & Simmons, 2014; Smaldino & McElreath, 2016). The format of the lecture also allows instructors to leave out subtopics within the lecture if they believe they cannot deliver the full lecture in less than 1 hr. We also hope to investigate the effectiveness of the teaching materials using a less invasive approach than the one described above, as mere participation in a study about educational outcomes might have enhanced students’ attention toward the concepts.

Conclusion

We hope that the development of this lecture serves as a first step in integrating discussions about replication and reproducibility into contemporary undergraduate curricula. In our lecture, we covered the development and emergence of these issues while providing a template for instructors to add or subtract material as they see fit. The continued development and refinement of the lecture will give instructors an additional tool that keeps both them and their students up to date.

Author Contribution

William J. Chopik and Ryan H. Bremner contributed equally to the writing of this article.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

References

- Achenbach, J. (2015). Many scientific studies can’t be replicated. That’s a problem. *The Washington Post*. Retrieved from <https://www.washingtonpost.com/news/speaking-of-science/wp/2015/08/27/trouble-in-science-massive-effort-to-reproduce-100-experimental-results-succeeds-only-36-times/>
- Asendorpf, J. B., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J. J. A., Fiedler, K., . . . Wicherts, J. M. (2013). Recommendations for increasing replicability in psychology. *European Journal of Personality*, 27, 108–119.
- Carey, B. (2015). Many psychology findings not as strong as claimed, study says. *New York Times*. Retrieved from <http://www.nytimes.com/2015/08/28/science/many-social-science-findings-not-as-strong-as-claimed-study-says.html>
- Carter, T. J., Ferguson, M. J., & Hassin, R. R. (2011). A single exposure to the American flag shifts support toward republicanism up to 8 months later. *Psychological Science*, 22, 1011–1018.
- Charles, D. (2008). Study links warm hands, warm heart. *All Things Considered*. Retrieved from <http://www.npr.org/templates/story/story.php?storyId=96041598>
- Diener, E., & Biswas-Diener, R. (2016). The replication crisis in psychology. In E. Diener (Ed.), *NOBA Project, Psychology as a Science*. Retrieved from <http://nobaproject.com/modules/the-replication-crisis-in-psychology>
- Finkel, E. J., Eastwick, P. W., & Reis, H. T. (2015). Best research practices in psychology: Illustrating epistemological and pragmatic considerations with the case of relationship science. *Journal of Personality and Social Psychology*, 108, 275–297.
- Frank, M. C., & Saxe, R. (2012). Teaching replication. *Perspectives on Psychological Science*, 7, 600–604.
- Gelman, A., & Loken, E. (2013). *The garden of forking paths: Why multiple comparisons can be a problem, even when there is no “fishing expedition” or “p-hacking” and the research hypothesis was posited ahead of time*. Unpublished manuscript provided by Andrew Gelman.
- Grahe, J. E., Reifman, A., Hermann, A. D., Walker, M., Oleson, K. C., Nario-Redmond, M., & Wiebe, R. P. (2012). Harnessing the undiscovered resource of student research projects. *Perspectives on Psychological Science*, 7, 605–607.
- Hochmuth, C. (2011). Study: U.S. flag “primes” voters toward republican viewpoints. *Fox News Politics*. Retrieved from <http://www.foxnews.com/politics/2011/07/20/study-us-flag-primes-voters-toward-republican-viewpoints.html>

- John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, 23, 524–532.
- Jonas, K. J., & Cesario, J. (2016). How can preregistration contribute to research in our field? *Comprehensive Results in Social Psychology*, 1, 1–7.
- Klein, R. A., Ratliff, K. A., Vianello, M., Adams, R. B., Jr., Bahník, Š., Bernstein, M. J., . . . Nosek, B. A. (2014). Investigating variation in replicability. *Social Psychology*, 45, 142–152.
- Lynott, D., Corker, K. S., Wortman, J., Connell, L., Donnellan, M. B., Lucas, R. E., & O'Brien, K. (2014). Replication of "Experiencing physical warmth promotes interpersonal warmth" by Williams and Bargh (2008). *Social Psychology*, 45, 216–222.
- Nosek, B. A., & Bar-Anan, Y. (2012). Scientific Utopia: I. Opening scientific communication. *Psychological Inquiry*, 23, 217–243.
- Nuijten, M. B., Hartgerink, C. H. J., van Assen, M. A. L. M., Epskamp, S., & Wicherts, J. M. (2016). The prevalence of statistical reporting errors in psychology (1985–2013). *Behavior Research Methods*, 48, 1205–1226.
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349, aac4176.
- Pashler, H., & Wagenmakers, E. J. (2012). Editors' introduction to the special section on replicability in psychological science: A crisis of confidence? *Perspectives on Psychological Science*, 7, 528–530.
- Roberts, B. W. (2015). The perverse incentives that stand as a roadblock to scientific reform. *In-Mind Magazine*, 3. Retrieved from <http://www.in-mind.org/article/the-perverse-incentives-that-stand-as-a-roadblock-to-scientific-reform>
- Schnall, S., Johnson, D. J., Cheung, F., & Donnellan, M. B. (2014). Commentary and rejoinder on Johnson, Cheung, and Donnellan (2014a). *Social Psychology*, 45, 315–320.
- 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, 1359–1366.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2013). *Life after P-hacking*. Paper presented at the Meeting of the Society for Personality and Social Psychology, New Orleans, LA.
- Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: A key to the file-drawer. *Journal of Experimental Psychology: General*, 143, 534–547.
- Smaldino, P. E., & McElreath, R. (2016). The natural selection of bad science. *Royal Society Open Science*, 3, 1–17.
- Spellman, B. A. (2015). A short (personal) future history of revolution 2.0. *Perspectives on Psychological Science*, 10, 886–899.
- Wagenmakers, E. J. (2015). A perfect storm: The record of a revolution. *In-Mind Magazine*, 3. Retrieved from <http://www.in-mind.org/article/a-perfect-storm-the-record-of-a-revolution>
- Williams, L. E., & Bargh, J. A. (2008). Experiencing physical warmth promotes interpersonal warmth. *Science*, 322, 606–607.
- Yong, E. (2016). Psychology's replication crisis can't be wished away. *The Atlantic*. Retrieved from <http://www.theatlantic.com/science/archive/2016/03/psychologys-replication-crisis-cant-be-wished-away/472272/>