A Short (Personal) Future History of Revolution 2.0

Barbara A. Spellman
University of Virginia

Abstract

Crisis of replicability is one term that psychological scientists use for the current introspective phase we are in—I argue instead that we are going through a revolution analogous to a political revolution. Revolution 2.0 is an uprising focused on how we should be doing science now (i.e., in a 2.0 world). The precipitating events of the revolution have already been well-documented: failures to replicate, questionable research practices, fraud, etc. And the fact that none of these events is new to our field has also been well-documented. I suggest four interconnected reasons as to why this time is different: changing technology, changing demographics of researchers, limited resources, and misaligned incentives. I then describe two reasons why the revolution is more likely to catch on this time: technology (as part of the solution) and the fact that these concerns cut across social and life sciences—that is, we are not alone. Neither side in the revolution has behaved well, and each has characterized the other in extreme terms (although, of course, each has had a few extreme actors). Some suggested reforms are already taking hold (e.g., journals asking for more transparency in methods and analysis decisions; journals publishing replications) but the feared tyrannical requirements have, of course, not taken root (e.g., few journals require open data; there is no ban on exploratory analyses). Still, we have not yet made needed advances in the ways in which we accumulate, connect, and extract conclusions from our aggregated research. However, we are now ready to move forward by adopting incremental changes and by acknowledging the multiplicity of goals within psychological science.

Keywords

It has been an interesting time to be Editor of Perspectives on Psychological Science. *My term (2010 to 2015) has unexpectedly coincided with some major concerns about our science. This short personal[1] future history describes some of what I saw at the revolution and includes my predictions of what longer pieces about the history of psychology will say in the future.*

> There is a war between the ones who say there is
> a war and the ones who say there isn't.
> – Leonard Cohen ("There Is a War")

Psychological science is currently going through a major introspective stage. Some people call it a "crisis" (of confidence or of replicability), and others deny that term is applicable. I call it a revolution. It is not a revolution in the sense of the "cognitive revolution" or of a Kuhnian paradigm shift (Kuhn, 1962) because it is not about the content of our science.[2] Rather, it is about the values we

hold as we conceptualize, implement, analyze, and share our science. Because this revolution relies on creating more open interaction between people and laboratories, and because how we do our science now so heavily depends not only on individual computers but also on the Internet, I call what is currently happening *Revolution 2.0* (Spellman, 2012b, 2013a, 2013d).

I predict that any (good) future history of this revolution will not read like a history of a scientific revolution. Instead, it will read like a history of a political revolution. This article elucidates some of the factors I see as common to our Revolution 2.0 and a prototypical political revolution. The analogy is not to a political revolution like the American Revolution, involving the overthrow of an external colonial power; rather the analogy is to a

Corresponding Author:
Barbara A. Spellman, University of Virginia School of Law, 580 Massie Road, Charlottesville, VA 22903-1738
E-mails: bas6g@virginia.edu; spellman@virginia.edu

revolution like the French or Russian Revolution, a revolution that overturns the status quo within one country and leaves the same people to function in a differently structured environment.

This revolution, like all revolutions, did not begin from nothing. Like all revolutions, it has precipitating events, past harbingers, and underlying structural factors that enabled a revolution to take root now rather than in times past. Like many revolutions there has been fear, anger, and confusion; excesses and extremes; some (metaphorical) head rolling; and movement back to an acceptable middle that will, incrementally, become the new way of proceeding.

Precipitating Events

Someone must have been living in a cave to not be aware of events within psychological science (and, in fact, in all of science) in the early 2010s that might have contributed to this revolution. They have been documented and discussed in many places (see the many papers in the November 2012 special section on replicability in *Perspectives*, Pashler & Wagenmakers, 2012, and the blogs of Ed Yong, e.g., Yong, 2012), so let me just quickly note a few:

- Failures to Replicate: The recognition that many findings, especially some that were ground breaking and well-cited, failed to replicate across laboratories, in addition to the increasing frustration with the inability to publish replication failures (or even successes).
- Questionable Research Practices: The pointed illustration by Simmons, Nelson, and Simonsohn (2011) that, with enough leeway built into a study, researchers could show just about anything; that many in the field knew of people (themselves or others) who used these practices (John, Loewenstein, & Prelec, 2012); and that some practices (e.g., not reporting all variables measured) have been not only accepted but also encouraged within the field.
- Standard Statistics: Increased dissatisfaction with the use of null hypothesis significance testing (NHST), which intensified after the publication of Daryl Bem's (2011) paper on precognition in a prestigious social psychology journal.
- Open Science (and Open Access): The inability to obtain the data of other researchers for reanalysis or inclusion in meta-analyses, despite publication guidelines stating that authors should be willing to share data for such purposes. (This concern is sometimes conflated with that of wanting scientific publications to be publicly available for no fee. The first can be thought of as "open science" and the second as "open access.")

- Fraud: Some high-profile cases of fraud were galvanizing early on, particularly the case of the social psychologist Diederik Stapel, which broke in 2010. The final report about his actions (Levelt Committee, 2012) found fraud in over 50 of his publications.
- Other Fields: Although psychologists, and social psychologists in particular, seemed to be suffering from the "spotlight effect" (i.e., believing that everyone was staring at them for these mishaps), it turns out that problems of nonreplicability had been running rampant across the biological sciences and, most scarily, in medicine for years (Ioannidis, 2005).

It is important to note here that what triggers a revolution need not be what the revolution is actually about. So, for example, although I believe that the Stapel case and other revelations of fraud were important motivators for action, they will not be key to the kinds of changes that will ultimately result from this revolution.

Prehistory

It turns out, of course, that none of these calls for alarm or reform is particularly new (although some may be better articulated now). Like all political revolutions, we can look with hindsight to the unsuccessful precursors of the current movement. These events have also been documented elsewhere, particularly in the introductory sections of many recent articles, so I mention only a few here. (The table in the Appendix shows a comparison between the worries of the past and present.)

Psychologists have long been concerned about our statistical tools, especially NHST. Indeed, the debate about its value goes back to its adoption and there were loud discussions about it in the 60s and 70s. Arguments for Bayesian analysis—or at least for supplementing NHST with other statistics—have been ongoing. In more recent decades, editors of various journals have attempted to change reporting statistics, including Tony Greenwald (1976, at *Journal of Personality and Social Psychology; JPSP*), Geoff Loftus (1993a, at *Memory & Cognition*), and James Cutting (at *Psychological Science*). Greenwald also wanted authors to submit their full analyses to *JPSP* (e.g., entire analysis of variance [ANOVA] tables rather than a few selected results from a multifactor ANOVA). He also told them to expect to have their data available for sharing for at least 5 years after publication. Greenwald's initiatives lasted all of 3 years before he was relieved of his editorial duties.

Researchers have also long expressed frustration in getting others to send them data for reanalysis or meta-analyses. Concern with the positive result rate and the inability to publish negative findings (or any kinds of replications) has a long history as well. Relatedly, the "file

drawer problem" is a well-known major flaw in all meta-analysis (Rosenthal, 1979).

In 1998, Norbert Kerr publicized the term *HARKing*: hypothesizing after the results are known. He pointed out the dangers of the then-common standard practice of presenting data as if it confirmed a hypothesis that you had all along. The likely pervasiveness of this practice was documented by Bones (2012).[3]

Worries about power are not new nor are worries about questionable research practices. And, of course, as psychologists, we should remember that problems of cross-laboratory replicability have haunted us since Introspectionism during the formative years of our science.

Why Might This Time Be Different?

To understand why Revolution 2.0 began when it did, and why it might actually have lasting effects, we need to look not only at the temporal convergence of the precipitating events, but also at the status quo in psychological science circa 2010.

I believe that there are two major factors behind the current large push for change: changes in technology and changes in the demographics of psychology researchers. These two factors interact with some structural characteristics in the field—namely, limited resources and perverse incentives. I also believe that there is a hugely underrated third factor that makes this time different—recognition that we (psychologists) are not alone.

Changing technology (as part of the problem)

Technology now is quite different from the extant technology when "the rules of the game called psychological science" (Bakker, van Dijk, & Wicherts, 2012) were developed. As researchers, we now have the ability to create[4] and access more information than ever before. When I give talks about changing science, I often quiz my audience (Spellman, 2013d). I have asked hundreds of psychology researchers questions such as:

Did you ever think that you could run 100 subjects in a day? How about 1,000?

Did you ever think that you could input and analyze all your data in an hour?

Did you ever think that you could sit at your desk and collect all of the articles you want to read and cite in 1 minute?

Did you ever think that with the push of a button you could send your research to colleagues across the globe in 1 second?

For people who started in the field before the new millennium, these transformations were magical, and as they raise their hands in answer to my questions some people smile, some groan, and some sigh. Perhaps that is because the great increase in the amount and speed of research enabled by the new technologies has also created problems.

More information is a good thing—or at least most scientists think so. But increasing the speed of acquiring information can have both beneficial and harmful consequences.

One unexpected consequence of the speed of research and dissemination was more people learning about others' failures to replicate studies. Yes, people used to fail to replicate studies, but such failures were (and still often are) typically seen as a failure of the replicating researcher to properly implement the study. Discovering that other labs had also failed to replicate the study was often the result of fortuitous late-night conference conversation. With research speed came more attempts to replicate, and with dissemination speed came swifter and broader communication through e-mail, Twitter, blogs, etc. In addition, advances in data mining techniques made it simpler to study the overall research literature itself. Statistical analyses of large swaths of the literature showed, for example, the unsettling relation between effect sizes and sample sizes (e.g., Fraley & Vazire, 2014) and the implausibility of the large percentage of hypothesis-confirming studies.

Of course, another obvious consequence was that with more research came more studies for publication and, in turn, more competition to produce more articles. Ultimately, articles packaging the best stories and prettiest data were more likely to be accepted.

Changing demographics

In the last few decades, psychology has been a booming business. Psychology departments grew throughout the 1990s and 2000s (at least before the "great recession" of 2008). Psychology and psychologists have moved into business, law, and policy schools. Psychological findings have been reported in visible publications—such as the *New Yorker* (e.g., in articles by Malcolm Gladwell and Jonah Lehrer) and the expanded *New York Times* Science Times section (Clark & Illman, 2006). And psychologists themselves have written popular books about their own research.

The demographics of psychology academia have also changed. Of course, women and minorities are still underrepresented as faculty members—at least relative to their presence in the larger population and, for women, relative to their presence as graduate students. But the grip of the "old boys club" is weaker. Gone are the days

when Professor A at a top university could call Professor B at another top university and say, "I've got a great graduate student who needs a job next year." And Professor B would say, "Sure. We have a position. Tell him to pack his bags." Perhaps the opening of our field, the increasing number and diversity of researchers, and the variety of laboratories and types of training have led to less implicit trust in statements such as, "I did it and so, of course, I did it correctly." It's not that people suspect fraud so much as they want more transparency in the process from people whom they had not attended graduate school with and did not know. "So, how exactly did you do that?"[5]

And, finally, the younger generation of academics was raised on the speed and connectivity of computers and the Internet. They are used to sharing more information and doing things faster. Thus, again, overall, there are many more people trying to publish many more studies.

Limited resources

The abundance of new researchers and new research has caused problems with various types of limited resources. One such limited resource is the number of subjects in university subject pools. Research demands have outpaced the growth of student bodies. Plus, there are new pressures for more diverse participants in studies to improve generalizability.[6] Researchers have begun using more online platforms, survey services, and, of course, Amazon Mechanical Turk[7] to increase their sample sizes.

A second limited resource is grant money. There are many more researchers but funding availability has stagnated. On the one hand, with the increase of researchers wanting funding for expensive fMRI (and other brain and biological) measures, more money has gone to fewer researchers. On the other hand, with computerized testing and simpler data entry and analysis, a great deal of research has become less resource intensive in terms of both money and time.

But to many people, the most limited resource is printed pages in journals—the remnants of a 1.0 publishing processes in a 2.0 world[8] (see Priem, 2013). New journals have provided more outlets, but not enough to meet the demand (at least in the eyes of the authors), and the rejection rates remain high. Many journals have implemented triage systems. The promise of fast reviews makes the sting of rejection hurt less, but it might exacerbate the problem of too many submissions—if it is a top journal and its rejections are quick, then why not gamble and send a paper there first?

Short form empirical papers in psychology—popularized by *Psychological Science* (which began publishing in 1990)—have become more common, with the goals of speeding publication and publishing more research. But that format itself exacerbates other problems (Ledgerwood & Sherman, 2012). One is the fragmented publication of research programs (and the concept of the "minimal publishable unit"). Researchers might prefer to publish a set of experiments as two fast short publications rather than as one slower longer (more comprehensive) one. This does not help create an integrated science. Another exacerbated issue is the problem of truncated reports with key omissions. When a paper has a tight word limit, it is easy to cut method details, mentions of pilot studies or measures that "didn't work," or even references. Missing method details are a source of nonreplicability, leaving out pilot studies or measures denies readers access to the full record of what does and does not work, and cutting references does not help create an integrated science.

There is also a puzzle: Although paper journals are limited in pages, we do not live in a purely paper world any more. Why can't journals publish more and publish longer by publishing electronically? Several journals started producing online supplements, online discussions or commentaries, and even entire online publications. The uptake for online supplements was fine early on, because they were supplementing an accepted archival peer-reviewed paper publication. But, despite, or perhaps because of, the limitlessness of the resource, interest in online-only publishing is mixed, and there are many people who prefer that their thoughts appear in print.

There is a final relevant piece of the status quo to consider: Who has controlled these limited resources? The answer, for the most part, is the established folks from Generation 1.0. They are the heads of associations, the reviewers on grant panels, and the editors of journals. Why would they want to change the status quo when the status quo allowed them to succeed? (I'm not saying that age and longstanding success are the only predictors of whether someone is part of Generation 1.0 or 2.0, but I suspect huge correlations.) I got my first associate editor position at *Psychological Science* in 2002. A couple of years later, I went to a conference and ran into a well-known, well-published psychologist who had been a friend but who had recently stopped talking to me. I asked him why. He said, "No friend has ever rejected a paper of mine before." My first thought was, "What a sad view of friendship," and my second was, "What a sad comment on the state of our science."

Misaligned incentives

Another pressure point is the misaligned incentives in scientific publication. We assume that researchers want to do true, accurate, and important science and be appropriately rewarded for it. But the rewards (credit, tenure, fame) only come if the research is published, and that research is only published if it is novel, hypothesis-confirming, and the data are "pristine"

(Giner-Sorolla, 2012). There has previously been no reward for solid research that "didn't work" or for replication, and there has not been much reward for so-called "incremental research."

This situation creates "a disconnect between what is good for scientists and what is good for science" (Nosek, Spies, & Motyl, 2012, p. 616). Even without assuming any kind of intentional fraud, scientists can unwittingly become victims of motivated reasoning or hindsight bias – valuing data, methods, and analyses that support their hypotheses more than those that do not and even misremembering what had been hypothesized before learning the results (Nosek et al., 2012.) And if you know that other people are engaged in questionable research practices, wouldn't it be fair if you did it too (John et al., 2012)?

Of course, the misalignment of incentives is not simply in publishing. Because publishing is key to jobs, tenure and promotion, future grant success, and professional awards and status, some people might feel the incentives for publication to be stronger than the incentives for truth telling. The problem of misaligned incentives in science is not a new one, but given the rest of the status quo— researchers with expectations of sharing, more competition for limited jobs and publication outlets—the bar and the stakes are higher than ever. (See Engel, 2015, for a game-theoretical analysis of scientific disintegrity as a "public bad" and Diederik Stapel's autobiography for an explanation of why even a prominent psychologist might succumb to minor, then major, fraud.)

Why Might a Revolution Be Catching on This Time?

Perhaps it is availability bias that makes me think the revolution is catching on this time—I mostly talk to people who are already using or promoting changes. However, with so many organizations, journals, and granting agencies on board, it is hard to believe that it is only in my head. If something is really happening, then why?

Technology (as part of the solution)

I described earlier how technology has set the stage for the revolution. But it turns out that technology can also fix many problems (Spellman, 2013d). Most obviously, technology solves the problem of the limited resource of space—particularly publication space. Entire journals can be published online with no restrictions on the size of the articles and no space-imposed limits on how many articles can be published.

In addition to publication space, technology helps solve other problems with open science. People can post data in repositories that allow others access to it for reanalysis or use in meta-analyses. Researchers can make their hypotheses public before they run their studies and share videos of methods and working scripts for data analysis.

Of course, technology would not be part of the solution unless there were ways to use it. I believe that the (relatively) early creation of the Open Science Framework (OSF; established 2011) has provided an existence proof that preregistration, open methods, data, and analysis plans can be implemented without much trouble. Psychfiledrawer .org (established 2012; see Spellman, 2012a) has shown how unpublishable replications can be made public (even though people were often afraid to post their research). In addition, several journals with word restrictions on manuscripts have moved quickly to allow people to publish details of their work as online supplements. Some of the barriers to change are disappearing.

We are not alone

The other big reason for the current acceptance of (some) reform in psychology is the recognition that it is not only psychology that is having problems. In the earliest days of the revolution, two areas within psychology felt especially pressured. Neuroscience had been famously attacked for the analysis methods used by many researchers (Vul, Harris, Winkielman, & Pashler, 2009). Social psychology felt particularly vulnerable because of the fraud allegations, the failures to replicate, and the discussion (especially of statistics and theory) after the publication of Bem's (2011) precognition paper. But researchers have begun to realize that many of these problems are not only in neuroscience and social psychology but are also occurring in other areas of psychology and in other social and life sciences as well. (See Fiedler's, 2011, aptly titled "Voodoo Correlations are Everywhere," a play on the original initial title of Vul et al.)

Political scientists, for example, have long been worried about sharing data and replication (see, e.g., King, 1995), and their journals have responded with useful guidelines. Empirical social scientists with converging concerns have joined together to form The Berkeley Initiative for Transparency in the Social Sciences, a "network of researchers and institutions committed to improving the standards of openness and integrity in economics, political science, psychology, and related disciplines" (http://www.bitss.org/about/mission/). And in 2013, the National Science Foundation's Directorate for Social, Behavioral, and Economic Sciences established a subcommittee on Replicability in Science (National Science Foundation, 2015).

Several years ago, I suggested that psychologists propose a symposium about reforms in science at the

American Association for the Advancement of Science (AAAS) convention. A common reply was, ""Let's not air our dirty laundry in public." "But it's not just our dirty laundry, " I said, "Everyone has this problem [mentioning Ioannidis] and psychology is in a good position to suggest what to do [mentioning Nosek]]." The last few AAAS conventions have had many "new science" talks and symposia, and none were sponsored by psychology. This will change at the 2016 convention with a symposium endorsed by two sections: Social, Economic, and Political Sciences and Psychology.

Several years ago, I discovered the blog "Retraction Watch" (http://retractionwatch.com/), which documents retractions of papers in scientific journals. For a few months, they heavily covered Diederik Stapel. Soon after, I heard Ivan Oransky, one of the blog's cofounders, give a talk to a bunch of psychologists in which he showed some of the retraction notices about Stapel. A social psychologist spoke up, "Maybe we have so many retractions because we are better than other groups at uncovering fraud." Oransky assured us that Stapel, with a retraction count in the mid-two-digits, was not close to being one of the top overall fraudsters. Said the social psychologist, "Well, maybe we just have less fraud." (Is that a real life example of HARKing?)

So, other fields have fraud, and other fields have failures to replicate. Even medical clinical trials with their required preregistration have not been doing well. Publishers, federal granting agencies, and the Food and Drug Administration were concerned. The fact that it is not just a few scattered psychologists calling for reform, but, rather, a much larger movement in which psychology is facing the same challenges as other social and life sciences, is, I think, the other major factor in why this time is different and why the various changes in how we do science are catching on now whereas they had not in the past.

Taking Sides: Battle Lines and (Perceived) Extremism

As with many political revolutions, neither side has behaved well. Generation 2.0 has shouted loudly, pointed fingers, and called for major upheavals. Generation 1.0 has felt threatened by the finger pointing and by worries that psychological science would suffer. And many people have asked "What will happen to my own career if the rules were to be changed now?"

In addition, although Generation 2.0 has shouted loudly, they have not done so in unison: Different people have different concerns and see different solutions. Several stringent reforms have been vehemently proposed and, in turn, vehemently criticized as not only unwise and unfair but also as detrimental to the development of science.

Replications

A lot of energy for change came from people who, as a group, have been dubbed "The Replicators." They believe that direct replications—positive or negative—should be treated as research and have a chance to be published, be valued, and be part of the "cumulative record."

Generation 1.0 has argued against having people replicate their research and has quarreled with the idea of replications in general. Many senior scientists have stated flat out that they do not want people trying to replicate their research (Spellman, 2013b). They justify this in two ways:

1. People who replicate other peoples' research are likely bad or incompetent researchers who have no ideas of their own.
2. People will only try to replicate other peoples' research in order to show that it doesn't work.

One weakness with these arguments was exposed in a conversation I had with a very successful colleague who has said to me (more than once), "I have never tried to replicate someone else's research." In my audience quizzes, I sometimes ask whether the following story sounds familiar. "A graduate student runs into a faculty member's office and says, 'I have a great idea for a study. I just read this really interesting paper and I think that if we tried to do X rather than Y, we could show that Z.' The faculty member agrees that it's a great idea and says to the student, 'Yes, you should definitely try that. But first . . .'" Everybody in the audience knows what comes next: "First you should try to replicate the original finding." Sometimes the students try it and sometimes they fail. And sometimes they try many times and fail many times and then find themselves behind on their first year project or master's thesis and have nothing to show and nothing to publish. My colleague agreed that of course he had engaged in *that* type of replication (and, indeed, has had students who had encountered that problem). So, it was false that he had never tried to replicate someone else's study; he had tried to do so many times, but he hadn't mentally coded it as such because he didn't see himself as that kind of person.

Unsurprisingly, some people have argued that they naturally replicate their own research before they publish anything. That is certainly a good thing, but it does not necessarily mean that others will be able to replicate it (e.g., the researchers might not have reported all important methodology; the results might not be generalizable across different types of participants).

I thought that I would never understand how a scientist could say, "I don't want people trying to replicate my research." In some other fields, like chemistry and physics, a finding is not considered to be solid until someone

else has replicated it using the original methods (or so I'm told). But I started to understand the apprehension when observing the tactics of some of the data detectives.

Data detectives (or the "gotcha gang")

When suspicions of fraud arose around Diederik Stapel and others, the data detectives went to work. Looking at published means and variability can give a clue, and looking at the actual data can give a huge clue, that some kinds of fraud (or at the least, some very serious questionable research practices) are going on.

The data detectives also looked at the power, effect sizes, and patterns of p values in studies and, without stating that there must be fraud, still would suggest that something very likely must be wrong. (My former graduate student Liz Tenney said, ".04 is the new .06.") And that's when the cries of "They are on a witch hunt!" and "This is McCarthyism!" began. Many researchers fear that as soon as data detectives start looking at your paper, you are doomed because they can always find something, and it will look like you had done something wrong, even if you hadn't. (A data detective once asked me, "Well, is there something wrong with a 'witch hunt' if there really are witches?")

I said, "There should be no ex post facto laws." That is, researchers should not be punished for actions they took in the past that were permitted or standard practice when they conducted and published the research but that are deemed to be "crimes" now. Amusingly, some of those then-standard practices came from reliance on Bem's (2003) otherwise-very-nice paper called "Writing the Empirical Journal Article." One example of an unnecessary attack can be seen in the November 2012 issue of *Perspectives*. Francis (2012) critiqued a paper (Galak & Meyvis, 2011) for containing *too many* successful replications within it. Galak and Meyvis (2012) coolly replied that they did, in fact, have a file drawer of relevant unpublished studies, some more or less "successful," and that they would have been happy to share them if asked. Mentors, reviewers, and journal editors have taught us that not everything can be, or should be, shared in papers.

When I made the ex post facto argument, a member of the "gotcha gang" accused me of being against cleaning up potential errors in the literature. I responded that I was a big fan of cleaning up the literature, just not a big fan of reviling people who had followed standard practice. (But we certainly do need better ways to do the cleaning.)

Other reforms: Open science and rigid rules

There is also the perception that some of the 2.0 reforms would involve rigid rules that would damage both scientists and science. (See the nicely titled, "Psychologists are open to change yet wary of rules," Fuchs, Jenny, & Fiedler, 2012.)

Require open data. One fear-inducing proposal is that researchers' data should be open—that is, available to reviewers and readers alike. There was tremendous initial pushback against this proposal, because (I think) people believed it was motivated by the desire to catch mistakes or fraud. When the shouting died down, various other supporting reasons for open data could be heard (e.g., availability for reanalysis and meta-analysis, that a federal granting agency had paid for the research, etc.) and various opposing reasons were suggested as well (e.g., inability to deidentify confidential data, the time and money spent gathering it, etc.).

Guarantee of sample size or power. Another proposal is a requirement for studies to have a certain sample size or power. As with the requirement of open data, no one actually believed that it could possibly be applicable to every study. Yet the yelling continued for a while until the strawpeople were disemboweled and the middle ground was recognized: More power is useful and desirable but, clearly, not feasible in many (important) research situations, and such research should not be devalued. (See a nice description of that policy in Vazire, 2015.)

Require the use of X, Y, or W statistics. There is continued shouting about which types of statistics should be used and reported in our journals. There is little defense for considering only inferential tests and their p values as sufficient. But what else? Adding effect sizes or confidence intervals? Providing internal meta-analyses? Moving to Bayesian statistics? All have pros and cons; all have proponents and opponents. There will be evolving discussions, but in the meantime it would be good if the empirical journals could make it easier for authors to know what each journal wants before submission. (And, of course, consistency across journals would be good too.)

Preregistration. Generation 2.0 has also called for the preregistration of studies and hypotheses. Perhaps some people meant it for all studies and hypotheses; perhaps it was only interpreted that way. Regardless, again, there were vehement objections at the start. Researchers want to be able to learn from their data and explore it in ways that they had not considered when first implementing the study. "Where is the leeway for serendipity?" they asked. Generation 2.0 has since made it clear that no one was forbidding any particular analyses; rather, researchers just need to clearly distinguish between which were confirmatory and which were exploratory.

Types of replication. Replicators have pushed specifically for more direct rather than conceptual replications and have questioned the value of the latter (e.g., Pashler & Harris, 2012). Indeed, I've been told about an advisor who instructs graduate students, "Never do a direct replication; that way, if a conceptual replication doesn't work, you maintain plausible deniability." Generation 1.0 has argued that direct replication is close to useless and that, like preregistration and the overconcern with false positives, limiting replications to direct replications would mean the end of creativity and generativity in our science (e.g., Fiedler, Kutzner, & Krueger, 2012).

Moving Forward

I would love to be able to describe what is currently happening in the revolution at this moment. However, I can't do that because things are constantly and quickly changing (plus at *Perspectives* we still have a print lag). Every week there are new relevant publications suggesting possible changes to how we should conduct, analyze, and report our research; new technological developments that can make those possibilities a reality; and new initiatives (by journals, granting agencies, and other organizations) intended to improve science. For example, in the last month, the Open Science Collaboration (2015) published its "Reproducibility Project," describing the results of the replication of 100 social and cognition psychology experiments. Diederik Stapel's retraction count went up to 55. Psychfiledrawer .org reached half a million views. The incoming editor of *Social Psychological and Personality Science* announced changes to the manuscript evaluation processes (Vazire, 2015). Several new preconferences and conferences on these topics have been planned. And dozens of blogs have been written, posted, and (often heatedly) commented on.

This continual change is why it is premature to write anything other than a "future history." It is, indeed, possible that what is happening now, like so many calls for change in the past (see Table 1), are just tiny bumps in the road that will be ignored and forgotten and not affect future science. But I suspect that some things are here to stay even if, now, they are only in the earliest stages of development.

Most important, it seems that the screaming has died down a bit, although it certainly has not stopped.[9] The reasons for this might be good (more understanding) or bad (people are dropping out of the conversation), or, again, it might be simply a matter of what I hear. But I think it has happened, and I think that at least some of the reduction is probably because some people recognize that the extreme caricature of what was viewed as the "opposition" is not correct.

So maybe Generation 1.0 sees that 2.0's tactics are not all about finding fraud. They acknowledge that direct replication can be valuable but still think that 2.0 is obsessed with false positives. They worry that the new science could inhibit creativity and new discoveries. But some people in 1.0 are embracing the idea that, in general, finding ways to make our studies more replicable is a useful goal. Generation 2.0 has certainly learned that running and interpreting direct replications (e.g., knowing what counts as a successful replication) is not as simple as initially thought, even when the previous authors are on board (Open Science Collaboration, 2015; Simons, Holcombe, & Spellman, 2014). I think that we have learned much from these endeavors but that it is time for the massive general programs of direct replications to stop and for us to be more selective in investing our ideas and resources. Finally, what counts as the right statistics to use is still a bone of contention, even within Generation 2.0. So certainly neither side has everything figured out.

It is also more apparent now that various areas of psychology, other disciplines, many journal and organizations, and even some grant agencies are on board with some of the changes to the status quo. (See the Table in Spellman, 2015, which lists the special sections in *Perspectives* related to methods over the last few years, showing how the discussion, or at least the submission and acceptance content, has changed.)

The most obvious changes have been to the journals. More of them are allowing, encouraging, requiring or even rewarding (e.g., Eich, 2014; Spellman, 2013c; Vazire, 2015) such things as:

- full descriptions of methods (giving unlimited words or the ability to post videos of the studies);
- making data sets and analysis code available;
- preregistration of studies (e.g., guaranteeing publication if certain criteria are fulfilled); and
- registered replications (e.g., for *Perspectives*, see Simons et al., 2014).

Especially interesting is that the predicted doomsday of all these things being required for every paper has not arrived. Some journals have required some changes to statistical presentations and descriptions. But dozens of journals in psychology and other sciences (including *Science*) have signed onto the Transparency and Openness guidelines, a framework for developing publication practices that would incrementally increase openness and (hopefully) consequently replicability in science (Nosek et al., 2015). The guidelines consider such practices as open data, materials, and analysis; preregistration of studies and analysis plans; replication; and citation standards for borrowed data, materials, and code. The guidelines provide suggestions not requirements, and different fields and different journals are free to proceed at their own pace in adopting none, some, or all of them.

What Is NOT Happening?

Although much is happening, there are still a few things that I hoped would happen in this revolution that have not yet occurred, though perhaps they will soon. Many of the changes I see address improving individual empirical studies—what I call "making better bricks." But our scientific endeavor is not simply about the bricks; it is about creating better walls and buildings and palaces. That is, we need better ways to accumulate, connect, and extract conclusions from our aggregated research. And we need to be better and faster at self-correction (Ioannidis, 2012).

I have said all of this in different places (e.g., Spellman, 2010, 2012a, 2012c, 2013e), so let me just touch on a few aspects of it here.

1. We need to stop losing important information. The lessening of restrictions on the length of some method sections is a start, as is being able to find and view (some) unpublished replications—we need to be able find them and value them appropriately. We still do not know enough about the unpublished variations (e.g., "pilot studies") of published studies. We need to know more about what does not work in addition to what does. This information can help us to find support for new theories, refine our current theories, and devalue or discard less-supported theories (something we are not at all good at; Ferguson & Heene, 2012; Greenwald, 2012). Progress in science needs all of those things.

2. Relatedly, we need to be better at compiling our results. We should have better ways to call for papers to include in meta-analyses and more systematic ways to save the work that goes into them (see Braver, Thoemmes, & Rosenthal, 2014). We should create sharable (and citable) databases for related studies in such a way that people can add to and modify them and then analyze them in different ways as the database grows. This is a familiar practice in some other fields (e.g., The Supreme Court database in political science).

3. We need to better connect our findings across areas and subareas and sub-subareas of our field. Our keyword system has become worthless, and we now rely too much on literal word searches that do not find similar (or analogous) research if the same terms are not used to describe it (see, e.g., Ranganath, Spellman, & Joy-Gaba, 2010). We should also keep track of the purpose of our citations. We cite other papers for many reasons—for example, general background, use of methods, and consistency with findings or theories. We could easily keep track of the reasons for citing, which would both simplify our research and help us understand whether some findings are being confirmed, expanded, limited, or disconfirmed in subsequent studies (similar to what is done for legal opinions; Anicich, 2014; Spellman, 2012c).

What Will the Future Look Like?[10]

After the revolution, we will come to a sensible middle ground. Standard practices may change, but we will recognize that those new default rules are certainly not applicable to all of our science all of the time. As Generation 2.0 fills in leadership positions we will see continuing evolution not revolution—at least for a while.

Individually, we will become more like academics in other fields. Like chemists and physicists, we will understand that it is not an insult when others try to replicate our research—it is standard science. And, indeed, we should be flattered that they think it is worth spending their time to do so. Like philosophers, we will understand that when people take issue with our work it is not personal, it is about the ideas. And, like . . . well, like good psychologists, we will recognize that with regard to our own research, we ourselves are all subject to those biases that some of our psychology colleagues study. Our training of undergraduates and graduate students will incorporate this knowledge and these values plus, I hope, a bit more history and philosophy of science than it has in the past. (A set of syllabuses for such classes is available on the Open Science Framework, 2015.)

Most important, I think that we will recognize that psychological science has a multiplicity of goals and that those goals must be sought in different ways. We want to understand how minds work and we want to understand how to apply what we know in the real world: It is likely that some subtle and difficult-to-replicate phenomena might be existence proofs that tell us something about the first; repeating the research and looking for moderators and mediators of the effects may help us with the second. We will value both data and theory. We will value both confirmation and exploration. We will realize that we have already picked a lot of the low-hanging fruit and that we can investigate new levels of complexity with our new methods for data collection (from brain imaging to wearable devices) and analysis. We will build better bricks and better buildings. And we will, carefully and seriously, be involved in finding ways to apply what we know to important issues and policies (see Teachman et al., 2015, this issue). After we stop shouting, we will (finally) have the seat at the table that our science deserves.

[It has been an interesting time to be Editor of *Perspectives*. But, then again, it should always be an interesting time to be engaged in science.]

Appendix

Comparing Past and Present Efforts to Change Psychological Science

Table A1. A comparison of past and present efforts to change psychological science. Note that this version of the table is a "draft" and is constantly being updated; see https://docs.google.com/document/d/1lmnYIcavpXjXo2GA2m7kytKdnZxJPoXVWLoYauxFu5s/edit

Issue	Older statements (pubs)—early and/or prominent	Older efforts	Lasting consequences	Newer statements (pubs, blogs, accounts)—early and/or prominent	Newer efforts
Crisis of replicability	Greenwald (1976)			Ioannidis (2005); Yong (2012)	
Positive result rate	Sterling (1959, 1995)	A few "journals of negative results" established		Fanelli (2010, 2011)	
File drawer problem	Rosenthal (1979)		Development of statistics to account for it in meta-analyses	This has been a continuous concern; Spellman (2010)	PsychFileDrawer.com, other general repositories (see efforts to get them published)
(Not) publishing replications		1970s *Replications in Social Psychology* (two issues)		Makel, Plucker, & Hegarty (2012)	*Cortex*; *Social Psychology* (special issue; Nosek & Lakens, 2014); *Perspectives* (RRRs); *Judgment & Decision Making*; others
Distinguishing confirmatory from exploratory analyses; preregistration	Kerr (1998)	Development of data techniques in SEQ modeling	Clinical trials.gov (required beginning 2007)	Wagenmakers, Wetzels, Borsboom, van der Maas, & Kievit, 2012	The American Psychological Association publication manual (6th ed., 2009) asks for all results; mentioned in TOP guidelines
Questionable research practices	Mills (1993)			Simmons, Nelson, & Simonsohn (2011); John, Loewenstein, & Prelec (2012)	
Power	Cohen (1962); Sedlmeier & Gigerenzer (1989)				Schimmack (2012)
Reporting full methods	Greenwald (1976)	Greenwald (provide full description with journal submission)	Movement to short reports works counter to this goal.		Many journals allow online supplements; *Psychological Science* (badges); mentioned in TOP guidelines

(Continued)

Table A1. *(Continued)*

Issue	Older statements (pubs)— early and/or prominent	Older efforts	Newer statements (pubs, blogs, accounts)— early and/or prominent	Lasting consequences	Newer efforts
Open data	Loftus (1993b)	Greenwald and Loftus ask that it be made available.			*Judgment and Decision Making* (the Editor, Jonathan Baron, requires it); Open Science Framework and other repositories; *Psych Science* (badges): mentioned in TOP guidelines
NHST Critiques	Going back to the beginning and big again in the 1960s (e.g., Meehl, 1967; Lykken, 1968; Rozeboom, 1960)			Continuing critiques Cohen (1994).	
Supplementing NHST	Loftus (1993a)	Loftus (1993b) asks for more figures with error bars, CIs, no NHST	Cumming (2014)	American Psychological Association publication manual (5th ed., 2001) author guidelines recommend effect sizes and CIs (see Finch, Thomason, & Cumming, 2002)	*Psych Science*; Cumming (2014) and Cumming's APS tutorial videos (http://www .psychologicalscience.org/ index.php/members/new- statistics)

Note: TOP = Transparency and Openness Promotion; NHST = null hypothesis significance testing; CI = confidence interval.

Acknowledgments

Declaration of Conflicting Interests

Notes

1. Because this piece was intended to be short, I have kept the citations minimal; because it is personal, many citations are to articles published in *Perspectives*.

2. See Sturm and Mülberger (2012) for a description of the claims of "crisis" in psychology going back over 100 years.

3. This spoof by the pseudonymous Arina K. Bones accentuates the point documented by Fanelli (2010, 2011)

4. Technology also makes it easier to create fake data with desired characteristics. Of course, it also makes it easier to expose fake data.

5. More than one person has pointed out an analogy to me here: As a society, we are skeptical of politicians or business-people who say "I can do X, trust me" with no details or explanation. Should it be different with scientists?

6. The article on "WEIRD" subjects by Henrich, Heine, and Norenzayan (2010) has been a hugely important recent influence on research practices apart from the revolution.

7. It is likely that an article published in *Perspectives* (Buhrmeister, Kwang, & Gosling, 2011) contributed substantially to the increase in researchers' use of MTurk.

8. This idea comes from a talk by Jason Priem during the 2012 academic year at Columbia University.

9. A few weeks ago, someone on Facebook asked, "Remember back when there was name calling?" Yes, I do. Whatever day today is, the answer is "Yesterday." I'm not convinced that this behavior will end soon.

10. I don't study precognition. Perhaps we will end up in the scientific utopia described by Brian Nosek and colleagues in two of their three planned papers: the first on scientific communication (Nosek & Bar-Anan, 2012) and the second on restructuring incentives in publishing (Nosek et al., 2012). The third will focus on crowdsourcing science.

References

American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: Author.

American Psychological Association. (2009). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.

Anicich, E. M. (2014). What lies within: Superscripting references to reveal research trends. *Perspectives on Psychological Science, 9,* 682–691.

Bakker, M., van Dijk, A., & Wicherts, J. M. (2012). The rules of the game called psychological science. *Perspectives on Psychological Science, 7,* 543–554.

Bem, D. J. (2003). Writing the empirical journal article. In J. M. Darley, M. P. Zanna, & H. L. Roediger III (Eds.), *The compleat academic: A practical guide for the beginning social scientist* (2nd ed., pp. 171–201). Washington, DC: American Psychological Association.

Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology, 100,* 407–425.

Bones, A. K. (2012). We knew the future all along: Scientific hypothesizing is much more accurate than other forms of precognition. *Perspectives on Psychological Science, 7,* 307–309.

Braver, S. L., Thoemmes, F. J., & Rosenthal, R. (2014). Continuously cumulating meta-analysis and replicability. *Perspectives on Psychological Science, 9,* 333–342.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk a new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6,* 3–5.

Clark, F., & Illman, D. L. (2006). A longitudinal study of the New York Times Science Times Section. *Science Communication, 27,* 496–513. doi:10.1177/1075547006288010

Cohen, J. (1962). The statistical power of abnormal-social psychological research: A review. *Journal of Abnormal and Social Psychology, 65,* 145–153.

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist, 49,* 997–1003.

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science, 25,* 7–29.

Eich, E. (2014). Business not as usual. *Psychological Science, 25,* 3–6.

Engel, C. (2015). Scientific disintegrity as a public bad. *Perspectives on Psychological Science, 10,* 361–379.

Fanelli, D. (2010). "Positive" results increase down the Hierarchy of the Sciences. *PLoS ONE, 5*(3). doi:10.1371/journal.pone.0010068

Fanelli, D. (2011). Negative results are disappearing from most disciplines and countries. *Scientometrics, 90,* 891–904.

Ferguson, C. J., & Heene, M. (2012). A vast graveyard of undead theories: Publication bias and psychological science's aversion to the null. *Perspectives on Psychological Science, 7,* 555–561.

Fiedler, K. (2011). Voodoo correlations are everywhere—Not only in neuroscience. *Perspectives on Psychological Science, 6,* 163–171.

Fiedler, K., Kutzner, F., & Krueger, J. I. (2012). The long way from α-error control to validity proper: Problems with a short-sighted false-positive debate. *Perspectives on Psychological Science, 7,* 661–669.

Finch, S., Thomason, N., & Cumming, G. (2002). Past and future American Psychological Association guidelines for statistical practice. *Theory & Psychology, 12,* 825–853.

Fraley, R. C., & Vazire, S. (2014). The *N*-pact factor: Evaluating the quality of empirical journals with respect to sample

size and statistical power. *PLoS ONE, 9*(10), e109119. doi:10.1371/journal.pone.0109019

Francis, G. (2012). The psychology of replication and replication in psychology. *Perspectives on Psychological Science, 7*, 585–594.

Fuchs, H. M., Jenny, M., & Fiedler, S. (2012). Psychologists are open to change, yet wary of rules. *Perspectives on Psychological Science, 7*, 639–642.

Galak, J., & Meyvis, T. (2011). The pain was greater if it will happen again: The effect of anticipated continuation on retrospective discomfort. *Journal of Experimental Psychology: General, 140*, 63–75.

Galak, J., & Meyvis, T. (2012). You could have just asked: Reply to Francis (2012). *Perspectives on Psychological Science, 7*, 595–596.

Giner-Sorolla, R. (2012). Science or art? How aesthetic standards grease the way through the publication bottleneck but undermine science. *Perspectives on Psychological Science, 7*, 562–571.

Greenwald, A. G. (1976). An editorial. *Journal of Personality and Social Psychology, 33*, 1–7.

Greenwald, A. G. (2012). There is nothing so theoretical as a good method. *Perspectives on Psychological Science, 7*, 99–108.

Greenwald, A. G., Gonzalez, R., Harris, R. J., & Guthrie, D. (1996). Effect sizes and p-values: What should be reported and what should be replicated? *Psychophysiology, 33*, 175–183.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral & Brain Sciences, 33*, 61–83.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine, 2*(8), e124. doi:10.1371/journal.pmed.0020124

Ioannidis, J. P. A. (2012). Why science is not necessarily self-correcting. *Perspectives on Psychological Science, 7*, 645–654.

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. doi:10.1177/0956797611430953

Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review, 2*, 196–217.

King, G. (1995). Replication, replication. *PS: Political Science & Politics, 28*, 444–452.

Kuhn, T. (1962). *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press.

Ledgerwood, A., & Sherman, J. W. (2012). Short, sweet, and problematic? The rise of the short report in psychological science. *Perspectives on Psychological Science, 7*, 60–66.

Levelt Committee. (2012, November 28). *Flawed science: The fraudulent research practices of social psychologist Diederik Stapel*. Retrieved from http://pubman.mpdl.mpg.de/pubman/item/escidoc:1569964/component/escidoc:1569967/Eindrapport_Commissie_Levelt.pdf

Loftus, G. R. (1993a). Editorial comment. *Memory & Cognition, 21*, 103.

Loftus, G. R. (1993b). A picture is worth a thousand *p* values: On the irrelevance of hypothesis testing in the microcomputer age. *Behavior Research Methods, Instruments, & Computers, 25*, 250–256.

Lykken, D. T. (1968). Statistical significance in psychological research. *Psychological Bulletin, 70*(3, Pt. 1), 151–159.

Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in psychology research: How often do they really occur? *Perspectives on Psychological Science, 7*, 537–542. doi:10.1177/1745691612460688

Meehl, P. E. (1967). Theory-testing in psychology and physics: A methodological paradox. *Philosophy of Science*, 103–115.

Mills, J. L. (1993). Data torturing. *The New England Journal of Medicine, 329*, 1196.

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., . . . Yarkoni, T. (2015). Promoting an open research culture. *Science, 348*, 1422–1425. doi:10.1126/science.aab2374

Nosek, B. A., & Bar-Anan, Y. (2012). Scientific utopia: I. Opening scientific communication. *Psychological Inquiry, 23*, 217–243.

Nosek, B. A., & Lakens, D. (Eds.). (2014). Special issue: Replications of important results in social psychology. *Social Psychology, 45*, 137–252.

Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science, 7*, 615–631.

National Science Foundation. (2015). *Social, behavioral, and economic sciences perspectives on robust and reliable science: Report of the subcommittee on replicability in Science Advisory Committee to the National Science Foundation Directorate for Social, Behavioral, and Economic Sciences*. Retrieved from http://www.nsf.gov/sbe/AC_Materials/SBE_Robust_and_Reliable_Research_Report.pdf

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*, 6251. doi:10.1126/science.aac4716

Open Science Framework. (2015). *Course syllabi for open and reproducible methods*. Retrieved from https://osf.io/vkhbt/

Pashler, H., & Harris, C. (2012). Is the replicability crisis overblown? Three arguments examined. *Perspectives on Psychological Science, 7*, 531–536.

Pashler, H., & Wagenmakers, E.-J. (Eds.). (2012). Special section on replicability in psychological science: A crisis in confidence? *Perspectives on Psychological Science, 7*, 528–687.

Priem, J. (2013, March 28). Scholarship: Beyond the paper. *Nature, 495*, 437–440. doi:10.1038/495437a

Ranganath, K. A., Spellman, B. A., & Joy-Gaba, J. A. (2010). Cognitive "category-based induction" research and social "persuasion" research are each about what makes arguments believable: A tale of two literatures. *Perspectives on Psychological Science, 5*, 115–122.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin, 86*, 638–641.

Rozeboom, W. W. (1960). The fallacy of the null-hypothesis significance test. *Psychological Bulletin, 57*, 416–428.

Schimmack, U. (2012). The ironic effect of significant results on the credibility of multiple-study articles. *Psychological Methods, 17*, 551–566. doi:10.1037/a0029487

Sedlmeier, P., & Gigerenzer, G. (1989). Do studies of statistical power have an effect on the power of studies? *Psychological Bulletin, 105*, 309–316.

Simmons, J. P., Nelson, L. N., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data

collection and analysis allows presenting anything as significant. *Psychological Science, 22*, 1359–1366.

Simons, D. J., Holcombe, A. O., & Spellman, B. A. (2014). An introduction to registered replication reports at *Perspectives on Psychological Science. Perspectives on Psychological Science, 9*, 552–555. doi:10.1177/1745691614543974

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.

Spellman, B. A. (2010). Letter from the incoming editor. *Perspectives on Psychological Science, 5*, 3–4. doi:10.1177/1745691609356780

Spellman, B. A. (2012a). Introduction to the special section: Data, data, everywhere . . . Especially in my file drawer. *Perspectives on Psychological Science, 7*, 58–59. doi:10.1177/1745691611432124

Spellman, B. A. (2012b, December 8). *Research Revolution 1.0— Let's move on*. Retrieved from https://morepops.wordpress.com/2012/12/08/research-revolution-1-0-lets-move-on/

Spellman, B. A. (2012c). Scientific utopia . . . or too much information? Comment on Nosek and Bar-Anan. *Psychological Inquiry, 23*, 303–304.

Spellman, B. A. (2013a, August 26). *Barcelona—and should Revolution 2.0 go grassroots?* Retrieved from https://morepops.wordpress.com/2013/08/26/barcelona-and-should-revolution-2-0-go-grassroots/

Spellman, B. A. (2013b, March 5). *But I don't want people to try to replicate my research*. Retrieved from https://morepops.wordpress.com/2013/03/05/but-i-dont-want-people-to-try-to-replicate-my-research/

Spellman, B. A. (2013c, February 27). *My first post-revolution action letter*. Retrieved from https://morepops.wordpress.com/2013/02/27/my-first-post-revolution-action-letter/

Spellman, B. A. (2013d, May 27). *Research Revolution 2.0— The Current "crisis": How technology got us into this mess and how technology will help us out*. Retrieved from https://morepops.wordpress.com/2013/05/27/research-revolution-2-0-the-current-crisis-how-technology-got-us-into-this-mess-and-how-technology-will-help-us-out/

Spellman, B. A. (2013e). There is no such thing as replication but we should do it anyway. Comment on "Recommendations for increasing replicability in psychology." *European Journal of Personality, 27*, 120–144. doi:10.1002/per.1920

Spellman, B. A. (2015). Introduction to the special section on methods: Odds and end. *Perspectives on Psychological Science, 10*, 359–360.

Sterling, T. D. (1959). Publication decisions and their possible effects on inferences drawn from tests of significance—Or vice versa. *Journal of the American Statistical Association, 54*, 30–34.

Sturm, T., & Mülberger, A. (2012). Crisis discussions in psychology—New historical and philosophical perspectives. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 43*, 425–433.

Teachman, B. A., Norton, M. I., & Spellman, B. A. (2015). Memos to the president from a "council of psychological science advisers." *Perspectives on Psychological Science, 10*, 697–700.

Vazire, S. (2015). Editorial. *Social Psychological & Personality Science*. Advance online publication. doi:10.1177/1948550615603955

Vul, E., Harris, C., Winkielman, P., & Pashler, H. (2009). Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspectives on Psychological Science, 4*, 274–290.

Wagenmakers, E. J., Wetzels, R., Borsboom, D., van der Maas, H. L., & Kievit, R. A. (2012). An agenda for purely confirmatory research. *Perspectives on Psychological Science, 7*, 632–638.

Yong, E. (2012, May 16). Replication studies: Bad copy. *Nature*. Retrieved from http://www.nature.com/news/replication-studies-bad-copy-1.10634