"Everybody Knows Psychology Is Not a Real Science"

Public Perceptions of Psychology and How We Can Improve Our Relationship With Policymakers, the Scientific Community, and the General Public

Christopher J. Ferguson
Stetson University

AQ: au
AQ: 1

In a recent seminal article, Lilienfeld (2012) argued that psychological science is experiencing a public perception problem that has been caused by both public misconceptions about psychology, as well as the psychological science community's failure to distinguish itself from pop psychology and questionable therapeutic practices. Lilienfeld's analysis is an important and cogent synopsis of external problems that have limited psychological science's penetration into public knowledge. The current article expands upon this by examining internal problems, or problems within psychological science that have potentially limited its impact with policymakers, other scientists, and the public. These problems range from the replication crisis and defensive reactions to it, overuse of politicized policy statements by professional advocacy groups such as the American Psychological Association (APA), and continued overreliance on mechanistic models of human behavior. It is concluded that considerable problems arise from psychological science's tendency to overcommunicate mechanistic concepts based on weak and often unreplicated (or unreplicable) data that do not resonate with the everyday experiences of the general public or the rigor of other scholarly fields. It is argued that a way forward can be seen by, on one hand, improving the rigor and transparency of psychological science, and making theoretical innovations that better acknowledge the complexities of the human experience.

Keywords: psychological science, replication, bias, public perceptions

That psychological science has struggled to find its voice among policymakers, other scientists, and the general public is an issue that is difficult to ignore. Those of us who are involved in psychological science deeply believe in the scientific value of the study of human behavior. But, on a fundamental level, many within other fields of scholarship, as well as the general public, still question whether psychology is a science at all. The question of whether psychology is a "real" science is not just questioned among the masses, but has been debated in the *New York Times* (Gutting, 2012), *L.A. Times* (Berezow,

2012), and similar outlets and may have contributed to reduced government funding of social science research (Stratford, 2014). Concerns about psychology's status as a science come not only from among policymakers and the general public, but also from other areas of science (e.g., see Yong, 2012, and Wai, 2012, for several examples). Psychology's status as a science is clearly not accepted as a *given* either among policymakers, the general public, or other scientists.

The reasons for this are manifold, ranging from misperceptions of psychology among the general public, the continued popularity of psychological myths, the visibility of therapeutic techniques with poor empirical support, and lack of clear delineation between pop psychology and research psychology (Lilienfeld, 2012; Lilienfeld, Lynn, Ruscio, & Beyerstein, 2009). These are all external problems, although they might be exacerbated by issues such as psychological researchers' reluctance to engage the public (Lilienfeld, 2012; Weeks, 2005).

In the current article, however, I discuss a slightly different issue, namely that some of the problem may be self-inflicted from within psychological science itself, with issues ranging from the replication crisis to an *overselling* of psychological science, particularly on morally or culturally valenced issues to a mechanistic world-view that does not speak to the complexities of the human condition. Although a difficult topic to consider, I argue that psychological science, both theoretically and methodologically, is not quite ready for prime time.

At the outset, however, I wish to be clear that although my observations may be critical, they are offered with a sense of optimism. Our field has much to offer, and the desire among policymakers and the general public to understand human behavior has never been greater. Reflec-

**Christopher J.
Ferguson**

tion, criticism, and remaking are part of science. With a hard look at our field, I sincerely believe we can work together to improve and reach our full potential as a scientific discipline.

I note upfront that factors that influence policymakers, other scientists, and the general public as groups may vary. For instance, it is doubtful that methodological issues such as the replication crisis or psychological science's aversion to the null are kitchen-table topics in the average household. However, such issues undoubtedly influence the opinions of other scientists who may view our field as less rigorous than hard-science fields such as physics, chemistry, or biology. These opinions of other scholars may be transmitted both to policymakers and the general public through public comments, written editorials, and academic lectures to students. Furthermore, psychology's internal problems (e.g., the replication crisis, problematic policy statements) have received remarkable press coverage in recent years (e.g., Berezow, 2012; Chambers, 2014; Firth & Firth, 2014; Gutting, 2012; Hendricks, 2014), and some of this public discussion has arguably influenced decisions of policymakers regarding psychological science (Nyhan, 2014). Citizens in the general public may remain unaware of the specific internal problems and controversies of psychological science. Yet a pervasive stream of bad news and negative comments from science journalists, policymakers, and scientists in other fields will contribute to a more general negative impression of psychological science among the general public.

In this article, I will discuss three broad areas of concern, namely methodological issues, theoretical issues, and the marketing of psychological science. These are a disparate range of concerns, but they have in common that they contribute to the perception that psychology struggles with scientific rigor. I argue that collectively the nature of these issues damages the credibility of our science in the eyes of others. Similarly, they have in common that they are self-inflicted rather than externally inflicted. In the spirit of Simmons, Nelson, and Simonsohn (2011), I will note that I do not excuse myself from the various issues and errors discussed within and believe most of the issues are those of good-faith human nature. I hope that this essay will spark an understanding that we can all improve and we can work together to make our science better. With that qualifier, I turn first to methodological issues.

Methodological Issues

In this section, I address methodological pitfalls for the field of psychology that have potentially reduced confidence in the reliability of the field among policymakers, other scientists, and the general public. By methodological issues I do not necessarily refer to specific methodologies, such as for laboratory studies, although these can be of issue as well. For instance, the validity of laboratory studies has often been debated, particularly in relation to their generalizability to real-life behaviors (e.g., Ritter & Eslea, 2005; Shipstead, Redick, & Engle, 2012; Wolfe & Roberts, 1993; although see Mook's (1983) famous defense of "external invalidity"). This generalizability issue, although methodological in nature, actually goes to the marketing of psychological science, particularly when artificial laboratory studies are improperly generalized to the real world. The methodological issues I refer to here are not specific to individual studies, but refer instead to broad issues regarding the conduct of psychological science itself and how these methodologies may influence perceptions by the general public.

The Replication Crisis

Roughly since the 2011 publication of an article suggesting that humans could foretell the future (Bem, 2011 and see Wagenmakers, Wetzels, Borsboom, & van der Maas, 2011 for comment), psychology has been gripped with what has been called a *replication crisis*. Replicability issues in science are neither new nor limited to psychology (see Ioannidis' (2005) famous essay on false positives in science). However, the publication of an article about ESP in the top journal of social psychology arguably brought home the likelihood that many other false-positive articles are being published but are not easily identifiable because of less fantastical subject matter. As such, several areas, particularly within social psychology, began to be identified as difficult to replicate ranging from social priming (Doyen, Klein, Pichon, & Cleeremans, 2012; Pashler, Coburn, & Harris, 2012) to media effects (Lang, 2013; Tear & Nielsen, 2013) and embodied cognition (Johnson, Cheung, & Donnellan, 2014). This, in turn, led to public and highly acrimonious debates between the replicators and those whose studies were not replicated about which much as been written (e.g., Bartlett, 2014; Etchells, 2014) and ultimately invited unfavorable comparisons once again to the "hard" sciences (Chambers, 2014).

To be clear, the problem is not that some areas of psychology have ultimately proven difficult to replicate. This can happen in any field, but the replication crisis has unveiled a surprising resistance to the concept of replication within psychological science (Elson, 2014). For instance, with respect to the Bem (2011) precognition study mentioned earlier, replication efforts were initially rejected by the editor of the publishing journal, with the replicating authors told that the journal "does not publish replications" (French, 2012). These included highly acrimonious exchanges between scholars, even those not directly involved in the replication efforts, and essays by prominent scholars (e.g., Mitchell, 2014) advocating that failed replications are scientifically unworthy and should not be published. As Elson (2014) notes, the replication crisis is not a crisis because some areas of psychology have not been replicated, but rather that the debate over replication has revealed an academic culture in which a central tenet of science, replicability, does not appear to be universally valued.

Aversion to the Null

Related to the replication crisis is psychological science's well-known aversion to publishing null results. Aversion to the null can occur both at the level of the editor/journal (publication bias) and because of authors who themselves suppress null findings for various reasons. The issue of publication bias has been known for some time (Rosenthal, 1979). Author-level suppression of null findings has received less attention and may occur because of perceived difficulty in publishing null findings (e.g., Franco, Malhotra, & Simonovits, 2014; van Lent, Overbeke, & Out, 2014), perceived difficulty in interpreting inconsistent results when they occur (e.g., Crotty, 2014), and ideological biases because of theoretical rigidity or the perception that positive findings would support an advocacy goal. It has become the culture of our field that either an empirical test provides evidence for a theory, or the empirical test failed (Schimmack, 2012). This is not an adequate scientific culture.

Like many of the issues described in this article, aversion to the null is not unique to psychology, although there is some evidence that it is particularly common in behavioral and psychological science (Fanelli, 2010). The proportion of studies published in psychological science that support the authors' a priori hypotheses appear to be unusually high. For instance, Fanelli (2010, p. 4) stated, based on his analyses of published reports ". . . the odds of reporting a positive result were around five times higher for papers published in Psychology and Psychiatry and Economics and Business than in Space Science." Fritz, Scherndl, and Kühberger (2012) found that from 395 randomly selected psychological studies, sample size correlated with effect size at $r = -.46$, typically an indication of "sample size chasing" (increasing sample sizes to just adequate size to obtain statistical significance for a given effect size) to achieve statistical significance (statistical analyses generally assume that larger samples have smaller sampling errors and thus more power to detect effects that are assumed to be "true" rather than because of sampling error), or traditional "file drawer" issues in which null studies or results are simply not reported. Having larger standard errors, smaller samples produce more extreme effect sizes, and in the presence of publication bias, more extreme effect sizes that attain statistical significance are more likely to be published. Effect size is generally independent of sample size and thus the expected correlation between effect size and sample size should be zero. A negative correlation between sample size and effect size indicates that publication selectivity is at play, and samples that did not produce statistically significant results are being declined from publication, either by journals or because of authors' practices themselves.

As Bakker, van Dijk, and Wicherts (2012) note that as many as 96% of publications in psychology provide data to support the authors' a priori hypotheses, a track record not warranted by the statistical power of the studies involved. Taken together, these results suggest a strong tendency toward rejecting studies with null results or authors using questionable researcher practices (described later) to convert null findings to statistically significant findings, a tendency that is difficult to explain purely based on quality differences in studies (e.g., Mitchell, 2014 and reply by Neuroskeptic, 2014).

The issue appears even in top journals. In a recent analysis of psychology articles appearing in the journal *Science* (Francis, Tanzman, & Matthews, 2014), the authors found that the success rate for hypothesis tests as indicated by achieving $p = .05$ was much greater than would be expected given the power and design of the studies included. The authors concluded that approximately 83% of published multi-experiment psychology articles in *Science* had suspicious statistically significant outcomes possibly indicating the presence of questionable researcher practices (QRPs), particularly tweaking statistical analyses to achieve p values below .05 or suppressing inconvenient null findings. Thus, authors may feel pressure to achieve statistical significance in order to attain publication or to support beloved theories.

Arguments against publishing null results tend to argue either that null studies are of lower quality than statistically significant ones (although see Neuroskeptic, 2014) or that null results are more difficult to interpret owing to Type II error. Little evidence has emerged, however, that null studies are necessarily of lower quality, and low power causes unreliable results that may influence both null and statistically significant studies. Or, put another way, if we can only publish statistically significant results and our statistics offer no road to falsification, the process we are engaged in is not scientific. This skew in publication practices results in research bodies that do not represent real-world behaviors and may, in fact, be increasingly remote from behavior in the real world.

QRPs

Arguably, the incentive structure in psychological science publishing, which values statistical significance and repudiates falsification, results in pressure for researchers to convert nonsignificant findings into statistically significant ones (Nosek, Spies, & Motyl, 2012). Such an incentive structure coupled with a lack of standardization in methods, data extraction and analysis, or data interpretation may create a research scenario in which research results reflect experimenter expectancies to a greater extent than they do real-life behavior (Klein et al., 2012). That is to say experimenters themselves, sometimes fraudulently, but more often acting in good faith, may be able to "tilt the machine" to produce research results which reflect what they wanted to see rather than what actually is.

This issue was prominently brought to light by Simmons et al. (2011), who demonstrated that statistics can be potentially manipulated to produce statistically significant but absurd results. In one experiment, Simmons and colleagues demonstrated that, by manipulating statistical analyses for an experiment using real data, they could make it appear that listening to a Beatle's song made participants chronologically younger, an obviously impossible outcome. Using simulations, they further demonstrated how manipulating statistical designs could produce considerable spurious results. QRPs may include behaviors such as stopping data collection only when statistical significance is achieved (i.e., sample size chasing), selecting from among data analytic strategies those that produce statistical significance and rejecting those that do not, extracting data from individual measures in ways that maximize the likelihood of statistical significance, convenient exclusion or inclusion of outliers, convenient inclusion or exclusion of covariates in statistical analyses, selective reporting of positive but not null results from studies with inconsistent outcomes, running multiple independent experiments but only reporting those with significant results (something some scholars have publically admitted to), and so forth. Surveys of psychological researchers indicate that such behaviors are rather common (e.g., Fang, Steen, & Casadevall, 2012; John, Loewenstein, & Prelec, 2012), although most often in good faith rather than deliberately fraudulent. That is to say, researcher expectancies tend to direct researchers to select from among multiple possible analytic strategies, those that produce statistically significant results, while eschewing analytical strategies that produce null results. In some cases, research designs and statistical analyses may be selected specifically to offer maximal likelihood of support for a priori theoretical models or hypotheses (Fiedler, 2011). Such approaches most often increase Type I error results and should not be applied.

Several approaches and efforts have been put into place to try to nudge the field away from these types of practices. These include a priori registration of research designs through outlets such as the Open Science Framework, in-principle acceptance of articles based on methods and design before results are known, use of alternate statistical approaches such as Bayesian statistics (Dienes, 2014; Rouder & Morey, 2012), or greater emphasis on outcomes such as effect size or confidence intervals, although these too have the pitfalls. More recently one psychological journal, *Basic and Applied Social Psychology*, has effectively banned null hypothesis significance testing altogether (Trafimow & Marks, 2015). Not surprisingly, this has created significant discussion and debate (e.g., Haggstrom, 2015; Novella, 2015) because this represents a major departure from business as usual regarding statistical analysis in psychology. Whether this effort truly develops into new thinking and greater rigor or simply greater creativity in producing false-positives remains to be seen. However, these efforts are certainly welcome and should be encouraged whether they ultimately succeed or fail.

Lack of Transparency

It is often startling to outsiders to learn that psychological researchers are not required to submit datasets for review as part of the publication process nor are repositories of data kept for confirmation purposes. Although American Psychological Association (APA) ethics require researchers to maintain records for 5 years' post publication and to provide those records to other scholars on request, enforcement of this ethical principle is arguably spotty at best and dependent upon individual journals, which themselves may be reluctant to support a challenge to articles published in their pages. Some scholars (e.g., Wicherts & Bakker, 2009) contend that revisions to the APA ethical code made data sharing less likely, mainly by allowing published authors to place undue demands and restrictions on authors requesting data. Given that requests for data may be perceived as "hostile," there appear to be few incentives for scholars to provide requested datasets and many incentives for them not to, with the possible exception of their own personal consciences. More often than not, both the field of psychological science and the general public has little recourse but to take scholars at their word that their data says what their articles say it does.

Previous studies have suggested that positive researcher responses to requests for data tend to be low, ranging from a low of about 27% (Wicherts, Borsboom, Kats, & Molenaar, 2006) to a high of about 75% (Eaton, 1984). This higher figure was for summary statistics only (data necessary to calculate effect sizes), however, not the raw data. Most studies that have examined this issue have found low compliance figures closer to Wicherts et al. (2006) than Eaton (1984) (e.g., Craig & Reese, 1973; AQ: 2 Wolins, 1962). Wicherts, Bakker, and Molenaar (2011) noted that negative replies (or unresponsiveness) to requests for data were more common for articles published with apparent statistical errors or other evidence methodological problems.

Methodological Problems: What Can Be Done?

As noted earlier, the general public is unlikely to be watching every aspect of psychology's replication crisis and attendant concerns regarding QRPs, transparency, and pub-

lication bias, although policymakers and other scientists are likely paying more attention. However, enough of this information has likely trickled down to the general public to create the image of an investigative field that is defensive, ideologically rigid, nontransparent, and slippery in methodology. No field could survive such a public image with credibility intact.

Fixing this requires several things related to increased transparency, as well as a cultural shift for the field. First, and foremost, psychological science must repudiate its previous aversion to null results generally and failed replications specifically. The credibility of psychological science will not be restored until we are able to communicate our openness to ideas not working out, theories being falsified and, in general, simply being proved wrong. Fortunately, some positive steps are already being taken in this regard. Psychological scholars are beginning to stress the importance of direct replications (e.g., Simons, 2014), and journals such as *Perspectives on Psychological Science* are providing avenues for publishing direct replications. Nonetheless, a more general aversion to the null likely remains pervasive and will continue until psychological journals demonstrate a determinedness to publish null studies alongside statistically significant ones, so long as those studies are of equal methodological merit. Opportunities are being created for scholars to preregister their hypotheses and methodologies in advance of data collection and articles that avail themselves of these opportunities should be reinforced with particular consideration (Wagenmakers, Wetzels, Borsboom, van der Maas, & Kievit, 2012). Replication studies, both positive and negative, could be examined meta-analytically with preregistration status considered as a moderator to examine whether preregistration is likely to result in lower effect sizes.

As for data, it is understandable that an open posting of data to the general public may be precluded by a number of ethical, practical, and proprietary concerns. However, it is argued here that raw datasets should be provided alongside articles to journal editors during the manuscript review process. Many journals, including those of the APA and Society for Personality and Social Psychology, have policies that specifically require authors to share data with other scholars. However, the degree to which authors comply with these requirements or to which journal editors enforce them is unclear. If authors fail to share data, are their articles retracted? By requiring authors to submit datasets to journals upon submission, these ambiguities can be avoided. Furthermore, journals should retain the submitted datasets for the required period of 5 years. In such cases that data are questioned, the journal could act as a repository for the data in question and would be able to directly address concerns or possibly solicit statistical reanalysis if necessary. In accordance with APA ethics, journals would be expected to release datasets to reasonable requests for reanalysis by other scholars. This may help to reduce the impact of reluctant authors, although journals themselves may be reluctant to revisit previously published data. Although imperfect, this may be a crucial step to prevent disappearing datasets that would communicate greater openness and transparency on the part of the research community.

I advocate for a change of culture in psychological science because, at the moment, many scholars appear to believe many of the issues I describe to be normative and acceptable. Such a change can be difficult and slow to achieve. However, one crucial effort in such a change would involve the training of graduate students. Graduate students in psychology and the social sciences need to be informed, most appropriately during the research methods sequence, of the methodological issues in our field, including issues of replication, transparency, QRPs, and null aversion. How these issues detract from scientific progress should be openly discussed. Academic mentors can also demonstrate good scientific practices for students, promote open science, and encouraging trainee scholars to embrace and report research results whether they are theory confirmatory or not. The graduate school years are undoubtedly formative regarding the culture of academic science and thus addressing and instilling good practices during graduate training can be a powerful element of cultural change. Addressing these issues in graduate training may not prevent those determined to be fraudulent but may help reduce good-faith QRPs that may come across as business as usual in the current climate.

Theoretical Issues

In this section, I discuss several theoretical issues related to psychological science and its difficulty in speaking to the totality of human experience. Put simply, I argue that there often exists a gulf between psychological science's explanation for human behavior, which tends to be narrowly focused, mechanistic, and rigid, and the lives people actually live. This does not at all discount Lilienfeld's observation (2012) that people may often misunderstand psychological principles at work in their personal lives. However, I do argue that psychological science can sometimes miss the forest for the trees and make assumptions that the dubious effects seen under artificial laboratory conditions speak to real human experience. Thus, in this section I advocate for new and more complex theories of human behavior.

The Hammer Hypothesis

Psychological science sometimes has the tendency to overemphasize the impact of imitation and modeling and automatic processes on human behavior. Since the 1960s, it has been well understood that humans are capable of learning through processes of imitation and vicarious reinforcement. Indeed, the "Bo-Bo doll" studies of Bandura (e.g., Bandura, 1965) are psychological science classics taught to every introductory psychology student (although not beyond criticism; see Tedeschi & Quigley, 1996). In these classic studies, children were randomized to watch an adult model either hit a Bo-Bo doll in a playroom or play peacefully, were then frustrated, and finally given the opportunity to model the behavior in a similar playroom. Children who watched the adult actor engage in aggression

toward the Bo-Bo doll tended to imitate that behavior, particularly in conditions in which the model had been reinforced for the aggressive behavior. The amount of research conducted on social learning theory is so numerous (a PsycINFO subject search for "social learning theory" OR "social–cognitive theory" turned up over 1,500 hits) that it would be difficult to summarize even in book form.

Nonetheless, public discussions of imitative behavior, either between scholars or to the general public, too often come across as simplistic. Rather than a useful tool that people *can* use to learn useful skills or behaviors, imitation is often presented as something people, particularly children *must* do (e.g., Heusmann & Taylor, 2003) and in which automatic and senseless responding to even subtle stimuli has been emphasized (e.g., Bargh, & Chartrand, 1999). Some scholars may acknowledge the complexities involved given multiple potential models in a person's life (e.g., Kazdin & Rotello, 2009), yet modeling is still often discussed as a powerful, overweening and automatic (rather than chosen) force for shaping behavior.

I refer to this as the Hammer Hypothesis (in the sense of, if the only tool you have is a hammer, everything begins to look like a nail) in that psychologists may overemphasize the power and, perhaps more importantly, automaticity of modeling because social learning theory is the major, unique, contributing body of work that originated specifically from within psychology to explain human behavior. The critique here is not to suggest that modeling does not occur (which would be absurd) but that too often psychologists present it too simplistically, which contributes to the reputation of psychologists as perceiving individuals as programmable machines rather than complex, motivation-driven, agentic people. Imitation also is too often discussed in the absence of other factors whether biological (Pinker, 2002) or motivational (Przybylski, Rigby, & Ryan 2010). These simplistic portrayals of imitation or other automatic responses to stimuli may be particularly attractive in relation to moral issues, such as highlighting the potential "harm" to minors because of exposure to objectionable stimuli . . . the "harm" coming through imitation of those stimuli. However, this portrayal of imitation as senseless and automatic as opposed to purposeful and practical may contribute to the perception that psychology is out of touch with human experiences.

Theoretical Rigidity and Lack of Open Dialogue

A further problem occurs when theoretical models of human behavior become rigid and purveyors of such theories begin to see them as "truth" rather than heuristics that may need to be adjusted or, indeed, abandoned when a body of contradicting new data comes to light. This can be particularly likely for theories that are rooted to public policy or social morality advocacy goals. Theories that provide structure for advocacy goals, however well meaning, may effectively reverse the scientific process in which the publishing process begins with a desired endpoint and the

process of data collection is a mere proforma one to support the theoretical model. Although theory in science is of critical importance, it is understood that theories can at times obstruct scientific progress (Greenwald, 2012; Greenwald, Pratkanis, Leippe, & Baumgardner, 1986). This occurs when scholars become emotionally invested in a theoretical perspective. Inconclusive data may be interpreted as supportive of the theory and nonsupportive data ignored, criticized, or suppressed.

Theoretical rigidity can lead to "scientific inbreeding" and obligatory replication in which the replication process is subverted to provide support for a given theoretical model rather than to seriously test whether it is falsifiable (Ioannidis, 2012). Defenders of a given theoretical model may engage in same-team replications, and (whether purposefully or not, such as through peer review of journal articles) put pressure on other scholars to produce similar results (possibly through QRPs or publication bias; see Shooter, 2011). The result is a fairly common pattern: theory-glorifying literature that can stretch for years or, sometimes, decades, only to come under considerable doubt once more skeptical scholars attempt replication (Coyne & de Voogd, 2012).

Rather than a constant eye at testing theoretical models with falsification at the forefront, the result is often theoretical "camps" of defenders and supporters who can become acrimonious (examples of this in psychological literature are numerous enough that any reader should have little difficulty finding them). In particular, advocates of particular theoretical models can view skeptics as "heretics" and employ a reverse burden of proof on skeptics. This closing down of dialogue is anathema to science. Undoubtedly, theory defenders may feel their own reputations threatened should confidence in a theory (particularly one with which their name is closely linked) comes under doubt. But this preservation of ego should not be considered primary and used to diminish opportunities for open scientific dialogue in which skepticism even of "beloved" theories should be encouraged.

Thus, we are well served to remember that scholars are not immune to confirmation biases that are in play for all people. Nickerson (1998) documents that confirmation bias is near ubiquitous, including in science and provides examples of noted scientists who engaged in confirmation bias practices in their scientific work. Individuals, including scholars, are likely to spend longer scrutinizing evidence that challenges their prior beliefs (Edwards & Smith, 1996). Biased cognitions such as belief perseverance and myside bias appear to be independent of intelligence (Stanovich, West, & Toplak, 2013) and may be part of a process that makes scientific debates both less productive and more acrimonious.

Advocacy and Science

Perhaps one of the bigger challenges for academic psychology is the dual role that psychology often finds for itself in advocating for human welfare while at the same time attempting to find objective scientific facts. This duality is

not surprising given that many of the subjects open to scientific psychology touch upon concerns related to psychological wellness, social justice, and public policy. Arguably, many individuals with academic appointments and publishing in reputable scientific journals find themselves functioning both as scientists and as advocates for particular policy or social justice goals.

The difficulty is that advocacy and science are diametrically opposed in method and aim. On an idealistic level, science is dedicated to a search for "truth" theoretically even if that truth is undesired, inconvenient, unpalatable, or challenging to one's personal or the public's beliefs or goals. By contrast, advocacy is concerned with constructing a particular message in pursuit of a predetermined goal that benefits oneself or others. Because of these diametrically opposed processes, it is unlikely that either organizations such as the APA or individual scholars are likely to be successful at operating as both scientists and advocates at the same time. Fundamentally, because advocacy may be more immediately attention-getting than science, I propose that mixing science with advocacy almost inevitably ends in damage to the objectivity of the former.

This may not be immediately apparent to the scholars themselves for whom partnerships with advocacy groups may seem a "meeting of the minds" in pursuit of social agendas. Yet the urge to intermingle science with advocacy can be detrimental to the scientific process, reducing objectivity and impartiality. As Grisso and Steinberg (2005, p. 619) noted, "scientific credibility demands impartiality, whereas advocacy is never impartial." As the authors also note (p. 620), "when developmental scientists choose what they will study about children and their welfare, they are often motivated by personal beliefs and values about the importance of child protection," and these personal beliefs and goals may lead to scientists knowingly or unknowingly forcing data to fit these beliefs or goals. Close associations with advocacy groups, particularly via research funding, may further reinforce ideological values and remove the scholars further from objective science. Additionally, Brigham, Gustashaw, and Brigham (2004, p. 201) admonished that "advocacy, however, requires conviction. Advocates 'sell' their positions so that they might convince others of a specific solution to a problem. Uncertainty is unhelpful to the advocate as well as to the salesperson."

This objectivity is reduced further by the observation that the body of psychologists tends to be uniform in beliefs and agendas that may not reflect the general populace (Redding, 2001). Thus, the advocacy goals selected both by individual scholars and professional advocacy groups such as the APA may themselves be the product of self-selected political and social inclinations rather than representing a plurality of ideas and perspectives. Policy statements in pursuit of a particular set of ideological agendas, however well-meaning, can place the profession inadvertently at odds with the remainder of the populace. Were this a matter of pure science and data, this could be perceived as a brave stance in favor of "truth," but because this comes in pursuit of an advocacy agenda, the fragile nature of social science (see again Simmons et al., 2011) can become perverted in

pursuit of a social agenda that may represent a particular ideology rather than a pluralistic view of human welfare.

As a potential illustrating point of the dangers of scholars taking advocacy positions, in the 1990s, the Centers for Disease Control and Prevention (CDC) lost funding for gun-violence research as a result of its perceived biases on the issue, particularly from Republican congress members, following statements by CDC personnel that appear to specifically advocate gun control policy (e.g., New York Times, 1994; see Bell, 2013). Naturally, gun control/rights policy is complex, and it is reasonable to argue that opponents of gun control may have looked for any opportunity to defund studies of gun violence. However, where CDC officials strayed beyond careful objective and policy-neutral presentation of data into an advocacy agenda, the CDC made it that much easier for politicians to argue, even if cynically, for the defunding of gun studies. I argue here that attempts by psychological science to attempt to juggle science and advocacy simultaneously risk the same issues and damage the credibility of our field in the process.

Theoretical Issues: What Can Be Done?

The first answer to most of the issues I outline above is to seek ways to make psychological science more data-driven than theory-driven (although, to be clear, I am not advocating small sample exploratory studies likely to prove difficult to replicate). At present, psychological science is too often in danger of reversing the scientific process by proving how "true" theories are either for ideological or advocacy reasons. As a result, it is too easy to ignore inconvenient data, engaged in "obligatory" or "same-lab" replications, and nudge data to fit existing theories, rather than using data to falsify them. I advocate a straightforward process. Theories should be developed from preexisting data, with an active effort to search out data that may be problematic for a theory in development. Such theories should be able to provide clear, testable hypotheses, including clear guidelines for when those hypotheses would be falsified and when the theory itself should be considered to be falsified. Tests of a theory should be preregistered with equal opportunity for publication, whether supportive or not. This proposal is neither remotely novel nor should it be controversial but, at present, I do not suspect that it is the norm.

Reestablishing a culture of falsification would be critical in moving the reputation of our field forward. So too would be reconsidering our close proximity to advocacy goals. I strongly suggest that psychological researchers should avoid accepting funding from advocacy groups advancing particular policy or social advocacy agendas, however well-meaning these may be. It is probably impossible to avoid any pressure to produce certain results whatever funding source may be available, even with government funding, but avoiding obviously biased sources would be helpful.

If psychological science is to consider examining more complex theories of human behavior, which I gener-

ally endorse, this comes with a few caveats. First, such theories must remain empirically based and care must be taken to avoid hidden problematic assumptions. Second, more complex theories would require more sophisticated analyses, potentially driving up demands for power and larger samples. Individual experiments could test smaller pieces of larger theories, but in general, scholars must be aware of power needs and the potential for Type I errors to promulgate a given theory beyond its natural life span. Put simply, theory construction and testing needs to be done with far greater care than in the past. Balancing the development of more sophisticated theory with maintaining methodological rigor will not be a simple task.

Regarding the issue of advocacy, it is worth asking: When does data in psychology become sufficient to endorse a particular advocacy or political course of action? When should psychological researchers feel comfortable vocally supporting advocacy or policy agendas related to their field of study? This is obviously not an easy question to answer.

Three things should be acknowledged. First, policymakers and the public will often seek definitive answers and support for particular agendas before the data is clearly able to do so (which, in fact, may never happen) and may put pressure on scholars and groups such as the APA to support these agendas. Second, data and even scholarly consensus often changes rapidly in psychological science; thus, what is "true" one year may be falsified the next. Third, there is likely no existing shortage of social agenda advocates outside of the psychological research community and thus, little clear need for psychological researchers to join these ranks.

Thus, given the problems with advocacy in psychological science in the past, I suggest that the default position for psychological researchers should be to simply present and clarify research data carefully, *warts and all*, and restrain their statements to specific comments on data including corrections to any hyperbole or misstatements made by advocates or policymakers. Obviously, many psychological researchers may be displeased by this recommendation but, until our scientific culture changes so that we have a clearer, objective read of data even in socially sensitive areas, indulgence in advocacy will only continue to damage the objectivity of psychological science. If issues such as the replication crisis, lack of transparency, and QRPs appears to clear in the future, this stance could be reassessed then, but only with great caution.

This does not mean that scientists cannot provide information to interested policymakers, only that such information should be strictly data-based, include data that may conflict with the scholar's personal views, and be restrained from interpretations advocating a particular course of policy. In other words, "Just the facts, ma'am." This also does not mean that advocacy cannot be informed by science, naturally, it can. Furthermore, scholars can certainly advocate for causes in which they believe. However, it may be wise for scholars not to function as advocates on topics on which they do research. As such, I am

recommending a clear distinction between scholarly and advocacy activities.

Concerns about theory does not mean that theory has no value in psychology; of course, it does. We may simply need to consider alternate theoretical models that are less mechanistic and speak to a broader range of human experience. As part of this, we may need to acknowledge that our current methods may not yet be able to definitively answer all questions of interest. Similarly, we need to be careful not to imply that the small effect sizes obtained in most psychological data can be translated to a fully mechanistic view of human experience. The challenge for psychology forthcoming will be in becoming more data-focused while acknowledging that the data we have may not always fully capture the human condition.

Marketing of Psychology

Professional organizations are understandably interested in promoting the field of psychology. Indeed, this is one of the primary purposes for such institutions. At the same time, marketing approaches may lead to making bigger promises than can be delivered, and advocacy organizations such as the APA may display biases relevant to false positive claims of nonexistent social problems psychologists would nonetheless fix. This next section turns to the issue of how psychology is "sold" to the general public and how some of these efforts may backfire. This is not to say that the field should not reach out to the general public to explain what we can do—just that, here again, a cautious and honest approach may be most effective in relation to long-term perception.

Policy and Other Official Public Statements

One means for demonstrating that a profession is active in the global community and promote itself is to attempt to connect research to policy issues under debate in society. Policy statements take many forms and serve multiple purposes. For instance, professional organizations may release policy statements that pertain to professional ethics or that otherwise are intended to state positions directly relevant to the field. Policy statements may also pertain to issues less directly relevant to the field with the implication that research from the field can be helpful in guiding policy debates in other realms. In effect, these policy statements take a stance on a particular debated issue and imply that psychological science can or should be used by others to help inform their decisions as well. Such policy statements can appear beneficial to the field to the extent that they garner press attention and support from advocates of the position advanced by the policy statement.

The website of the APA's Public Policy Directorate (http://www.apa.org/pi/guidelines/index.aspx) lists approximately 70–80 active policy statements, in addition to several advocacy letters to government officials and a document detailing 52 resolutions concerned with the status of

Fn1 women in psychology.[1] By contrast, the American Society of Criminology (ASC) lists only three advocacy letters to government officials and two statements from the policy committee related to incarceration and the death penalty, which both carry the disclaimer that they represent the policy committee, not necessarily the ASC.[2] Thus, the **Fn2** APA appears to be particularly enamored of policy statements. Naturally, making direct comparisons between two fields of different sizes must be done with caution, but both psychology and criminology address issues related to public policy, yet take clearly different approaches as to the appropriate frequency of such policy statements and when a threshold of evidence has been reached to support them.

Although the opportunities for news media attention and political influence that come with policy statements are understandable, there are also risks associated with the use or, perhaps, overuse of policy statements. First, the field of psychology can develop a reputation for being opinionated, particular to the degree that policy statements extend into areas of human welfare, however well intended, that are beyond the strict realm of a science that strives to be objective. This is particularly risky when psychologists themselves do not represent a pluralistic view of social and political perspectives (Redding, 2001). Consistent with this observation, policy statements by the APA and other scientific organizations appear to support generally left-leaning positions likely to alienate social conservatives (the allegation that the APA colluded with the Bush administration to allow psychologists to participate in enhanced interrogations during the wars in Iraq and Afghanistan is the obvious, yet still problematic, counter example. Risen, 2014). Scholarly organizations such as the APA have consistently failed to distinguish, in their public policy statements, between scientific "truths" (which are few and far between) and policy positions that can be informed by, but not dictated by psychological science.

Second and related, the impact of each policy statement arguably becomes less the more policy statements are issued. An objective scientific organization taking the unusual step of releasing a policy statement on a particular issue can have dramatic impact, particularly when the reputation of the field as an objective neutral arbiter has otherwise been maintained. However, frequent use of policy statements, particularly on advocacy issues that are up for debate in the public sphere, can ultimately result in the field developing the reputation as an invested politicized entity and the impact of such statements is reduced.

Third, by becoming invested in particular social advocacy issues, the field may inadvertently engage in citation bias, only citing scientific evidence supporting a policy statement while ignoring scientific evidence that contrasts with the policy statement. Such policy statements may gravitate toward sweeping politicized statements while misrepresenting, intentionally or unintentionally, nuances or discrepancies in the research.

Thus, it is not surprising that many policy statements by both psychological and other health and behavioral health organizations have been criticized for inaccuracies, lack of clarity, and a tendency to put agenda before a careful representation of science (Goffin & Myers, 1991; O'Donohue & Dyslin, 1996; Olson, Soldz & Davis, 2008) and fail to distinguish whether hard data or personal opinion is guiding policy stances (Martel, 2009).

Policy statements may be further reduced in objectivity when they are written by specific groups of scholars highly invested in specific advocacy goals. It is perhaps this past point that is most critical. To the extent that professional organizations such as the APA attempt to "sell" their opinions on various social issues as well-replicated findings, such organizations can risk damaging the credibility of the field as an objective arbiter of information on human behavior.

I provide three examples of APA position and/or policy statements that are arguably problematic. The first concerns the reaction of the psychological community as exemplified by task forces and position statements promoted by the APA that appear to be critical of the Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (PRWORA). PRWORA was a bipartisan welfare reform effort of the 1990s that has generally been considered successful by politicians in both parties (e.g., The New Republic, 2006). Nonetheless, a task force of Division 35 (Psychology of Women) in apparent cooperation with the APA's Public Policy Office, released a position paper that appeared critical of PRWORA and portrayed many citizen's concerns with welfare as "myths" (Task Force on Women, Poverty, and Public Assistance, 1998). This is not to say that the position statement was without valid points, but rather that it took only a limited (and arguably left-leaning) perspective, appeared to unintentionally disparage alternate perspectives, and arguably attempted to apply psychological science to a real-world scenario where it may have been ill-suited. Although certainly well intended, this position statement ultimately appeared to place psychology in opposition to a bipartisan and largely successful (although certainly imperfect) effort to address a pressing social problem.

The second example concerns various policy statements by the APA on the issue of abortion. As a moral issue, many may agree with the APA's proabortion rights stance (for the record, this author is staunchly neutral on this issue). But should a scientific organization take a stance on a purely political/moral issue particularly when it (potentially selectively) cites scientific articles in support of this position? The APA has multiple policy statements dating from 1969, 1980, and 1989, and including a 2008 task force report on the research (APA Task Force on Mental Health & Abortion, 2008). The 2008 task force appears to provide scientific rationale for previous policy

---

[1] The actual number of policy statements is arguably much larger, including internal policies and older policy statements technically still on the books. A full listing of APA policy statements is provided at http://www.apa.org/about/policy/index.aspx.

[2] It is worth acknowledging that, in conversation with the current President of the ASC Chris Eskridge (personal communication, August 25, 2014), he noted that the ASC continues to debate whether to be more or less active in the policy arena.

statements, concluding that a single abortion for unplanned pregnancy did not impact women's mental health any more than continuing the pregnancy. While certainly noting potential issues for women with multiple abortions, the task force concluded that (p. 4) "In general, however, the prevalence of mental health problems observed among women in the United States who had a single, legal, first-trimester abortion for nontherapeutic reasons was consistent with normative rates of comparable mental health problems in the general population of women in the United States." Prolife groups have criticized the APA for being selectively critical of research that conflicts with their policy statements, ignoring methodological flaws in research supporting their position, and allowing tasks forces to be composed of scholars with a priori pro-choice moral stances (e.g., Ertelt, 2006) who downplayed results from studies associating abortion with subsequent mental health issues in women (e.g., Fergusson, Horwood, & Boden, 2008; Pedersen, 2007).

Current APA policy (initially implemented in 1969, updated in 1980, and to the knowledge of this author, remaining in effect) states "that termination of pregnancy be considered a civil right of the pregnant woman." It is unclear that it is appropriate that social science be used to make such a moral and policy determination. For instance, it is unclear the degree to which women's mental health consequences related to abortion are critically meaningful regarding whether abortion is moral. Nonetheless, the results of the 2008 task force appear to be geared to support a moral stance by the APA. Furthermore, there is concern that policy and task force statements by the APA presented a far more conclusive portrait of the research literature than may have been warranted by inconsistent data. It is not the intent of this article to take a position on abortion and mental health, only to suggest that policy statements such as these may create a "tail wags the dog" effect in which science is selected to support a preexisting policy instead of science being carefully and objectively communicated to policymakers and the general public. Moral positions that can certainly be informed by science but not determined by science and it is unclear whether it is wise for professional groups to endorse moral positions.

The third example concerns the APA's 1994 policy statement on media violence (APA, 1994) and subsequent 2005 policy statement on interactive media violence (APA, 2005). These statements, and others like it, have long been criticized for failing to include research findings that conflicted with the policy statement (see Freedman, 1984, for a discussion of conflicting findings predating the 1994 resolution), presenting inconsistent research findings as consistent, and extending often problematic research programs into a real-life social issue when the external validity of such conclusions were in doubt (see Dillio, 2014; Hall, Day, & Hall, 2011; O'Donohue & Dislin, 1996, for recent discussions). The tasks forces creating these statements, as in other controversial areas, have also been criticized for appearing stacked with members likely to endorse a particular view without counterbalancing of differing views. Unfortunately, this trend appears to have continued with

the task force currently reviewing the literature on video game violence, which likewise has been criticized as appearing to be purposefully weighted with individuals with a priori antigame views (Ferguson, 2014).Such policy statements arguably mistake science for moralization and indulge in "culture war" and damage the credibility of scientific organizations once research disconfirming the policy statements can be readily identified. Scientific organizations should, at minimum, present a full summary of research, including conflicting findings, not merely that supporting a policy statement.

There are many other historical examples of policy statements or other public statements that appear either to have little to do with psychological science, are simply left-leaning takes on policy issues, or seem simply unnecessary because they take positions few would disagree with. Examples of politicized policy statements include a 1982 policy on limiting the nuclear arms race; 1969, 1980, and 1989 policies supporting abortion; and various policies or other public statements supporting liberal issues, including affirmative action, equating Zionism with racism, same-sex couple rights, and the Affordable Care Act. Unnecessary resolutions (so-called because the position of any reasonable entity should have been obvious) include a 2008 resolution against genocide, a 2002 resolution stating that psychologists should work to reduce terrorism, and a 1999 resolution against child sexual abuse.[3] Some policy statements such as that on Zionism and arguably abortion may lack cultural sensitivity toward groups these may offend. It is not my position that these opinions by the APA are "wrong" (in fact, I share most of them), but rather that the APA simply expresses what are, indeed, opinions too often, obfuscates policy opinion for scientific fact, and consistently leans in a liberal political direction in a way that alienates many policymakers and the general public.

This is not to say the policy statements are never warranted, of course. Policy statements that directly relate to the practice of psychology are well within the purview of organizations such as the APA. Condemnations of therapies designed to alter sexual orientation are an example of policy statements that take a stance on an issue of social justice, while still grounded in empirical data and practice within the field. Policy statements in other areas less related to psychological practice may also be warranted so long as the policy is clearly identified as one of advocacy rather than science. However, frequent use or overuse of advo-

---

[3] However, it is granted this policy statement has a curious history and appears to be in response to an article published in *Psychological Bulletin* that appeared to minimize the impact of child sexual abuse (see Lilienfeld, 2002, for full history). In a letter to Congressman Tom Delay, then APA CEO Ray Fowler (1999) noted that this policy statement was intended to reaffirm the APA's position on child sexual abuse. Dr. Fowler also stated "We acknowledge our social responsibility as a scientific organization to take into account not only the scientific merit of articles but also their implications for public policy" which, aside from the obvious seriousness of the topic of child sexual abuse, I argue is a worrisome position for a scientific organization to take. Should we not publish valid scientific research if it is inconvenient for public policy objectives?

cacy agendas run the risk of undermining the scientific reputation of fields that release too many on too many issues while blurring the lines between science and advocacy.

Small Is Big

One issue that has plagued psychological science is the observation that so many results from empirical studies demonstrate only very small effects, the practical significance of which is unclear (Wilson & Golonka, 2012). For instance, in one recent survey of meta-analyses in psychology, the authors found that most meta-analyses found average effects (in terms of r^2 here for ease of understanding) in the range of 0%–4% (corresponding to rs below .1 through to .2, with few above this level) shared variance between independent and dependent variables, with many of the effects at the range of 1% and lower (Ferguson & Brannick, 2012). Bakker et al. (2012) reported slightly higher average effect sizes in meta-analyses (reported as approximately $d = .37$, p. 547), although the authors suggest that these effect sizes are likely inflated by QRPs. Too many papers have been written on the topic of effect sizes and their interpretation to do remote justice to covering them here. However, the concern expressed here is that, too often, the field has constructed pat explanations for why "small is big"; that is to say, the small effects typically demonstrated in psychological science remain of tantamount practical significance.

One of these is the argument that small effects spread over a wide population can still be of great significance even if only a small proportion of people are influenced. For example, if a study finds that viewing more negative words on Facebook is associated with a decline in mood of 0.07% (e.g., Kramer, Guillory, & Hancock, 2014), the argument sometimes seen is that, in a population of, say, one billion worldwide users, if Facebook even influenced only 0.07% of them, that is still 700,000 people dramatically influenced (and presumably endangered) by Facebook. However, this argument fundamentally misrepresents the effect sizes. The r^2 demonstrates the typical impact on the *average user*, not the proportion of users impacted in a more dramatic fashion. Presumably few individuals would note 0.07% (that is about 1/15th of a percent, not 7% reduction in mood from one day to the next; Grohol, 2014).

The other common and equally problematic argument seen sometimes is the temptation to compare psychological effect sizes with those found in important medical studies, with the implication that psychological effects sizes are *bigger* than those perceived in vital medical studies (e.g., Meyer et al., 2001; Rosnow & Rosenthal, 2003) and have been used to defend weak claims in psychology up through and including parapsychology (Bem & Honorton, 1994). That these claims are "good news stories" for psychology is obvious, but they are problematic on two levels.

First, the calculations made to compare medical to psychological effects are based on statistical analyses that are now known to be flawed (Block & Crain, 2007; Fer-guson, 2009; Hsu, 2004; Kraemer, 2005). Without delving too far into the "inside baseball" of the statistics, the problem was one of nonequivalence, namely the tendency for medical studies to use different effect size indices (typically odds ratio or OR) based on different types of data than those commonly employed in psychological science (most typically either Pearson r or Cohen's d). Thus, to make comparisons it was necessary to convert OR into "r" or "d." In most cases, this transmogrification was done using formulas that include sample size as a denominator term (despite the fact that sample size and effect size should be independent). The problem arises with medical epidemiological studies in which the number of participants reaches into the tens and even hundreds of thousands (such as the Salk vaccine trials or the physician's aspirin study, two studies falsely highlighted as having "small" r effect sizes). Using sample size as a denominator term for the conversion of OR to r dramatically truncates the resultant r value to near zero. That this seemed "good news" to psychologists used to defending r values in the .1 to .2 range seems to have acted as a block to asking whether such results were too good to be true and might have been the product of flawed formulas. Related to this is the observation that most of the medical samples in question used skewed samples, detecting relatively rare events such as cancer in large populations. It is inappropriate to make direct comparison between samples with differing distributions, using very different types of outcome analyses. Unfortunately, these comparisons between medical and psychological research are statistically meaningless.

However, and as the second point, even had they been meaningful statistically, they were nonetheless an "apples and oranges" comparison. Psychological research very often has to contend with "proxy" measures that estimate the actual behavior of interest. Neither laboratory measures of behavior sampled under unrealistic and unnatural circumstances nor self-report surveys are a true measure of most of the behaviors that interest us. Thus, even for the best measures of behavior, there are issues related to reliability and validity. However, many medical epidemiological studies measure mortality or pathology rates that are not proxy measures; put simply, death is a perfectly reliable and valid measure of death. Thus, overzealous comparisons between psychological and medical research are fraught and potentially do more to make the field appear desperate rather than rigorous.

Juvenoia

Juvenoia, or the fear of youth among older adults, is a term coined by sociologist David Finkelhor (2010) and involves society's efforts to find "evidence" that youth today (in any generation) are, somehow, psychologically worse than in previous generations. Psychological science can unwittingly promote juvenoia through claims insinuating unfavorable psychological characteristics of the current generation of youth, compared with previous generations. Research directed at such claims may appear to be partic-

ularly lucrative in regards to grant funding, book sales, prestige from older adult members of society, and so forth.

Examples of this in psychological science include the disparagement of the recent generation of youth as experiencing a "narcissism epidemic" (e.g., Twenge & Campbell, 2009), being less empathic (Konrath, O'Brien, & Hsing, 2011) or incapable of rational thought because of alleged structural deficits in the brain (Galván, 2013). Such narratives about youth may fit well with typically generational struggles in which older adults may enjoy hearing "evidence" that youth are less rational or desirable than older adults are. However, the validity of these lines of inquiry that apparently disparage youth have often been questioned (e.g., Grijalva et al., 2015; Males, 2009; Willoughby, Good, Adachi, Hamza, & Tavernier, 2013), and statistical evidence suggests that the current generation of youth are psychologically healthier on most outcomes, compared with previous generations (Childstats.gov, 2014). Coupled with an apparent overeagerness to condemn elements of youth popular culture as "harmful" based on weak data (see Adachi & Willoughby, 2013; Hall, Day & Hall, 2011; Steinberg & Monahan, 2011, for discussions), psychology may appear to be overly willing to contribute to generational squabbles, particularly at the expense of youth, rather than functioning as fair arbiters of careful science.

Juvenoia among psychological researchers is unlikely to go unnoticed among youth themselves, who may be exposed to a litany of scholars describing them or the culture they value in reproachful terms. Such exposures to comments by psychological researchers may contribute to a longer term sense of unease about the objectivity and validity of psychological science.

Death by Press Release

The last issue I mention related to the marketing of psychological research is what I call "death by press release" or the tendency, when talking about research in news media, to make statements that go beyond the data and that minimize the limitations of psychological research so as to make it sound "better" than it actually is. There is an element of human nature to this, with researchers necessarily being excited about their research, and public relations departments pushing for flashier press releases more likely to get headlines.

Such press releases may be good news in the short-term for researchers but may backfire in many ways. For instance, dramatic claims in news media may work against the credibility of psychological science when those claims are later found to have been based on faulty methodology (e.g., Brown et al., in press). Similarly, exaggerated claims made in press releases may encourage other scholars to find greater fault with the original study than had the study been marketed more conservatively, perpetuating that image that psychological scholars squabble continuously over small data. Although resultant news reports are likely to make for good theater, it is unlikely they make for good science. Furthermore, given that corrections to faulty science are less likely to receive coverage than the initial press releases

(Brown et al., in press), exaggerated press releases likely contribute to "myths" about psychology rather than a careful and systematic understanding of psychological processes. It has been observed that new research often experienced a decline effect (Shooter, 2011) wherein many initial results ultimately prove difficult to replicate. With this in mind, exaggerated press releases on novel but unreplicated findings can do more to misinform than inform the general public.

Marketing of Psychology: What Can Be Done?

The issues surrounding the marketing of psychology all focus on the difficulty in balancing a desire to attract attention to our work while appropriately retaining a conservative and objective perspective. Much can be done simply by changing the culture related to marketing psychology. Certainly, psychological science should be active in promoting "truth" and questioning "myths" related to human behavior (Lilienfeld et al., 2009). At the same time, psychological science must be careful, in its eagerness to promote itself, not to inadvertently create new myths that come from within psychological science as opposed to pop psychology.

Some of the issue is a failure to learn from history, such as on the issue of juvenoia. Scholars making dramatic negative statements about "kids today" is nothing new, yet our field persists in such behavior. Similarly, "small is big" arguments, particularly those based on comparisons to medical science, have been known to be problematic for some time, yet their use persists. I suggest that these points to inherent and difficult struggles between presenting what we want to believe about our field and its contribution to society and what we are actually capable of saying based on data.

Furthermore, professional organizations such as the APA may wish to rethink both their approach regarding the number and purpose of policy statements, particularly where they speak to broader issues not directly related to the practice of psychology. This is not to say policy statements on social issues are never warranted, but existing policy statements do not appear to have been constructed through a transparent and objective process, often contain inaccuracies or distortions of the science, and appear more as political statements than they do scientific statements. It is not my contention that the advocacy causes that the APA and other organizations have supported are not good causes (again, I agree with most of them), but rather that the mission of such groups as advocacy groups and purveyors of objective and careful scientific information are often directly opposed and noticeably so. Less may indeed be more in this particular realm.

Concluding Thoughts

It would be easy to view the proceeding as a litany of complaints about psychological science. However, they are offered with a spirit of optimism. I fully agree with Lilienfeld (2012) that psychological science, presented carefully

and strategically, is recourse to the plethora of folklore and myths about human behavior that persist in the general public. At the same time, there are some straightforward fixes we can consider as we bring psychological science further into the public consciousness.

The suggestions I offer are, by no means, impossible ones. By and large, I am suggesting something of a change in culture. This change in culture can and should begin with the way we train graduate students, by having open discussions about these issues during graduate research training. At the root of the issue is understanding that, particularly given the emotional, moral, and political overtones of our field, we need to hew more closely to the data (rather than advocacy or even theory) and eschew grandiose statements where they are not warranted. At the same time, we would benefit from exploring new ways to provide a more sophisticated understanding of the human condition that moves beyond the reflexive mechanism that has typified so much of the history of psychological science. Furthermore, by taking the lead on issues of transparency, replication, and falsification we can demonstrate our field's commitment to fundamental and nonnegotiable elements of any science. These cultural fixes are doable and doable in our scientific lifetime.

It is not my intention to be Pollyannaish either. Any cultural change can be difficult, particularly when there are incentives to resist that change. News media may continue to grant headlines to dramatic novel (but ultimately unreplicable) claims. Grant funding may continue to go to scholars who make big promises, or exaggerated claims of pressing social issues that may not, in fact, exist. I argue that just because change may be difficult does not mean we should not try. Even scholars acting as individuals, trying to improve the objectivity of their own work, preregistering their studies for instance, can have a cumulative effect.

Psychology is a topic about which policymakers and the general public are hungry for knowledge. As Lilienfeld (2012) so aptly noted, the general public remains wary of psychology as a field however, and too often gets information from pop psychology and other unscientific outlets. With care and caution, psychological science can position itself to take the lead in guiding the general public's understanding of human behavior in future years, but it will take equal measures of sophistication, open-mindedness, careful adherence to scientific principles, and humility.

REFERENCES

Adachi, P. C., & Willoughby, T. (2013). Do video games promote positive youth development? *Journal of Adolescent Research, 28,* 155–165. http://dx.doi.org/10.1177/0743558412464522

American Psychological Association. (1994). *Violence in mass media.* Retrieved from http://www.apa.org/about/policy/media.aspx

American Psychological Association. (2005). *Resolution on violence in video games and interactive media.* Retrieved from http://www.apa.org/about/governance/council/policy/interactive-media.pdf

AQ: 3 American Psychological Association. (2015). *Abortion resolutions.* Retrieved from http://www.apa.org/about/policy/abortion.aspx

APA Task Force on Mental Health and Abortion. (2008). *Report of the APA Task Force on Mental Health and Abortion.* Washington, DC: Author.

Bakker, M., van Dijk, A., & Wicherts, J. M. (2012). The rules of the game called psychological science. *Perspectives on Psychological Science, 7,* 543–554. http://dx.doi.org/10.1177/1745691612459060

Bandura, A. (1965). Influence of models' reinforcement contingencies on the acquisition of imitative response. *Journal of Personality and Social Psychology, 1,* 589–595. http://dx.doi.org/10.1037/h0022070

Bargh, J., & Chartrand, T. (1999). The unbearable automaticity of being. *American Psychologist, 54,* 462–479. http://dx.doi.org/10.1037/0003-066X.54.7.462

Bartlett, T. (2014). Replication crisis in psychology research turns ugly and odd. *Chronicle of Higher Education.* Retrieved from http://chronicle.com/article/Replication-Crisis-in/147301

Bell, L. (2013). Why the Centers for Disease Control should not receive gun research funding. *Forbes.* Retrieved from http://www.forbes.com/sites/larrybell/2013/02/12/why-the-centers-for-disease-control-should-not-receive-gun-research-funding/

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. http://dx.doi.org/10.1037/a0021524

Bem, D., & Honorton, C. (1994). Does psi exist? Replicable evidence for an anomalous process of information transfer. *Psychological Bulletin, 115,* 4–18. http://dx.doi.org/10.1037/0033-2909.115.1.4

Berezow, A. (2012). Why psychology isn't science. *The L.A. Times.* Retrieved from http://articles.latimes.com/2012/jul/13/news/la-ol-blowback-pscyhology-science-20120713

Block, J. J., & Crain, B. R. (2007). Omissions and errors in "media violence and the American public." *American Psychologist, 62,* 252–253. http://dx.doi.org/10.1037/0003-066X.62.3.252

Brigham, F. J., Gustashaw, W. E., III, & Brigham, M. S. (2004). Scientific practice and the tradition of advocacy in special education. *Journal of Learning Disabilities, 37,* 200–206. http://dx.doi.org/10.1177/00222194040370030301

Brown, N., MacDonald, D., Samanta, M., Friedman, H., & Coyne, J. (in press). A critical reanalysis of Fredrickson et al.'s study of genomics and well-being. *Proceedings of the National Academy of Sciences of the United States of America.*

Chambers, C. (2014). Physics envy: Do "hard" sciences hold the solution to the replication crisis in psychology? The Guardian. Retrieved from http://www.theguardian.com/science/head-quarters/2014/jun/10/physics-envy-do-hard-sciences-hold-the-solution-to-the-replication-crisis-in-psychology

Childstats.gov. (2014). *America's children: Key national indicators of well-being, 2010.* Retrieved from http://www.childstats.gov/

Coyne, J. C., & de Voogd, J. N. (2012). Are we witnessing the decline effect in the Type D personality literature? What can be learned? *Journal of Psychosomatic Research, 73,* 401–407. http://dx.doi.org/10.1016/j.jpsychores.2012.09.016

Crotty, D. (2014). When crises collide: The tension between null results and reproducibility. *The Scholarly Kitchen.* Retrieved from http://scholarlykitchen.sspnet.org/2014/09/10/when-crises-collide-the-tension-between-null-results-and-reproducibility/

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology.* Retrieved from http://journal.frontiersin.org/article/10.3389/fpsyg.2014.00781/full. http://dx.doi.org/10.3389/fpsyg.2014.00781

Dillio, R. (2014). A critical miss: Video games, violence, and ineffective legislation. *First Amendment Studies, 48,* 110–130. http://dx.doi.org/10.1080/21689725.2014.950496

Donnellan, M., Trzesniewski, K. H., & Robins, R. W. (2009). An emerging epidemic of narcissism or much ado about nothing? *Journal of Research in Personality, 43,* 498–501. http://dx.doi.org/10.1016/j.jrp.2008.12.010 **AQ: 4**

Doyen, S., Klein, O., Pichon, C.-L., & Cleeremans, A. (2012). Behavioral priming: It's all in the mind, but whose mind? *PLoS ONE, 7,* e29081. http://dx.doi.org/10.1371/journal.pone.0029081

Eaton, W. O. (1984). On obtaining unpublished data for research integrations. *American Psychologist, 39,* 1325–1326. http://dx.doi.org/10.1037/0003-066X.39.11.1325

Edwards, K., & Smith, E. E. (1996). A disconfirmation bias in the evaluation of arguments. *Journal of Personality and Social Psychology, 71,* 5–24. http://dx.doi.org/10.1037/0022-3514.71.1.5

Elson, M. (2014). How reliable is psychological science? Retrieved from http://blog.thingswedontknow.com/2014/07/how-reliable-is-psychological-science.html

Ertelt, S. (2006). American Psychological Association criticized for pro-abortion position. Retrieved from http://archive.lifenews.com/nat2090.html

Etchells, P. (2014). Psychology's replication drive: It's not about you. *The Guardian*. Retrieved from http://www.theguardian.com/science/head-quarters/2014/may/28/psychology-replication-drive-methods-bullying

Fanelli, D. (2010). "Positive" results increase down the Hierarchy of the Sciences. *PLoS ONE, 5,* e10068. http://dx.doi.org/10.1371/journal.pone.0010068

Fang, F. C., Steen, R. G., & Casadevall, A. (2012). Misconduct accounts for the majority of retracted scientific publications. *Proceedings of the National Academy of Sciences of the United States of America, 109,* 17028–17033. http://dx.doi.org/10.1073/pnas.1212247109

Ferguson, C. J. (2009). Is psychological research really as good as medical research? Effect size comparisons between psychology and medicine. *Review of General Psychology, 13,* 130–136. http://dx.doi.org/10.1037/a0015103

Ferguson, C. J. (2014). Is video game violence bad? *The Psychologist, 27,* 324–327.

Ferguson, C. J., & Brannick, M. T. (2012). Publication bias in psychological science: Prevalence, methods for identifying and controlling, and implications for the use of meta-analyses. *Psychological Methods, 17,* 120–128. http://dx.doi.org/10.1037/a0024445

Fergusson, D. M., Horwood, L. J., & Boden, J. M. (2008). Abortion and mental health disorders: Evidence from a 30-year longitudinal study. *The British Journal of Psychiatry, 193,* 444–451. http://dx.doi.org/10.1192/bjp.bp.108.056499

Fiedler, K. (2011). Voodoo correlations are everywhere: Not only in neuroscience. *Perspectives on Psychological Science, 6,* 163–171. http://dx.doi.org/10.1177/1745691611400237

Finkelhor, D. (2010, October). The internet, youth deviance and the problem of "juvenoia." Paper presented at the Justice Studies Colloquium, ●●●. Retrieved from https://vimeo.com/16900027

Firth, U., & Firth, C. (2014). A question of trust: Fixing the replication crisis. *The Guardian*. Retrieved from http://www.theguardian.com/science/occams-corner/2014/may/28/question-trust-fixing-replication-crisis-experimenter-reputation

Fowler, R. (1999). Letter to Tom Delay. Retrieved from http://www.leadershipcouncil.org/1/rind/apa.html

Francis, G., Tanzman, J., & Matthews, W. J. (2014). Excess success for psychology articles in the journal science. *PLoS ONE, 9,* e114255. Retrieved from http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0114255. http://dx.doi.org/10.1371/journal.pone.0114255

Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science, 345,* 1502–1505. http://dx.doi.org/10.1126/science.1255484

Freedman, J. L. (1984). Effect of television violence on aggressiveness. *Psychological Bulletin, 96,* 227–246. http://dx.doi.org/10.1037/0033-2909.96.2.227

French, C. (2012). Precognition studies and the curse of the failed replications. *The Guardian*. Retrieved from http://www.theguardian.com/science/2012/mar/15/precognition-studies-curse-failed-replications

Fritz, A., Scherndl, T., & Kühberger, A. (2012, April). *Correlation between effect size and sample size in psychological research: Sources, consequences, and remedies*. Paper presented at the 10th Conference of the Austrian Psychological Society, Graz, Austria.

Galván, A. (2013). The teenage brain: Sensitivity to rewards. *Current Directions in Psychological Science, 22,* 88–93. http://dx.doi.org/10.1177/0963721413480859

Goffin, S. G., & Myers, M. (1991). A landscape of concerns: A national survey of position papers on children's issues. *Child Welfare: Journal of Policy, Practice, and Program, 70,* 529–540.

Greenwald, A. G. (2012). There is nothing so theoretical as a good method. *Perspectives on Psychological Science, 7,* 99–108. http://dx.doi.org/10.1177/1745691611434210

Greenwald, A. G., Pratkanis, A. R., Leippe, M. R., & Baumgardner, M. H. (1986). Under what conditions does theory obstruct research progress? *Psychological Review, 93,* 216–229. http://dx.doi.org/10.1037/0033-295X.93.2.216

Grijalva, E., Newman, D. A., Tay, L., Donnellan, M. B., Harms, P. D., Robins, R. W., & Yan, T. (2015). Gender differences in narcissism: A meta-analytic review. *Psychological Bulletin, 141,* 261–310. http://dx.doi.org/10.1037/a0038231

Grisso, T., & Steinberg, L. (2005). Between a rock and a soft place: Developmental research and the child advocacy process. *Journal of Child and Adolescent Psychology, 34,* 619–627. http://dx.doi.org/10.1207/s15374424jccp3404_4

Grohol, J. (2014). Emotional contagion on Facebook? More like bad research methods. *PsychCentral*. Retrieved from http://psychcentral.com/blog/archives/2014/06/23/emotional-contagion-on-facebook-more-like-bad-research-methods/

Gutting, G. (2012). How reliable are the social science? The New York Times. Retrieved from http://opinionator.blogs.nytimes.com/2012/05/17/how-reliable-are-the-social-sciences/

Haggstrom, O. (2015). Intellectual suicide by the journal Basic and Applied Social Psychology. Retrieved from http://haggstrom.blogspot.com/2015/02/intellectual-suicide-by-journal-basic.html

Hall, R. C., Day, T., & Hall, R. C. (2011). A plea for caution: Violent video games, the Supreme Court, and the role of science. *Mayo Clinic Proceedings, 86,* 315–321. http://dx.doi.org/10.4065/mcp.2010.0762

Hendricks, V. (2014). Scientific research can be prone to bubbles too: Neuroscience risks being the next one. *The Conversation*. Retrieved from http://theconversation.com/scientific-research-can-be-prone-to-bubbles-too-neuroscience-risks-being-the-next-one-33797

Heusmann, L., & Taylor, L. (2003). The case against the case against media violence. In D. Gentile (Ed.), *Media violence and children: A complete guide for parents and professionals* (pp. ●●●). New York, **AQ: 6** NY: Praeger.

Hsu, L. M. (2004). Biases of success rate differences shown in binomial effect size displays. *Psychological Methods, 9,* 183–197. http://dx.doi.org/10.1037/1082-989X.9.2.183

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS. Med, 2,* e124. Retrieved from http://www.plosmedicine.org/article/info:doi/10.1371/journal.pmed.0020124

Ioannidis, J. P. (2012). Scientific inbreeding and same-team replication: Type D personality as an example. *Journal of Psychosomatic Research, 73,* 408–410. http://dx.doi.org/10.1016/j.jpsychores.2012.09.014

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. http://dx.doi.org/10.1177/0956797611430953

Johnson, D. J., Cheung, F., & Donnellan, M. (2014). Does cleanliness influence moral judgments? A direct replication of Schnall, Benton, and Harvey (2008). *Social Psychology, 45,* 209–215. http://dx.doi.org/10.1027/1864-9335/a000186

Kazdin, A., & Rotello, C. (2009). I spy daddy giving someone the finger: Your kids will imitate you, use it as a force for good. Retrieved from http://www.slate.com/articles/life/family/2009/01/i_spy_daddy_giving_someone_the_finger.html

Klein, O., Doyen, S., Leys, C., de Saldanha da Gama, P., Miller, S., Questienne, L., & Cleeremans, A. (2012). Low hopes, high expectations: Expectancy effects and the replicability of behavioral experiments. *Perspectives on Psychological Science, 7,* 572–584. http://dx.doi.org/10.1177/1745691612463704

Konrath, S. H., O'Brien, E. H., & Hsing, C. (2011). Changes in dispositional empathy in American college students over time: A meta-analysis. *Personality and Social Psychology Review, 15,* 180–198. http://dx.doi.org/10.1177/1088868310377395

Kraemer, H. C. (2005). A simple effect size indicator for two-group comparisons? A comment on r equivalent. *Psychological Methods, 10,* 413–419.

Kramer, A., Guillory, J., & Hancock, J. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences of the United States of America, 111,* 8788–8790. http://dx.doi.org/10.1073/pnas.1320040111

Lang, A. (2013). Discipline in crisis? The shifting paradigm of mass communication research. *Communication Theory, 23,* 10–24. http://dx.doi.org/10.1111/comt.12000

Lilienfeld, S. O. (2002). When worlds collide: Social science, politics, and **AQ: 7** the Rind et al. (1998) child sexual abuse meta-analysis. *American* **AQ: 8**

AQ: 5

Psychologist, 57, 176–188. http://dx.doi.org/10.1037/0003-066X.57.3.176

Lilienfeld, S. O. (2012). Public skepticism of psychology: Why many people perceive the study of human behavior as unscientific. *American Psychologist, 67,* 111–129. http://dx.doi.org/10.1037/a0023963

Lilienfeld, S., Lynn, S., Ruscio, J., & Beyerstein, B. (2009). Mythbusting in introductory psychology courses: The whys and the hows. *Essays from e-xcellence in teaching* (Vol. 9). Retrieved from http://teachpsych.org/ebooks/eit2009/index.php

Males, M. (2009). Does the adolescent brain make risk taking inevitable? A skeptical appraisal. *Journal of Adolescent Research, 24,* 3–20. http://dx.doi.org/10.1177/0743558408326913

Martel, M. M. (2009). The ethics of psychology's role in politics and the development and institution of social policy. *Ethics & Behavior, 19,* 103–111. http://dx.doi.org/10.1080/10508420902772694

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., . . . Reed, G. M. (2001). Psychological testing and psychological assessment. A review of evidence and issues. *American Psychologist, 56,* 128–165.

Mitchell, J. (2014). On the emptiness of failed replications. Retrieved from http://wjh.harvard.edu/~jmitchel/writing/failed_science.htm

Mook, D. G. (1983). In defense of external invalidity. *American Psychologist, 38,* 379–387. http://dx.doi.org/10.1037/0003-066X.38.4.379

Neuroskeptic. (2014). On "on the emptiness of failed replications." Retrieved from http://blogs.discovermagazine.com/neuroskeptic/2014/07/07/emptiness-failed-replications/

New York Times. (1994). Public health approach to crime urged. Retrieved from http://articles.sun-sentinel.com/1994-10-16/news/9410150290_1_gun-control-public-health-approach-dr-mark-rosenberg

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2,* 175–220. http://dx.doi.org/10.1037/1089-2680.2.2.175

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. http://dx.doi.org/10.1177/1745691612459058

Novella, S. (2015). Psychology journal bans statistics. *Real Clear Science.* Retrieved from http://www.realclearscience.com/2015/02/27/psychology_journal_bans_statistics_263538.html

Nyhan, B. (2014). To get more out of science, show the rejected research. *New York Times.* Retrieved from http://www.nytimes.com/2014/09/19/upshot/to-get-more-out-of-science-show-the-rejected-research.html?abt=0002&abg=0

O'Donohue, W., & Dyslin, C. (1996). Abortion, boxing and Zionism: Politics and the APA. *New Ideas in Psychology, 14,* 1–10. http://dx.doi.org/10.1016/0732-118X(95)00025-C

Olson, B., Soldz, S., & Davis, M. (2008). The ethics of interrogation and the American Psychological Association: A critique of policy and process. *Philosophy, Ethics, and Humanities in Medicine, 3,* 3. http://dx.doi.org/10.1186/1747-5341-3-3

Pashler, H., Coburn, N., & Harris, C. R. (2012). Priming of social distance? Failure to replicate effects on social and food judgments. *PLoS ONE, 7,* e42510.

Pedersen, W. (2007). Childbirth, abortion and subsequent substance use in young women: A population-based longitudinal study. *Addiction, 102,* 1971–1978. http://dx.doi.org/10.1111/j.1360-0443.2007.02040.x

Pinker, S. (2002). *The blank slate: The modern denial of human nature.* New York, NY: Penguin.

Przybylski, A. K., Rigby, C. S., & Ryan, R. M. (2010). A motivational model of video game engagement. *Review of General Psychology, 14,* 154–166. http://dx.doi.org/10.1037/a0019440

Redding, R. E. (2001). Sociopolitical diversity in psychology. The case for pluralism. *American Psychologist, 56,* 205–215. http://dx.doi.org/10.1037/0003-066X.56.3.205

Risen, J. (2014). *Pay any price.* Boston, MA: Houghton Mifflin Harcourt.

Ritter, D., & Eslea, M. (2005). Hot sauce, toy guns and graffiti: A critical account of current laboratory aggression paradigms. *Aggressive Behavior, 31,* 407–419. http://dx.doi.org/10.1002/ab.20066

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86,* 638–641. http://dx.doi.org/10.1037/0033-2909.86.3.638

Rosnow, R. L., & Rosenthal, R. (2003). Effect sizes for experimenting psychologists. *Canadian Journal of Experimental Psychology, 57,* 221–237. http://dx.doi.org/10.1037/h0087427 **AQ: 9**

Rouder, J. N., & Morey, R. D. (2012). Default Bayes factors for model selection in regression. *Multivariate Behavioral Research, 47,* 877–903. http://dx.doi.org/10.1080/00273171.2012.734737

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

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin, 138,* 628–654. http://dx.doi.org/10.1037/a0027473

Shooter, J. (2011). Unpublished results hide the decline effect. *Nature.* Retrieved from http://www.nature.com/news/2011/110223/full/470437a.html

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. http://dx.doi.org/10.1177/0956797611417632

Simons, D. J. (2014). The value of direct replication. *Perspectives on Psychological Science, 9,* 76–80. http://dx.doi.org/10.1177/1745691613514755

Stanovich, K. E., West, R. F., & Toplak, M. E. (2013). Myside bias, rational thinking, and intelligence. *Current Directions in Psychological Science, 22,* 259–264. http://dx.doi.org/10.1177/0963721413480174

Steinberg, L., & Monahan, K. C. (2011). Adolescents' exposure to sexy media does not hasten the initiation of sexual intercourse. *Developmental Psychology, 47,* 562–576. http://dx.doi.org/10.1037/a0020613

Stratford, M. (2014). Symbolic slap at social sciences. *Inside HigherEd.* Retrieved from http://www.insidehighered.com/news/2014/06/02/house-passes-nsf-funding-bill-takes-slap-social-sciences#sthash.hh40CGu6.dpbs

Task Force on Women, Poverty, and Public Assistance. (1998). Making "welfare to work" really work: Improving welfare reform for poor women, families and children. Retrieved from http://www.apa.org/pi/women/programs/poverty/welfare-to-work.aspx

Tear, M. J., & Nielsen, M. (2013). Failure to demonstrate that playing violent video games diminishes prosocial behavior. *PLoS ONE, 8,* e68382. http://dx.doi.org/10.1371/journal.pone.0068382

Tedeschi, J. T., & Quigley, B. M. (1996). Limitations of laboratory paradigms for studying aggression. *Aggression and Violent Behavior, 1,* 163–177. http://dx.doi.org/10.1016/1359-1789(95)00014-3

The New Republic. (2006). Fared well. *The New Republic.* Retrieved from http://www.newrepublic.com/article/fared-well

Trafimow, D., & Marks, M. (2015). Editorial. *Basic and Applied Social Psychology, 37,* 1–2. http://dx.doi.org/10.1080/01973533.2015.1012991

Twenge, J. M., & Campbell, W. (2009). *The narcissism epidemic: Living in the age of entitlement.* New York, NY: Free Press.

van Lent, M., Overbeke, J., & Out, H. J. (2014). Role of editorial and peer review processes in publication bias: Analysis of drug trials submitted to eight medical journals. [Advance online publication]. *PLoS ONE, 9,* e104846. http://dx.doi.org/10.1371/journal.pone.0104846

Wagenmakers, E. J., Wetzels, R., Borsboom, D., & van der Maas, H. L. (2011). Why psychologists must change the way they analyze their data: The case of psi: Comment on Bem (2011). *Journal of Personality and Social Psychology, 100,* 426–432. http://dx.doi.org/10.1037/a0022790

Wagenmakers, E., Wetzels, R., Borsboom, D., van der Maas, H. J., & Kievit, R. A. (2012). An agenda for purely confirmatory research. *Perspectives on Psychological Science, 7,* 632–638. http://dx.doi.org/10.1177/1745691612463078

Wai, J. (2012). Can psychology be considered a science? Psychology Today. Retrieved from http://www.psychologytoday.com/blog/finding-the-next-einstein/201206/can-psychology-be-considered-science

Weeks, N. (2005). A few minutes of your time: Tips on communicating scientific information to the media. *The Observer.* Retrieved from http://www.psychologicalscience.org/index.php/uncategorized/a-few-minutes-of-your-time-tips-on-communicating-scientific-information-to-the-media.html

Wicherts, J., & Bakker, M. (2009). Sharing: Guidelines go one step forwards, two steps back. *Nature, 461,* 1053. http://dx.doi.org/ 10.1038/4611053c **AQ: 10**

Wicherts, J. M., Bakker, M., & Molenaar, D. (2011). Willingness to share research data is related to the strength of the evidence and the quality

of reporting of statistical results. *PLoS ONE, 6,* e26828. http://dx.doi.org/10.1371/journal.pone.0026828

Wicherts, J. M., Borsboom, D., Kats, J., & Molenaar, D. (2006). The poor availability of psychological research data for reanalysis. *American Psychologist, 61,* 726–728. http://dx.doi.org/10.1037/0003-066X.61.7.726

Willoughby, T., Good, M., Adachi, P. J., Hamza, C., & Tavernier, R. (2013). Examining the link between adolescent brain development and risk taking from a social-developmental perspective. *Brain and Cognition, 83,* 315–323. http://dx.doi.org/10.1016/j.bandc.2013.09.008

Wilson, A., & Golonka, S. (2012). *The small effect size effect: Why do we put up with small effects?* Retrieved from http://psychsciencenotes.blogspot.com/2012/08/the-small-effect-size-effect-why-do-we.html

Wolfe, J., & Roberts, C. (1993). A further study of the external validity of business games: Five-year peer group indicators. *Simulation & Gaming, 24,* 21–33. http://dx.doi.org/10.1177/1046878193241004

Wolins, L. (1962). Responsibility for raw data. *American Psychologist, 17,* 657–658. http://dx.doi.org/10.1037/h0038819

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