From crisis of evidence to a "crisis" of relevance? Incentive-based answers for social

psychology's perennial relevance worries

Roger Giner-Sorolla

University of Kent

Contact information: Email rsg@kent.ac.uk; School of Psychology, University of Kent, Keynes College, Canterbury, Kent CT 27NP, United Kingdom

Abstract

Current controversies in social psychology have sparked the promotion of new rules for evidence in the field. This "crisis of evidence" echoes prior concerns from the 1970's about a so-called "crisis of social psychology", with such issues as replication and statistical significance once more under examination. I argue that parallel concerns about the *relevance* of our research, raised but not completely resolved in the 1970's crisis, also deserve a fresh look. In particular, the advances made in the current crisis of evidence came about because of changes in academic career incentives--particularly publishing. Today, many voices in psychology urge greater respect for relevance in topics, methods and communication, but the lack of clear and concrete incentives to do so has stood in the way of answers. I diagnose the current incentive structures, propose partial solutions that are within the reach of journal editors and professional societies, and conclude by discussing the links between relevance and evidence, as well as special challenges to the relevance of social psychology post-2016.

From crisis of evidence to a "crisis" of relevance? Incentive-based answers for social

psychology's perennial relevance worries

> If the multitude of social-psychological findings cannot aid the planners of society, it is apparently not because we have been researching the wrong topics. It must be that our data are not generalizable to the objects of our studies in their natural, ongoing states. This is a basic inadequacy of methodology rather than direction, and it will not be resolved by pontifical edicts from any source about what to study and where. -- Silverman (1971), p. 583

Whether social psychology research can be relevant to anyone but social psychologists is a question often put in the form of a "relevance crisis." Can our ways of research create solutions to real-world social problems? Do our findings reach a wider audience, or echo in a sealed chamber? Crisis questions of many different kinds have reared up in social psychology, time and again, first in the 1930's, and reaching a peak in the early 1970's. Recent years have seen continued criticism of the discipline of social psychology for failing to value research in applied settings, for relying on limited populations and artificial measures, and for not effectively communicating a great deal of its findings to the public. Regardless of whether the fault is seen in the content of the research itself, its methodology, or its dissemination, the common element in relevance-based critiques is a failure to connect theory with the lives of everyday people. Psychology seems to draw recurring criticism as a discipline that fails to deliver things of interest and utility to end-users and members of the public, whether the fault is in the ecological validity of the research itself, its applicability to problems, or its dissemination.

Whether or not these criticisms are justified, let alone constitute a "relevance crisis," is difficult to establish, and quite subjective. It is easy to point to success stories, but harder to judge how typical they are. It is, however, clear that highly respected figures in psychology—

best-selling authors, presidents of professional organizations—are prone to expressing discontent that relevance is not more respected within academic psychology, as I will show in more detail. Thus, by definition, relevance in psychology has an image problem. However, it is not often that these criticisms include concrete suggestions to change the incentives in the field. As a model for change, I will look at another avenue of criticism of social psychology often called a "crisis" – the recent focus on reforming standards of evidence, including changes to reporting and statistics, and increased promotion of replication. We can label this set of issues the "evidence crisis," in contrast with the "relevance crisis" which I will later describe.

A brief overview of the "evidence crisis" and what it has changed

The first appearance of a crisis of evidence in social psychological research came in the 1960's and 1970's. Social psychologists began to express doubts about the methodological validity of their research: this was evident, for example, in debates on the usefulness of the attitude construct in predicting behaviour (Wicker, 1964), and the influence of experimenter demand both overt and subtle (Orne, 1962; Rosenthal, 1963). These doubts shook the field, generating solutions as radical as to abandon the natural science paradigm in psychology (Gergen, 1973; N. C. Smith, 1970). Others urged greater sophistication within existing paradigms, such as a more precise re-specification of the attitude-behaviour problem (Ajzen & Fishbein, 1977).

However, the criticism with the most resonance today went beyond the specific methods of social psychology, and focused on the common practice of declaring an effect to be true based on a single statistically significant null hypothesis test of $p < .05$. This criterion of truth, because it encouraged publication and analytic bias in favour of significant results, was seen as risking unsupported conclusions (e.g., Sterling, 1959; Lykken, 1968). At the same time, some authors of

literature reviews observed that single published experimental results often failed to replicate, leading to an overall equivocal literature (e.g., Sears & Freedman, 1967, on selective exposure effects). Numerous solutions were proposed in psychology journals: publishing null as well as significant results (Greenwald, 1975), improving the statistical power of studies (Cohen, 1962), reporting power, effect sizes and confidence intervals (Tversky & Kahneman, 1971), and internal replication of effects within a lab (Lykken, 1968; N. C. Smith, 1970). Social and personality psychologists such as Sears, Lykken, and Greenwald were well-represented among these critics and reformers.

Through the following decades, it became clear that many of the concerns had not led to real changes in reporting. Statistical power of typical experiments did not greatly improve (Sedlmeier & Gigerenzer, 1989; Fraley & Vazire, 2014). While replications were urged in theory (e.g., Fishman & Neigher, 1982), papers that only replicated existing results were not easily published (Makel, Plucker & Hegarty, 2012), and setting up special journals for them was not a great success (Giner-Sorolla, 2012). Reporting of effect sizes was also rare (Greenwald, Gonzalez, Harris, & Guthrie, 1986). The long-awaited recommendations of the APA Task Force on Statistics (Wilkinson and APA Task Force, 1999) were not generally put into practice across psychology journals (Cumming, Fidler et al., 2007; Vankov, Bowers & Munafo, 2014). The APA recommendations included requests to report effect sizes for all major tests, eschew post-hoc testing, refrain from making causal claims on the mere basis of multivariate modelling, and over 70 other desiderata of publishing. Ten years after the Task Force report, out of the 10 most cited articles in APA's *Journal of Personality and Social Psychology* in 2009 as determined by Google Scholar (April 3, 2017), four reported no effect size or confidence interval for their main

findings, and one reported effect sizes for some tests but not others. One has to wonder whether the comprehensiveness of these guidelines was a hindrance to their implementation.

And yet after the first crisis social/personality psychology did adopt one change of practice to quell doubts about significance testing. Articles with multiple studies, each replicating a common hypothesis with a significant effect, became the norm at the most prestigious journals. Increasingly, journal editors asked for additional studies supporting the point, or asked single-study authors to come back when they had more evidence. Although not often enshrined in explicit policies or editorials, the new quasi-requirement was visible enough to complain about (Wegner, 1992), and spread through word-of-mouth communication. This hand on the levers of journal publishing changed the look of social psychology. For example, the *Journal of Personality and Social Psychology* (*JPSP*) increased its proportion of multi-study articles steadily from the 1970's and ended up in 2007 with an average of 3.3 studies per article (Reis & Stiller, 1992; Witte & Brandt, 2009).

The story of the second evidence crisis in 2011 has been told more comprehensively elsewhere (Spellman, 2015). Briefly, after Bem's (2011) nine-study paper purporting to establish the existence of precognition was published in *JPSP*, the multi-study solution seemed an inadequate safeguard against false positives. That year, too, several high-profile fraud cases in social psychology threw kerosene on the fires of reform. Subsequent analyses showed that selective reporting of analyses and studies was almost certainly going on (Schimmack, 2012), that it could be a problem for inference even in multi-study papers (Simmons, Nelson & Simonsohn, 2011), and that the problem was aided by the exclusive evaluation of research by *p*-value thresholds, ignoring statistical power and effect sizes (Cumming, 2014). Also, concerns in the 2000's that multi-study publishing slowed down the field had led some journals, such as

Psychological Science and *Journal of Experimental Social Psychology* (*JESP*), to introduce

formats for short reports of single studies, partially repealing the multi-study solution. The short-

articles trend, too, had its critics (Ledgerwood & Sherman, 2012).

Post-2011, however, much more attention than before has been paid to the role of

publishing in moving the discipline past incomplete reporting and irreproducible results (e.g.,

Nosek & Bar-Anan, 2012). Suggestions for change have been kept simple (e.g., 21- or 18-word

statements enforcing disclosure of all measures, manipulations, and exclusions; LeBel,

Borsboom, et al, 2013; Simmons, Neilson & Simonsohn, 2012) and concrete (e.g., a "new

statistics" requiring reporting of confidence intervals around effect sizes; Cumming, 2014). The

Center for Open Science (COS) has codified the Transparency and Openness Principles (TOP)

guidelines and Open Science badges, which spell out concrete practices that journals can

officially adopt at various levels (Nosek, Alter, et al., 2015), such as complete reporting of

results, sharing of data and materials, and encouragement of replications and pre-registration.

The COS has also chosen to hand out incentives in the most direct way possible, by offering one

thousand grants of one thousand dollars to researchers across the sciences who pre-register their

studies (Preregistration Challenge, 2017).

Uptake of these proposals has been mixed, with journals supporting guidelines more in

principle than in practice (Center for Open Science, 2018). However, enough time has passed

post-crisis to observe that journal editorial teams now usually feel a need to do something

different about evidence: requiring disclosure of all measures and manipulations, moving away

from the $p < .05$ criterion, requesting verbatim materials for review, requiring confidence

intervals, asking for statistical power to be addressed in each article, requiring a certain level of

statistical power, and opening the door to direct replications and preregistered studies (for

various subsets of these innovations, see these selected editorials: at *JESP*, Giner-Sorolla, 2016; at *JPSP*, Kitayama, 2017 and Cooper, 2016; at *PSPB*, Crandall, Leach, Robinson and West, 2017; at *Psychological Science*, Eich, 2014 and Lindsay, 2015; at *SPPS*, Vazire, 2016). At a couple of years' distance from the earliest of these changes, there is some evidence that adopting them has moved publication in the desired direction (Tressoldi, Cumming, Fresc, & Giofré, 2016).

In an earlier essay (Giner-Sorolla, 2012) I gave evidence for a narrowing "bottleneck," pitting increasing numbers of social psychologists against much more slowly increasing numbers of journal and career slots (see also Pratkanis, 2017). The "crisis of evidence" has shown us that under severe competition, norms have a chance to change only when practices visibly enable the immediate achievement of markers of excellence such as grants, awards, jobs and articles. But norms have little chance to change when incentives are distant and vague; are couched in terms of approval rather than benefit; and ask researchers to act individually in taking unrewarded risks for the future good of the discipline. This lesson, I will argue, also applies to issues of social psychology's relevance.

Three facets of relevance criticism, in history

Writings mentioning a relevance "crisis" in social psychology date back to before the Second World War. Throughout this literature, evidence and relevance concerns have appeared in parallel. An early polemic (Britt, 1937) noted that students yawn at abstruse, idiosyncratic armchair theories but show far more interest in empirical answers to social questions -- "'To what extent do newspapers print "true" news?' 'What are the underlying myths and legends that cause friction between Gentile and Jew?'" (p. 464). Evidently, the shift towards empirical

methods was so profound that it is today taken for granted. Cantril's (1934) concerns were more methodologically relevant to today, and critically hinged on the relevance of research to everyday life: "The choice lies between thorough studies on authentic problems, conducted with the best techniques that can be devised, and controlled laboratory experiments on problems which have no relation to life itself" (p. 298).

After a couple of confident decades in research, producing many now-classic laboratory experiments, the crisis that Cantril anticipated came to life again. Elms (1975) reviewed over a dozen writings, mounting up over the previous decade, that proclaimed a "crisis stage" in social psychology--a term Elms attributed to Berkowitz (p. 967), one which also resonated with the larger 1970's zeitgeist of social and economic crises in Western society at large (Faye, 2012). Elms put the key issues under three headings. "Research difficulties" referred to evidence and interpretation issues of the kind we have already seen: generalizability, reproducibility, potential confounds in research. "Researcher expectancies" referred to the unrealistic expectation that social psychology can be tackled with simple theories and paradigms. Finally, relevance appears under the third heading, "outside pressures," including pressure to apply knowledge to social problems, coming from the many movements for ethnic, gender, sexual, and colonial liberation that flourished in the 1960's. "Outside pressures" also included concerns about the treatment of research participants, and a concern that relevance, in the shape of applied research, had become *too* important in the face of Federal cuts to research on basic questions in the United States.

Elms' article, and other literature of the period, suggests a scheme for subdividing relevance concerns.

- **Q**uestion relevance is a concern about the limited applicability of theories and associated methods to social and human problems. For example, the worry often arises that basic lab research cannot easily be reproduced under messier field conditions, or is not amenable to crafting effective interventions.

- S**ample relevance** is a concern about the limited generalizability from findings based on research samples to other human populations. Most often, this involves the recurring concern that research has overused samples of convenience, which can give a culturally, educationally, and demographically skewed picture of psychology. Social psychology is particularly vulnerable to sample relevance concerns, because it deals with culturally and socially influenced topics such as attitudes, prejudice and politics.

- **Communication relevance** is a concern that we are not letting laypeople know about psychological research enough, or in the right way. For example, one might worry that our research is often written in jargon, that it is not represented when pundits are called upon, or that it is misrepresented and oversimplified if it does reach the public.

In the following sections I will describe the history and present status of appeals to each relevance concern. In particular, I hope to show that influential voices in social psychology have pointed out deficiencies in each standard of relevance up to the present day.

Question relevance

Question relevance has been at the heart of struggles about how to define the field of psychological research. In most sciences, "basic" and "applied" research are differentiated from each other in methodological priorities (lab vs. field setting; conceptual vs. ecological validity) and in purpose (to create widely applicable theories vs. to solve a specific problem; Hoffman & Deffenbacher, 1993). By some accounts, North American psychology had turned toward valuing basic research and away from application in the early 20th century, when Titchener and Boring took the helm of the American Psychological Association (O'Donnell, 1979).

More broadly, post-World War II developments in the Western world saw strong governmental investments in basic science for a variety of reasons, including support for democratic and enlightened values, and the need to anticipate world-changing technological development at the basic science level (Pielke, 2012). But the decades since then have seen challenges to basic science from both the political left and right. Seeking fundamental insights has been criticized as irrelevant, both to pressing social problems and to the pragmatic use of public funds. In the 1970's, in particular, psychology saw many calls for greater development of ecologically valid and applicable methods. While some contributors to the dialogue defended basic theory development (e.g. Smith, 1973), others called for the experiment to be seen as a specific context that may not generalize, rather than a default basis for knowledge (e.g., Silverman, 1971; Sherif, 1970). Helmreich (1975) called applied social psychology an "unfulfilled promise," pointing out that in the then-new *Journal of Applied Social Psychology*, most articles were lab experiments. While they were possibly relevant to applied issues, these studies, in Helmreich's view, did not show the kind of external validity that might come from field settings or tests of applications.

Another critical trend around the same time, mostly among European psychologists, questioned the epistemic positivism, mechanism, and individualism underpinning the laboratory experiment. These critics offered more radical solutions to question relevance (see Parker, 1989, for a review, as well as Rijsman & Stroebe, 1989 and other papers in that issue of the *European Journal of Social Psychology*). For example, some argued for abandoning nomothetic aims, in favour of a more idiographic social psychology based on methods used in history (Gergen, 1973). Some drew on theory and methodology from linguistics, arguing for more sensitivity to the intentions and context of the persons studied (Harré & Secord, 1972). Others promoted a focus on interactions rather than individual processes in explaining social phenomena (Moscovici, 1972). These insights gave rise to a parallel school of critical social psychology that has intersected with positivistic social psychology mainly in European publishing and academic circles, and even then, it has to be said, without a great deal of mutual understanding.

In the post-crisis environment, voices from the non-Western world were also heard, working out how their perspectives might fit into the disagreement between American and European social psychologies. Moghaddam (1987) identified a more sensitive approach to cultural issues and a greater need for economic and social relevance as potential features of a "Third World" psychology. South Africa presented an example in the 1990's and 2000's of a country undergoing political upheaval where "relevance" had been a near-constant watchword among psychologists (de la Rey & Ipser, 2004). Long and Foster (2013), through analysis of South African professional addresses, showed how the political tone of calls for relevance changed from supporting the anti-apartheid struggle to adjusting to the new realities of a majority-rule market society. This well-taken analysis reminds us that "relevance" of research

questions is in itself a morally and politically ambiguous topic. It can as easily be attained by giving advice to the exploiter, as by giving succour to the exploited.

Historians and sociologists of 20th century science have described basic-applied relations as potentially following two models: one-way influence, in which basic science provides ideas for a wide range of applications, or two-way influence, in which insights from application additionally feed back into and inform basic science (Godin, 2006; Roll-Hansen, 2017; see also Medin, 2012). In social psychology, Lewin's (1951) vision of a program of *action research*, with mutually influencing theoretical and applied strands, is an often-cited example of the two-way ideal. However, the reality of status relations in psychology tends to be more one-way. Basic research gains more respect and attention in the field at large because it is "upstream," influencing applications and behavioural technology, but without a reciprocal requirement for basic researchers to stay aware of where and how their ideas are being used. Social psychology researchers may recognize the negotiation of the basic-applied boundary revealed in the interviews of Calvert (2006) with physicists and biologists. Among themselves, these scientists valued basic research most highly, but saw a need to present it with an applied spin to funders.

Prioritizing basic research in social psychology can happen even though applied work is frequently done, and even though it leads to important developments in practice (for example, in eyewitness identification, Wells, Memon & Penrod, 2006; for a more general overview, see Maton, 2017). Indeed, it is not my intention to downplay the actual importance and recognition of applied research. But we also should not underestimate a fundamental asymmetry in prestige. The friction between a de facto one-way model and an ideal Lewinian two-way model in social psychology regularly leads to public expressions of concern, from society presidents and other influential figures, about the low prestige of application in psychology (e.g., Buschini, Guimond,

& Breakwell, 2011; Gergen, 1973; Leary, 2015; Medin, 2012; Pettigrew, 2018). But these are seldom backed up by calls to change the structure of the one-way flow.

Consider one article reporting an address by the then-president of the Society for Personality and Social Psychology (Snyder, 1993). Snyder gives examples from lab and field work with colleagues, showing how social psychology, by developing theory based on specific motivations, can improve the recruitment and retention of volunteers. He cites Lewin and argues that this research is an example of ideas flowing between basic and applied settings. The model of volunteerism thus developed was certainly highly useful and translatable to other applied domains where people go out of their way to help others. Indeed, the basic element, drawn from motivational and self psychology, is easy to trace into practice.

But it is harder to see from this account how the applied phase informed the basic theory underlying the research. Reading between the lines, some of the dimensions in the model were initially derived from open-ended interviews as well as existing psychological ideas, so that mid-level theory at least took some inspiration from the field setting. However, Snyder's address does not make clear how the insights from practice have fed back to basic work on the multiple theories that convergently helped Snyder and colleagues to understand volunteering. This example is but one indication of the greater prestige of theory: even praise and recognition given to applied work often takes place following the one-way rather than the two-way model.

Another indication is the frequent selective quotation in psychology of Kurt Lewin's "There is nothing so practical as a good theory" (1951). As Craig (1996) notes, the quotation is often used "as a handy device for asserting that academic theory has practical importance in the world while avoiding any careful analysis [...]" (p. 65). Billig (2015) likewise notes that many textbooks use the phrase to show that social psychology is self-evidently relevant to the applied

world, rather than to open discussion of whether and how it is relevant as practiced. Few uses of this quotation, however, have gone on to mention the return path that Lewin specified in the same article, by which there is also nothing so theoretical as a good practice (with some exceptions, mostly from applied psychology: e.g., Labonte, Feather, & Hills, 1999; Lévy-Leboyer, 1988; Meyer, 1991).

In mainstream North American social psychology, the goal of relevance has been defined mainly as wider application, rather than adopting the more contextualized methods demanded by European critical responses. Specifically, the mid-2000's saw a surge of interest in research that was "translational," defined as directly concerning outreach and interventions, in response to US funding agency shifts in priorities (Markus, 2004a; Major, 2006). In Europe, the applied, policy-driven focus of the European Union's Framework Programmes had for some time provided incentives for researchers to follow translational goals (e.g., European Commission, 2002), while individual national bodies varied in their relative weightings of basic and applied research.

In this discourse, some viewpoints on psychology's applied potential have been sanguine, with Breckler (2006a) declaring that psychology was already translational, even while observing in the same journal and year that applied research suffered from diminished prestige in some eyes (Breckler, 2006b). Tashiro and Mortensen (2006) were more critical, in particular of social psychology. They showed in detail how methodological habits of basic research (e.g., its reliance on short- versus long-term intervention and single-session measurement) had reduced its credibility among applied researchers, while these habits applied as standards for evaluating research had likewise reduced the perceived prestige of applied research among basic researchers, and they urged each side to bend a little toward the other in a "translational" spirit. Major (2006) also noted that the kind of methods and samples privileged in basic social

psychology have been ill-suited to impressing funders of translational research. Others during this time voiced concerns that social psychology had moved away from measuring real behaviour (Baumeister, Vohs & Funder, 2007), taking aim both at explicit self-reports and implicit reaction-time measures, and arguing that concrete behavioural outcomes are more valid and convincing to an external audience.

More recent addresses and editorials show that issues of question relevance are as alive as they ever were -- but also that their solutions seem to be as far away from the mainstream as ever. In a meeting on applying behavioural science to public policy (Keller-Allen & Li, 2013), the behavioural economist Thaler opined that "[c]urrently the type of psychological research that is useful in the real world is not the type that is valued in the academic profession of psychology or in top journals." (p. 45). The psychologist Kahneman agreed, adding "[t]hen and now, applied psychology is not attracting the best talent or held in high regard at the best universities." (p. 46). It is possible to take a more optimistic view of the applicability of social psychology, especially by comparison to other social sciences. But enough important voices continue to draw attention to the reputational disadvantage of applied research that these concerns should be looked at seriously.

Sample relevance

Relying on samples of undergraduate students, while presuming generalizability to broader populations, was a habit of social psychological research that began to look bad in the 1960's. Because undergraduate populations in the USA and elsewhere were explicitly or de facto segregated by class, gender and race, the practice underrepresented the categories of people who were claiming the right to be heard in those days. More generally, it skewed psychology towards conclusions that might not apply to older, less affluent, or less educated populations, let alone

people outside the Euro-American world. Sample relevance was mentioned by various authors in the crisis literature of the 1960's and 1970's, including Elms (1975) and Sherif (1977), but often tangentially. Strong concern over sample relevance was redoubled in the 1980's, with two critical salvoes in the *Journal of Personality and Social Psychology*: first from Carlson (1984), who also criticized the relevance of lab research settings, and then from Sears (1986), who focused on social psychology's dependence on undergraduates, documenting both the prevalence of the practice and its likely drawbacks. In general, sample relevance can be seen as more of a concern the less "basic" the psychological processes in play, and the more socially and culturally influenced they are.

Does sample relevance make a difference? Reviewing meta-analyses of consumer research that could compare student and representative samples, Peterson (2001) noted that in 19% of cases one sample yielded an effect in the opposite direction than the other. In an additional 32% of cases, results were in the same direction, but one sample's effect size was twice or more that of the other (with student samples, in general, being less variable and yielding larger effects). Weiner, Kraus and Lieberman (2011) offer similar cautions on the basis of a number of systematic comparisons of mock juries composed of students versus community members. However, other scholars have argued that limitations of student-based research are minimal; for example, Druckman and Kam (2011) propose that failures to generalize can be overcome by identifying relevant variables on which student and other populations differ and including those variables as moderators or covariates in a common model.

The typical research sample in social psychology has changed since the 1970's. The 2010's have seen the rise of crowdsourced participants as a widely accepted alternative to the undergraduate sample of convenience -- mainly workers on Amazon Mechanical Turk (AMT),

although new data sources are being developed, such as Prolific Academic and Crowdflower (Peer, Samat, Brandimarte & Acquisti, 2017). Indeed, in major social/personality psychology journals, online samples have gone from being extremely rare in 2005 to comprising nearly 50% of study samples in 2015 (Anderson, Allen et al., 2018). By comparison, in a more systematic look at two journals in the pre-AMT era, 20 years after Sears' critique, Witte and Brandt (2008) showed that 87.5% of studies in *JPSP*, and 92.6% of studies in the *European Journal of Social Psychology* (*EJSP*), sampled student participants. Today, the monoculture of undergraduates decried by Sears has been replaced by a duo-culture of undergraduates and crowdsourced workers. If this is an improvement over the mainly-students landscape of research participation, it is still a long way from a truly diverse selection of research samples.

Have online samples solved our sample relevance problems? Initial validation studies gave positive reviews for AMT workers' data. These participants seemed accurate, attentive, and more diverse in age, gender and career than typical psychology students (Berinsky, Huber, & Lenz, 2012, Buhrmester, Kwang, & Gosling, 2011; Paolacci, Chandler, & Iperiotis, 2010). Also, Coppock (2018) and Mullinix, Leeper, Druckman and Freese (2015) have separately shown that AMT samples give very similar results to nationally representative samples in survey experiments drawing on political and social psychology for their hypotheses.

However, doubts have also emerged. The active population of AMT workers doing studies at any one time turns out to be only 7,300 or so (Stewart, Ungemach, et al., 2015), adding to concerns that worker data may be 'too good'--i.e., invalid--because of non-naïveté about the tasks and theories involved. Workers are familiar with common paradigms and frequently discuss experiments amongst themselves (Chandler, Mueller, & Paolacci, 2014). Making use of less intensively used platforms may reduce this concern (Peer et al., 2017), but then again,

naïveté is likely to reduce as more researchers use them. Lack of laboratory control over online participants also raises problems with some paradigms. For example, post-manipulation drop-outs can seriously bias conclusions if one condition seems more difficult than another to participants (Zhou & Fishbach, 2016). Online studies also are limited in ability to create an absorbing lab environment, especially in social psychology, whose key procedures often include personal interaction. Methodological limitations of online studies have likely affected the kind of research that is published in other, not necessarily positive ways (Anderson et al., 2018).

Additionally, renewed attention has focused on the cultural limitations of samples of convenience. Henrich, Heine and Norenzayan (2010) coined the memorable acronym WEIRD (Western, Educated, Industrialized, Rich, Democratic) to describe the socio-cultural background of most participants in psychological studies. While such samples are conveniently close at hand for researchers who are themselves WEIRD, Henrich and colleagues argued that they do not represent the life of most human beings -- now or ever -- and lead to skewed assumptions about human nature, even more so if the participant population is drawn primarily from specific groups (e.g., USA nationals) and social strata. Given that AMT workers from the USA are more educated and liberal-leaning than the population at large (Paolacci & Chandler, 2014), this source of data looks to provide no respite from the WEIRDness of convenience sampling. An increasingly common solution is to take advantage of the strong representation of Indian as well as US workers in Mechanical Turk and conduct cross-cultural research comparing the two nationalities (e.g. Eriksson, Coultas & de Barra, 2016; Nishi, Christakis, & Reed, 2016). But the representativeness of Indian crowdsourced workers within their own culture has not yet been studied comprehensively, and again, this solution only replaces a monoculture with, as it were, a two-culture culture.

Ultimately, whatever the actual validity of the convenience sample, its use also poses a problem of public relations for the discipline. Returning to Major's (2006) argument, a science based on undergraduates (or hard-driven semi-professional research participants) is vulnerable to criticism from observers, who might well wonder whether any new finding will generalize to their special population of interest. Even if most findings do generalize, some will not. A science that can point to general-population and cross-cultural replications of its major findings is one that will have an easier time making its case to others.

Finally, social psychologists are only recently awakening to the possibility that not just our participants, but our stimuli and measures, might be non-representative, echoing Carlson's (1984) joint critique of samples and methods. Statistical solutions, such as treating stimulus characteristics as random variables, provide one way to test generalizability but will usually require an increase in participant numbers and/or study length to make up for lost statistical power (Judd, Westfall & Kenny, 2012). Method relevance concerns are also sometimes awakened by a "paradigm monoculture" in which one arbitrary or artificial-seeming scenario comes into vogue to the exclusion of more realistic alternatives, such as the trolley dilemma vignette in moral psychology (Bauman, McGraw & Bartels, 2014).

Communication relevance

In the 1970's crisis, another worry was that social psychologists had isolated themselves from reality, expressing themselves in language not easily understood by non-specialists. Miller (1969), in addressing the American Psychological Association, coined the durable phrase, "giving psychology away," echoed in the titles of scores of books and articles since then. Miller's address argued that more psychologists needed to take on the job of telling findings to the public rather than to fellow scholars. Re-reading Miller's address is worthwhile today

because it highlights a more radical message often lost in the rote repetition of his signal phrase: that, ethically, politically and pragmatically, people outside academia are the best placed to apply psychological insights. Following this observation, he argued that psychologists should focus on communicating well-supported principles to that audience, rather than trying to craft interventions themselves.

An emphasis on communication also informed the immediately subsequent literature on relevance and application. Thoreson, Krauskopf, McAleer and Wenger (1972) decried the "circular communication system with its own language (jargon)" (p. 137) which prevailed even in applied psychology, and suggested a role for professional psychologists as "translators" between applied researchers and practitioners. Smith's (1973) essay on relevance included a sharp look at the "meta-psychologies" or philosophical messages of such theoretical stances as behaviourism and humanistic psychology, with the conclusion that a meta-psychology that sends an arrogant, controlling message to the public would be neither popular nor desirable. Communication, to these writers, offered a way to influence the public that would be more gentle, humanistic, and democratic than the human engineering envisioned by behaviourists and some social psychologists themselves.

Communication relevance today might be seen as a resolved concern, because social and personality psychology has been in the public eye over the past decades as never before. Gladwell's popular books on behavioural research themselves represent a "tipping point" (Gladwell, 2000) opening a flood of interest in social psychology findings, through media both conventional and new (e.g., TED talks, online articles at such outlets as Vox.com and Slate.com). Academic social and cognitive psychologists have also taken to writing popular trade books on their own research and expertise. Looking at the Amazon.co.uk top 100 best sellers in

"Psychology and Psychiatry" at the time of first writing (30 March, 2017), one finds such familiar names as Kahneman (#8), Dweck (#19), Cialdini (#40), and Feldman Barrett (#97), alongside Gladwell's own *Outliers* (#56). Notably, the concept of "implicit bias" developed in social psychology was mentioned by Hillary Clinton during the 2016 US Presidential debates, and continued to draw attention in the 2018 response of the Starbucks coffee chain to a racial bias incident involving its employees.

Nonetheless, even with a good amount of communication, there comes the risk of miscommunication. As much as social psychologists have been able to control the message by being their own authors and spokespeople, they often still have to communicate via journalists, popular press authors and university press releases. This can lead to inflated claims, wrong claims, and oversimplification, which in turn discourages any further search for communication relevance (Gernsbacher, 2006). Often, too, social psychology research is rebranded ("The Behavioural Insights Team is widely perceived as doing behavioural economics," says Daniel Kahneman. "They are actually doing social psychology" -- Harford, 2014), so that findings are communicated to the public without being linked to our discipline, leaving people in the dark about what we do. Finally, the message may not be getting out as much as we would like; Lilienfeld (2012) cites the perception that psychology research is not useful as an overarching reason why the public doubts psychology's status as a science.

Not only the media are to blame. Social psychologists with very different opinions on the crisis of evidence have voiced concerns about publishing attention-grabbing, media-friendly research with surprising or "sexy" results -- recapitulating Ring's (1967) concerns about "'fun-and-games' social psychology" at the onset of the previous crisis. Some define the fault of these studies as abandoning rigorous methodology (e.g., Asendorpf, Conner et al., 2013), others as

abandoning rigorous theorizing (e.g., Ellemers, 2013; Schaller, 2016). However, it is possible to agree with both points at once.

Against this background, one might wonder whether calling for more communication might be premature, until we can agree on things like standards of evidence and replicability. One concern is that early reporting of results might later be negated by evidence produced under improving standards of knowledge, leaving the public confused. For example, research into the psychological effects of assuming a confident bodily posture (Carney, Cuddy & Yap, 2015) looked like a role model for public engagement earlier in the decade. Social psychologist Amy Cuddy became a widely known public-facing star of the field through TED talks and network television appearances. But when a larger, independent study failed to replicate some of these effects on risk-taking and hormones (Ranehill, Dreber et al., 2015), and a key member of the lab stated her own lack of belief in these effects (Carney, 2016), media accounts called the whole enterprise into question (e.g., Peters, 2016). It seems that debunking narratives sell just as well as surprising effects, even though Ranehill et al. (2015) and further replication efforts (Cesario, Jonas, & Carney, 2017) have replicated the basic effect of posing upon feelings of power, if not downstream outcomes. The loser from this portrait of instability, perhaps, is the public image of social psychology as a whole.

Because of the perennial nature of exhortations to communicate, it is particularly important to look at some widely-spread views of what holds back communication. In 2014, the American Psychological Association (APA) president commissioned a Presidential Task Force to address the lack of public outreach and understanding about scientific psychology. As reported in the APA *Monitor* (Kaslow, 2014), initial feedback from this task force answered the question "What stops us?" (from communicating) with this explanation:

Countless barriers make it challenging for us to communicate our science. They include attitudinal barriers, such as the disconnect between scientists and practitioners and the view of some psychologists that translation and advocacy are not part of the role of the scientist or scientist-practitioner. There are also barriers related to the public's ability to understand science and mental health issues, as well as with psychologists' ability to communicate in ways that can be understood by non-psychologists. Many psychologists lack experience and comfort with communicating to the public or lack the academic institutional support to do so. (p. 5)

A similar, if more specific, analysis was offered at the 2016 conference of the Society for Personality and Social Psychology, in a symposium talk titled "Talking to humans: Is it a good idea?" (Gilbert, 2016). To summarize this talk is to lose all of the exuberant wit with which Gilbert expressed his views, but the essence of his argument was that scientists don't communicate with lay people because they are afraid of two things. First, they fear that other scientists will scorn them for watering down their message. Second, they fear that in communicating simplified or preliminary findings, they will commit a crime against the truth. Gilbert urges us nevertheless to go ahead and communicate what we know, because our work is important, the world is necessarily imperfect, and the process of communicating can help psychologists figure out for ourselves what is most important. For now, it is enough to note that both these US-based analyses decry an apparent communication gap in scientific psychology, and focus on the individual researcher's attitude and ability as culprits.

It is also interesting to note that recent Presidential addresses of the British Psychological Society (BPS) take a much more positive approach to engagement, reviewing communication success stories rather than exhorting people to engage (Hacker Hughes, 2016; Kinderman, 2017). This may be due to the BPS' greater structural engagement with the media and communication, relative to comparable US organizations. Attention to incentives seems to produce, and flow

from, an approach to improving communication that is systemic rather than focused on

individual fears, individual attitudes, or individual lack of ability.[1]

Incentives and relevance

We have seen that changing practices in research evidence standards are largely driven by

career incentives, and in particular the increased willingness of journals in social psychology to

give credit for replications, fully reported research, nonsignificant findings, and other

innovations that overall tend to improve the quality of our evidence base. When it comes to

relevance, then, a larger gamut of incentives needs to be looked at. In this next section I review

the existing incentives that work both for and against greater adoption of the three standards of

relevance, and offer suggestions for those who wish to promote relevance effectively.

Incentives for question relevance

There is no question that a great deal of highly useful applied psychology work has been

carried out in recent decades, both in psychology departments and in allied disciplines such as

communication and marketing. Still, as shown previously, there is a perception that applied work

is not respected as highly as basic research, when careers in social psychology are evaluated in

hiring and institutional assessment.

The objective basis for this disparity is hard to find. Metrics such as the Impact Factor of

journals give only equivocal evidence for it. In the Thompson Reuters Journal Citation Reports

(2017), the category "Psychology, Applied" has an aggregate Impact Factor of 2.27, compared to

---

1   I should mention that in a longer article based on her presidential address, Kaslow (2015) does mention, if briefly, the need to actively reward communication and remove disincentives.

2.08 for "Psychology, Social." In the Social category, moreover, many journals appear that could reasonably be classed as applied. While not as impactful as the flagship *JPSP* (5.73), the ratings of *Journal of Health and Social Behaviour* (2.62) and *Law and Human Behaviour* (2.87) are respectable enough compared to their basic social counterparts, while in the Applied category proper, *Journal of Applied Psychology* (4.64), *Journal of Management* (8.08) and *Personnel Psychology* (5.52) more than hold their own.

It may be that the reputation of a research publication, and by extension its field, depends not just on how many citations it gets, but who is citing it -- the "direction of gaze" of a field as measured by more specific citation patterns. High-ranking applied journals might not be seen as prototypical social psychology, because their authors and readers overlap with fields such as management and organizational psychology. To compare *JPSP* to *Journal of Applied Psychology* (*JAP*), while both score high in impact, from 2008-2017 *JAP* has cited *JPSP* 337 times, while *JPSP* has only cited *JAP* 41 times. *Personality and Social Psychology Bulletin* cites *JAP* 38 times during this period, and is cited by *JAP* 75 times, although its impact factor (2.50) is much lower than that of *JAP* (4.64). Where, then, does *JAP*'s impact rating come from? That journal's top ten sources of incoming citations are all journals in management, organizational psychology and business. Much of its knowledge is exported outside psychology altogether.

If basic psychology does not look well on applied psychology, then, it may be because basic psychology does not look upon it at all – the one-way model put into practice in the flow of citations. More generally, the true impact of applied psychology on the field as a whole can never be adequately measured by its merely scholarly impact, because its findings are taken up by users who do not leave a scholarly record. Instead, they communicate by newsletter articles, training manuals, evaluation documents, and even more ephemeral means. This may all seem

like a natural consequence of the one-way flow of knowledge, from theory to basic research to applied research to practice. But there is nothing necessary about valuing upstream over downstream locations in this ecology. Lewin, as already mentioned, wanted theory to learn and take validation from application just as much as the other way around.

Imagine a parallel universe where Titchener and Boring were hit by a horseless carriage circa 1915, so that the APA missed the sharp turn to experimentalism. Career expectations might then require that after validating a basic effect in the lab, some real-world application be tested. But in our universe, the kind of follow-up research urged on scholars is usually to elaborate and refine the evidence for the theory behind the effect. The privilege of theory is upheld by publication standards that require evidence of exact causal mechanisms, "process" variables, or moderating effects as necessary elements of research development. These are all elements more readily tested in the lab than in field settings, where conditions are messy and participants' patience to fill out additional measures is limited.

While it is true that field studies are not categorically barred from publication in "the major social psychological journals," evidence standards at those journals mean they often have to be packaged with a cohort of lab studies (e.g., Webber et al., 2018; Goff, Jackson, Di Leone, Culotta, & DiTomasso, 2014), whereas there is no reciprocal requirement for mostly lab-based studies to take to the field. While the call of McGuire (2009) to test basic theoretical questions in field settings is appealing, the added difficulty and expense of field research already weigh against this move, not to mention the risk that a field setting would be seen as lacking some crucial element of sophistication or control.

Standards of novelty and autonomy in publishing and career evaluation also undercut relevance. We award kudos to someone who has established a fresh theoretical idea, and the

acclaim for such innovation does not require proof of concept in the field. Lower marks go to someone who merely tries to reproduce other people's findings (Neuliep & Crandall, 1990, 1993). Under the one-way flow of ideas from basic to applied realms, applied research is cast in the role of conceptually replicating basic research. Seldom is basic research required to validate itself by building on systematic observations from the field.

Moreover, a practitioner seeking to adopt an effective intervention usually wants one that draws on various theories to have the best chance of success, rather than promoting a single view. The work on volunteering described by Snyder (1993), for example, drew on incentive theory, as well as a variety of self- and social-oriented motivational theories. These features work against the tacit requirement that research should be novel, causally pure, and represent an autonomous contribution. So, even if we could get basic research out into the field, practitioners may baulk when career incentives push the researchers into insisting that only their own, recent work be included as the basis for an intervention.

Add to all these disincentives the greater difficulty and expense of doing relevant research in the field, and it becomes obvious why there is a constant unresolved worry about social psychology's relevance. But on a more positive note, applied research questions and methods already have many incentives in their corner. Organizations such as the Society for the Psychological Study of Social Issues (SPSSI) at the research end, and the Behavioural Science Policy Association (BSPA) at the translation end, are dedicated to publishing, promoting, and rewarding relevant research topics and outreach to policymakers. Importantly, the recently founded SPARQ center at Stanford University (https://sparq.stanford.edu ; Conner, March 2014) not only directly enlists leading social psychologists to communicate directly with practitioners on issues of public relevance, but emphasizes the reciprocal experience of social psychologists

learning from practitioners what would be useful research to do, in a true application of the Lewinian model. Public funding is constantly under pressure to show results, and the recurring sense of a relevance crisis in the USA coincides with times when the federal government directed resources away from basic research, in the 1970's (Brim et al., 1969) and 40 years later. The effect of the scientific nihilism shown by the Trump administration to date is yet to play out fully, but it may well make an even stronger case for increasing the relevance of our research in the eyes of the public, rather than retreating into basic questions.

In the United Kingdom, the funding councils have made a major effort in the past ten years to increase relevance under the rubric of "impact." Successful grants now need to make a statement about the relevance of the proposed research, whether it be application outreach, or the communication of findings from basic research. Career evaluation in the UK also increasingly depends on one's ability to make the institution look good in the periodic national research assessment that determines distribution of funding to universities (currently, the Research Excellence Framework, or REF). As of 2013, the REF has included "impact cases" in its evaluation, with clear if quite stringent guidelines that reward research that was developed in academia and later proven to have been successfully applied outside that setting to change policy. Finally, over the past two decades or so, the massive FP6, FP7, and HORIZON 2020 grant schemes funded by the European Union have all emphasized the outcomes and uses of research, usually requiring active collaboration with business or other end-users.

Those who champion greater application of social psychology should, above all, work to remove barriers in publication and career evaluation. Few journal policies (including, admittedly, our own current policy at *JESP*) make clear how high-relevance tests of research can fit into criteria that explicitly stress internal validity and theoretical rigour. In a time when mediation,

moderation, or complex interactions are seen as tacit requirements, the more elegant designs typical of applied research require positive validation if they are to be encouraged. Those who sit on hiring, tenure, and promotion committees should consider how their standards might discourage a research approach that finds solutions driven by a problem, rather than creating the solution first and then looking for a problem.

Those who are satisfied with a one-way flow of information from basic to applied research will have no need to call for increased question relevance. But, for those who are not, it is time to go beyond merely saying "we should support applied research more." Effective change must affirm, overtly, that theoretical sophistication and real-world relevance are parallel grounds for evaluation. To do more than merely pay lip-service to Lewin's vision of a reciprocal flow of information, evaluators should reward those theories that seek out and incorporate insights from the field, and look for and reward explicit consideration of the reverse path from applied to basic research. If basic and applied research are indeed equal, then theory without relevance should be given as much critical scrutiny as relevance without theory.

Incentives for sample relevance

Most social psychologists would likely conduct their research using more representative samples, if they had unlimited budgets and help with recruitment. The incentive to use samples of convenience is above all economic. Undergraduates can work for free, as part of course requirements, and are readily available. Online participants also started out being paid rates well below the minimum wage in most Western countries, to the point where the typical left-liberal social psychologist (Duarte, Crawford, et al., 2015) risked being characterised in a modern-day Devil's Dictionary as someone who demands that McDonalds pay $15 an hour yet accepts an

AMT rate of payment of $3 an hour. Even if we heeded the call to pay a living wage to online workers (Williamson, 2016), the ease and speed of recruitment and the availability of a willing population make online worker pools easier to use than community samples. With less and less funding available, and career evaluation demanding more and more studies in more and more publications, requiring more and more participants (thanks to the evidence crisis), something has to give. That something is usually sample relevance.

Achieving sample relevance also can require additional effort: translating and back-translating for cross-national studies, making sure items are valid for all cultural backgrounds and educational levels studied, bringing the lab to the people or the people to the lab. Adding insult to injury, researchers often find that non-standard populations become a marked category in the review process (Hegarty & Bruckmüller, 2013), raising issues that can jeopardise acceptance at highly selective outlets. For example, if you recruited a sample of noodle vendors in Thailand in order to study certain basic processes, you would risk a reviewer asking "Why did you study Thai noodle vendors?" or "Can these results really generalize beyond the noodle stand?"-- questions almost never asked of "normal" but actually WEIRD participant populations such as North American introductory psychology students.

A few bright lights of incentive have lit the way toward more representative and diverse samples. Some journal editors have raised the issue in recent editorials; in particular, the recently appointed chief editor of *JPSP*'s Attitudes and Social Cognition section has spelled out concrete ways that replicating an effect across samples can now open the door to publication: "If the same effects occur in a sample that is very different from the original one, this will constitute a 'big plus' that could bring the paper above the threshold for publication in JPSP: ASC. Likewise,

demonstrating that an effect does not generalize across samples for theory-guided reasons would also be considered a strength." (Kitayama, 2017, p. 359).

Funding opportunities have also encouraged the use of more relevant samples. The best of these have explicitly funded research on large representative population cohorts, such as TESS (Time-sharing Experiments in the Social Sciences) a program funded by the US National Science Foundation and partnered with a large opinion research corporation. Since 2001, TESS has been taking competitive applications for one-shot experimental studies to be conducted on stratified panels of about 2,000 US residents, accepting about 24 a year, and taking on all responsibility and costs of data collection once the approved questionnaire has been submitted.

In an increasingly competitive multidisciplinary funding environment, it seems rash for a social psychologist to endanger a grant proposal by reducing costs through the use of convenience samples. Compared to the costs of salaries and overheads, or the large sums spent on scanning time in neuroimaging research, allocating grant budget to recruiting representative samples should be seen as a worthwhile investment. Initiatives would be especially welcome for redirecting funds going to Amazon into non-profit research enterprises similar to TESS, ones that recruit and maintain quality participants keeping in mind a researcher's needs.

Incentives for communication relevance

As with the other two relevance areas, career pressures within academia seem to discourage spending time on communication output to the public. Scholars are instead advised to focus their attention on journal articles, grants, and service activities such as journal reviewing or editing. A number of studies and commentaries have documented how mentors and peers in academia actively discourage public engagement (for a review, see Kaslow, 2015). How, then, is

there so much communication of social psychology in the media, if social psychologists are not supposed to be doing it?

For one, there is a certain intrinsic publicity value to media coverage. University administrations, in particular, see value in having media-star faculty and a flow of research news branded with the university name. To varying degrees, they encourage press releases, media contacts, and other forms of engagement; under the supervision of public information officers, these activities have increased at universities over the past decades (Dudo, 2015). Moreover, a large number of empirical studies across a variety of fields have found that scientists communicate with the media and public more often than might be assumed, that their interactions are largely enjoyable, and that much more often than not they leave with the impression that their findings were communicated accurately (Dudo, 2015).

To ask the opposite question, then: if public communication can be pleasurable, why aren't more people doing it? One reason is that research communicated to the public has to meet some standard of question relevance. While a certain level of theoretical and methodological sophistication has become state-of-the art, even necessary, in order to publish in leading journals, such articles are not likely to engage the public. Structural equation models and complex interaction effects have great merit when addressing precise scientific questions. But papers focused on such findings, or on tests between abstract theories, are not really cut out for the public eye, and their authors are aware of this.

Nonetheless, compared to cell biology or particle physics, social psychology deals much more with everyday concepts. Our research often proceeds from concrete ideas and insights with an immediate appeal – for example, why do people disapprove of "victimless" crimes? And do people in power think differently than their subordinates do? Nor does research have to be

immediately applicable to social problems in order to be interesting. The public is often eager to validate, challenge or expand their own ideas about basic concepts like emotions, memory, culture, and personality. Miller's (1969) "giving psychology away" address proposed that influencing lay psychological beliefs could itself be a revolutionary act. Schwartz (1997) has argued that introducing ideas from psychological research into the wider culture itself is a kind of intervention, or "idea technology," shaping expectations about what is possible in a self-fulfilling way. Moscovici's (2008) analysis of the influence emanating from the social representation of psychoanalytic concepts is one testimony to the power of such ideas.

The problem may not lie so much in encouraging communication, as in understanding when it is and is not high-quality. A practice may become devalued because scholars don't know how to evaluate it, as has been observed with replication studies (Brandt, IJzerman et al., 2014). If standards of evaluating media engagement are unclear, then media engagement risks being excluded entirely from consideration (e.g., when applying for jobs or promotion) out of a suspicion that credit for media engagement has accrued only for trivial, easy, inaccurate, or low-quality efforts. And bad media coverage does exist in science. A recent study of health research communication identified most exaggerations in the media as traceable to overstated press releases (Sumner et al., 2014), while other research shows that press reports often fail to take research quality into account (Wang, Bolland, Gamble & Grey, 2015). Worse still, some academics and universities collude with industry and the press in preparing highly questionable "research" releases on such topics as the equation for a perfect cup of tea, especially in the UK (Goldacre, 2009). Bad apples of research communication tend to be generalized into a whole rotten orchard that junior researchers are warned away from entering.

Kaslow (2015) called for a "culture shift" in psychology, which would involve increased training and research into effective outreach. Most concretely, she suggests introducing more honours and awards for individuals and institutions who communicate well. Indeed, awards for media engagement are already given by (among others) the Society of Personality and Social Psychology, the Australian Psychological Society, and the BPS. And, returning to an earlier topic, applied research in the public interest is honoured with annual awards by the APA itself, by the Association for Psychological Science, and by the Society for the Psychological Study of Social Issues.

While these competitive incentives have the desirable feature of being vetted for quality, arguably they do not go far enough to reward good-quality public communication as a normal, rather than exceptional, achievement in an academic career. Unlike the other facets of relevance, public communication does not have an easy means of evaluation, of the kind that peer review and journal prestige provide for published articles. Imposing a change by fiat on hundreds of job and advancement committees, and the advisors who prepare junior scholars to meet them, would be a difficult task at best. Institutions dedicated to systematic improvement of science communication can help: for example, the workshops and media events organized through The Alan Alda Center for Communicating Science, at Stony Brook University in New York. Another way forward would be to treat good-quality public communication as a normal and necessary requirement of public funding. The Economic and Social Research Council (ESRC) in Britain have done so through their impact requirements for grant proposals, which can be achieved via policy influence, communication with professionals, or communication with the public.

I can imagine a more ambitious project, still within the reach of a professional organization, that would help to validate good news engagement in the eyes of sceptical scholars.

Let's imagine a website that would be sponsored by one of our professional societies, called SPInsight, for "Social Psychology Insight." Psychologists would be invited to submit links to general-interest news and magazine articles about their research to SPInsight, supplied with clerical support and a board of peer editors. The editors, with the help of specialist reviewers where needed, would rate each article on accuracy and evidence base, and select a subset of the best for linking on the site. Selection onto SPinsight would be available as a benchmark of service to the field for career evaluators to consider, and could plausibly substitute for other activities such as reviewing or society office-holding. What's more, SPInsight could address concerns about communication by shifting its standards or creating special prestige sections--for example, giving special honours to articles that address broader issues than the findings of one lab, such as evidence from a meta-analysis, or findings that explain dissent within a field to the public in a measured and informative way.

Links between relevance and evidence

Relevance and evidence concerns do not just exist in parallel. We may very well need to address questions of relevance in order to move the evidence crisis past a number of impasses that have appeared over the past five years or so. That is, to solve some problems in the evidence crisis, we will have to make progress on relevance issues as well. Most broadly, seeing what we do as relevant to real-world outcomes can motivate bolder steps toward reproducible and solid research. Triviality is sometimes taken as a defence against accusations of poor practice, as in this interview with one psychologist whose media-friendly research on dancing had apparently overstated exploratory results: "As I said, we're psychologists," he said. "We haven't wrongly

reported a cure for cancer or anything like that. You know, it's not that important, is it really?" (Engber, 2017).

If we are to tell a better story about our research--especially to the ears of granting bodies-- this kind of excuse should be inadmissible. Taking on more important questions and using less controllable methods should naturally motivate more serious, less playful approaches to solutions. To take just one example, when working with a difficult-to-recruit but highly relevant sample, running a dozen underpowered "pilot" studies with slightly different methods and reporting only the ones that worked is not even a plausible option, as it is with students or AMT workers. Although it makes acquiring large numbers more difficult, the real-world sample is compatible with other characteristics of more valid and reproducible research (e.g., LeBel & Peters, 2011), such as taking greater care with the validity and reliability of measures, and pre-registering hypotheses to guide principled data analysis. Improved practices such as results-blind review might also encourage the complete reporting of results from more difficult data sets, even if hypotheses are not confirmed, and discourage the practice of scouring the data for unexpected findings which are then presented as if anticipated all along (Kerr, 1998).

The urgency of reproducibility issues in medical research derives not just from that field's life-or-death stakes, but also because it is a discipline that values research relevance. Clinical trials failing to confirm the findings implied by some areas of lab-based study have given a strong push forward to research reform in medicine (Begley & Ellis, 2012; Grignolo & Pretorius, 2016; Turner, Matthews et al., 2008). Most areas of basic social psychology do not have a corresponding cycle of "clinical trials," or confirmation in applied settings. Nevertheless, it is tempting to speculate that if we had paid more attention to the Lewinian process of uptake from practical application, we might have had fair warning about some of the well-publicized basic

effects in social psychology that have proven hard to replicate in other labs (e.g., behavioural

priming, Cesario, 2014), and have also scarcely been tested in the field.

Even more fundamentally, questions of statistical evidence in basic psychology have

reached an impasse that would be best broken with a dose of relevance. Some argue that

statistical reporting should be based on effect sizes with confidence intervals around them, rather

than directional effects significant according to *p*-value (Cumming et al., 2014). This has

implications for sample size. To establish a stable effect size, much larger samples are required

than when establishing a mere directional effect (Schönbrodt & Perugini, 2013). Meta-analytic

integrations of research also find it harder to come to consensus about a typical effect size,

compared to whether or not the effect plausibly goes in a particular direction (Fabrigar &

Wegener, 2016; Francis, 2017). Nonetheless, deciding on a target effect size before results are

known underlies power analysis and Bayesian analysis, necessary tools to satisfy some recent

journal requirements aimed at improving the evidence base of studies (e.g., Trafimow & Marks,

2015; Vazire, 2016).

But dealing with effect sizes is complicated by the fact that, in much basic research,

measures and contexts are abstract, making standardized effect sizes difficult to benchmark. How

much happiness is gained, exactly, when increasing 0.3 scale points on a 1-7 scale?

Psychophysics appears to be the only sub-discipline that has really developed insight into this

notion of "criterial magnitudes," or how much difference makes a difference (e.g., Ward, 1992).

Meta-analyses of typical social psychological effects certainly vary greatly in size according to

topic (Richard, Bond & Stokes-Zoota, 2003). Comparisons between similar effects in the lab and

the field show no consistent strength advantage to either, and there is considerable variability in

which setting yields stronger findings, especially in social psychology (Anderson, Lindsay & Bushman, 1999; Mitchell, 2012).

Finding an effect of a given size in the lab doesn't mean that the same size will be found in the real world. When studies are conducted in the real world, the outcomes are more likely to be concrete behaviours, decisions, and personal outcomes. Effect size is less difficult to grasp when the outcome variables are school grades, money donated to charity, colds not caught, or number of jobs applied for. Any of these outcomes, as an unstandardized effect size, is meaningful because it provides the basis for a cost-benefit analysis. If the cost is trivial -- for example, changing the wording in the next edition of a government leaflet -- then a positive benefit of any size justifies its adoption. In short, the "new statistics" make more sense when dealing with measures that live up to criteria of relevance.

A final word on relevance, values, and technocracy

As mentioned earlier, Hillary Clinton brought social psychological research to the forefront of relevance in the 2016 US Presidential debates. This was just one of many signals that mainstream governmental parties worldwide, both liberal and conservative, have been willing to listen to and make use of our science. The Behavioral Insights Team, established by David Cameron's government in the UK following earlier initiatives by the Blair government, was a so-called "nudge unit" using behavioural science (that in the view of Kahneman, as noted earlier in this essay, is social psychology) to encourage prosocial choices. The Obama administration was beginning to follow suit in the US by establishing the Social and Behavioral Sciences Team, now apparently discontinued by the Trump administration (Stillman, 2017; Social and Behavioral Sciences Team, 2017). Social psychology techniques for persuading people and attitude-behaviour techniques for getting out the vote were used to an unprecedented degree in both

Obama Presidential campaigns (Carey, 2012; Issenberg, 2012), with an advisory panel including Susan Fiske, Robert Cialdini, and Richard Thaler. The Hillary Clinton campaign continued this engagement, with voter encouragement techniques derived straight from social psychology's own implementation intention theory (Abcarian, 2016).

We all know how key political campaigns in 2016 ended. Behind the victorious Trump and Brexit campaigns lay a different type of research-driven campaign, one based on social media rather than personal or phone contact, and targeting voters with different messages based on personality rather than social psychology (most notably, Cambridge Analytica; Commons Select Committee on Culture, Media and Sport, 2017; Grassegger & Krogerus, 2017). The irony is that a resurgent populist right, predicated on a revolt against big government, bureaucrats and experts, made its own use of sophisticated behavioural science--and at a fraction of the cost of traditional campaigning. The undisputed message of 2016 was the emergence of new rifts pitting Establishment politicians against a populist style, on both the right and the left of politics. Social psychologists might reasonably feel existential threat. The relevance of our field seems decidedly endangered these days, especially given the air of technocracy attributed to some of our flagship behavioural interventions..

These anxieties have been felt before. Indeed, the crisis period in the 1970's also was a time of re-examining the ethics of social psychology. With the rise of radical and critical voices came a new look at the paradigms of deception and manipulation that characterized such classic studies as Milgram's obedience paradigm and the Stanford prison "experiment." Ethical practices and guidelines were recognized as often lacking or inadequate, calling social psychology into question not only as a knowledge enterprise but as a moral enterprise (Elms, 1975). In the 1980's, informed consent and debriefing, along with institutional review, were

established and generally acknowledged as acceptable procedural and regulatory answers to these research issues--with some dissent, largely emanating from behavioural economists who objected even to ethically shielded deception research (e.g., Davis & Holt, 1992; Hertwig & Ortmann, 2001). Eventually, deception paradigms waned in numbers, without completely disappearing (Nicks, Korn & Maineri, 2010). This decline may simply be a consequence of the more general retreat from highly relevant and engaging behavioural designs overall, as already noted (Baumeister et al., 2007).

However, deeper philosophical questions raised by the crisis of ethics were not so easily resolved. Although the experiment as a basic-knowledge process is a way to establish causality, it also lays the groundwork for the application of manipulations or interventions – or, in a more sinister light, the exercise of power by experts upon a "subject" group (now, more respectfully, called "participants"). No matter how well-intentioned they are, or how similar to everyday influence (e.g., Smith, Goldstein & Johnson, 2013), interventions present a problem for values of freedom and autonomy. Argyris (1975), writing in the full swing of the first "crisis," offered a trenchant analysis of some of the most issue-relevant social psychology research of the time, showing how it assumed a world of rational individuals controlling each other ("Model I") rather than one of free and informed choice ("Model 2").

Moral unease is especially great when forces outside our awareness or beyond our control are leveraged to create behavioural change, be they the second-hand Freudian "depth motivation" glorified by the 1950's advertising industry (Packard, 1957) or the more up-to-date heuristics and primes of the cognitive unconscious (e.g., Nairn & Fine, 2008). Indeed, social psychology's turn to the unconscious, first seen as a way of overcoming the limitations of self-report (Greenwald & Banaji, 1995), soon developed into assertions that most human behaviour

was determined by automatic forces (Bargh & Chartrand, 1999). There followed a wave of experimental research in social psychology using surprisingly small-scale environmental changes to produce relatively large-scale behavioural changes without participants being aware of the influence--for example, increasing creativity by sitting outside versus inside a large cardboard box, i.e. "thinking outside the box" (Leung, Kim et al., 2012). The popularity of these manipulations creates a loophole in the purportedly resolved ethics issues of the 1970's. They do not outright deceive their participants, nor do they fully disclose the influences on them. While there are questions about the replicability of this kind of finding in the wake of the second crisis of evidence (Cesario, 2014), it is perhaps reassuring, for advocates of human autonomy, that the literature of unconscious influence has seen few applications in field settings.

One way to resolve these meta-theoretical doubts is to return to a deeper reading of Miller's (1969) address, mentioned earlier. In his view, "giving away" psychology is also an act of giving up control, because "In the public view, I suspect, all this talk about controlling behavior comes across as unpleasant, if not actually threatening" (p. 1068). Instead, Miller advocates letting the public--and practitioners--use our insights as they will, effectively becoming the masters of their own intervention. It is through this avenue of communication that basic research could come into its own, as long as the concepts in play, such as emotions, morality, how we learn and remember, and social issues of current concern, are inherently understandable.

Most importantly, to convince us to reach out to the outside world, suitable incentives need to be there. Real-world relevance should be recognized as a publication-worthy contribution alongside theoretical precision and sophistication. Perhaps a fourth section of *JPSP* could be

established, to enshrine highly relevant research at the top of the publishing league.[2] Research using populations outside the convenient norm should also be given special consideration in publication, along with some degree of forgiveness for the restrictions that relevance places on the niceties of design. Communication based on a measured view of the literature, addressing public curiosity without betraying scientific integrity, should be evaluated and recognized as a positive service to the field--not just as a special decoration, but as an everyday activity. It is not realistic to expect psychologists to improve their behaviour simply out of a sense of duty. A sense of duty is already expressed *pro bono* in the editorial process and through other forms of service to the profession. But with increasing and continuing pressures on all levels of a career, it is difficult to ask people to take on much more for the common good. Even when researchers are safely ensconced in tenure or full professorship, they still are guided by scholarly habits that have not been unlearned from the days of everyday struggle, while now having the career interests of their own students and protégés to look after.

To conclude, I suspect that the metaphor of a "crisis" is overly flattering to the importance that the field places on relevance. After all, a crisis is a self-evident thing. The word implies that everyone knows something bad is going on and realizes something must be done about it. Relevance, however, has been less of a crisis in this sense. It has been more of a persistent low-level worry that stings every time a social psychologist is misidentified as an economist or a sociologist,[3] or every time we wish that people would listen to what we have to say about some social issue. But there have been no signals surrounding our relevance "crisis" that would by

---

[2] For this idea I have the remarks of the present journal's editors, Miles Hewstone and Tony Manstead, to thank.

3   Except that one time, once, when a noted US radio personality identified a famous fraud in social psychology as a "sociologist" (Limbaugh, Nov. 4, 2011). That felt better.

themselves spur on action, as rising waters or a marching army of the unemployed would do.

Investing in the relevance of our field, then, is more of a long-term concern that will ensure our

respect, increase the chances that governments and businesses will turn to us, and make us feel

better about ourselves when friends and relatives tell us that they have heard about our research.

Those who disagree with this analysis are free to do so, but then should commit explicitly

to a one-way model of influence from basic to applied psychology. Doing this should come with

the understanding that the one-way model entails a certain lower prestige for applied work,

because applied work must look to basic work while basic work is not obliged to look back at the

applied. But if this prospect is uncomfortable to the reader, if Lewin's action research and

mutuality of regard are things that strike you as worthwhile, I hope my analysis convinces you

that mere concern is not enough. There is instead a need for concrete incentives designed to

enhance the prestige of applied research, which would go hand-in-hand with greater visibility of

applied research, as it can inform the validation and modification of basic-level ideas in social

psychology. To achieve that goal, in turn, requires far-sighted leaders who are prepared to make

relevance worthwhile for the average social psychologist.

References

Abcarian, R. (June 6, 2016). Using persuasion and a dollop of science, Clinton volunteers dial

    for votes. *The Los Angeles Times*. Retrieved from

    http://www.latimes.com/local/abcarian/la-me-abcarian-clinton-volunteers-20160606-

    snap-story.html

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review

    of empirical research. *Psychological Bulletin*, *84*, 888-918.

Anderson, C. A., Lindsay, J. J., & Bushman, B. J. (1999). Research in the psychological

    laboratory: Truth or triviality? *Current Directions in Psychological Science*, *8*, 3-9.

Anderson, C. A., Allen, J. J., Plante, C., Quigley-McBride, A., Lovett, A., & Rokkum, J. N.

    (2018). The MTurkification of social and personality psychology. *Personality and Social*

    Psychology Bulletin, online publication. Retrieved from http://journals.sagepub.com

    /doi/abs/10.1177/0146167218798821.

Argyris, C. (1975). Dangers in applying results from experimental social psychology. *American*

    Psychologist, 30, 469-485.

Asendorpf, J. B., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J. J., Fiedler, K., ... &

    Perugini, M. (2013). Recommendations for increasing replicability in psychology.

    European Journal of Personality, *27*, 108-119.

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American*

    Psychologist, *54*, 462-479.

Bauman, C. W., McGraw, A. P., Bartels, D. M., & Warren, C. (2014). Revisiting external

    validity: Concerns about trolley problems and other sacrificial dilemmas in moral

    psychology. *Social and Personality Psychology Compass*, *8*, 536-554.

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior?. *Perspectives on Psychological Science*, *2*, 396-403.

Begley, C. G., & Ellis, L. M. (2012). Drug development: Raise standards for preclinical cancer research. *Nature*, *483*, 531–533. https://doi.org/10.1038/483531a

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.

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis*, *20*, 351-368.

Billig, M. (2015). The myth of Kurt Lewin and the rhetoric of collective memory in social psychology textbooks. *Theory & Psychology*, *25*, 703-718.

Brandt, M. J., Crawford, M., & Giner-Sorolla, R. (2016). Editorial: Special issue on confirmatory research. *Journal of Experimental Social Psychology*.

Breckler, S. J. (2006a). Psychology is translational science. *Monitor on Psychology, 37* , 22.

Breckler, S. J. (2006b). Embracing the many applications of psychology. *Monitor on Psychology, 37* , 24.

Brim, O. G., Jr., et al. (1969).. Knowledge into action: Improving the nation's use of the social sciences. ('Report of the Special Commission on the Social Sciences of the National Science Board) Washington, D.C.: National Science Foundation, 1969.

Britt, S. H. (1937). Past and present trends in the methods and subject matter of social psychology. *Social Forces*, *15*, 462-469.

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

Buschini, F., Guimond, S., & Breakwell, G. M. (2011). Social issues and social psychology: Distinctive pathways in applying social psychology to resolve major social problems. *Revue Internationale de Psychologie Sociale*, *23*, 5–15.

Calvert, J. (2006). What's special about basic research?. *Science, Technology, & Human Values*, *31*(2), 199-220.

Cantril, H. (1934). The social psychology of everyday life. *Psychological Bulletin*, *31*, 297-330.

Carlson, R. (1984). What's social about social psychology? Where's the person in personality research?. *Journal of Personality and Social Psychology*, *47*, 1304-1309.

Carey, B. (Nov. 12, 2012). Academic "Dream Team" helped Obama's effort. *The New York Times*. Retrieved from http://www.nytimes.com/2012/11/13/health/dream-team-of-behavioral-scientists-advised-obama-campaign.html?_r=0

Carney, D. R., Cuddy, A. J., & Yap, A. J. (2015). Review and summary of research on the embodied effects of expansive (vs. contractive) nonverbal displays. *Psychological Science*, *26*, 657-663.

Carney, D. R. (2016). My position on "power poses". Retrieved from http://faculty.haas.berkeley.edu/dana_carney/pdf_My%20position%20on%20power%20poses.pdf

Center for Open Science (2018). Implementing TOP. Retrieved from https://cos.io/our-services/top-guidelines/.

Cesario, J. (2014). Priming, replication, and the hardest science. *Perspectives on Psychological Science*, *9*, 40-48.

Cesario, J., Jonas, K. J., & Carney, D. R. (2017). CRSP special issue on power poses: what was the point and what did we learn?. *Comprehensive Results in Social Psychology*, *2*, 1-15.

Chandler, J., Mueller, P., & Paolacci, G. (2014). Nonnaïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers. *Behavior Research Methods*, *46*, 112–130. https://doi.org/10.3758/s13428-013-0365-7

Cialdini, R. B. (2009). We have to break up. *Perspectives on Psychological Science*, *4*, 5-6.

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

Commons Select Committee on Culture, Media and Sport (2017). *Disinformation and 'fake news': Interim Report.* Retrieved from

https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/363/36302.htm

Conner, A. C. (March, 2014). Stanford SPARQ sparks change. *APS Observer*. Retrieved from https://www.psychologicalscience.org/observer/stanford-sparq-ignites-change.

Cooper, M. L. (2016). Editorial. *Journal of Personality and Social Psychology*, 110, 431-434. http://dx.doi.org/10.1037/pspp0000033

Craig, R. T. (1996). Practical theory: A reply to Sandelands. *Journal for the Theory of Social Behaviour*, *26*, 65-79.

Crandall, C., Leach,.C. W., Robinson, M. and West, T. (2017). The new editorial team of PSPB addresses editorial philosophy. Retrieved from

http://www.spsp.org/publications/personality-and-social-psychology-bulletin/Jan2017-Editorial-Statement

Cumming, G., Fidler, F., Leonard, M., Kalinowski, P., Christiansen, A., Kleinig, A., . . . Wilson, S. (2007). Statistical reform in psychology: is anything changing? *Psychological Science, 18*, 230-232.

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

Davis, D. D. & Holt, C. A. (1992). *Experimental economics.* Princeton, NJ: Princeton University Press.

De la Rey, C., & Ipser, J. (2004). The call for relevance: South African psychology ten years into democracy. *South African Journal of Psychology*, *34*, 544-552.

de Witt Huberts, J. C., Evers, C., & de Ridder, D. T. (2013). Double trouble: restrained eaters do not eat less and feel worse. *Psychology & Health*, *28*, 686-700.

Druckman, J. N., & Kam, C. D. (2011). Students as experimental participants. *Cambridge Handbook of Experimental Political Science*, ed. J. N. Druckman, D. P. Green, J. H. Kuklinski, & A . Luria (pp. 41-57). Cambridge, UK.

Duarte, J. L., Crawford, J. T., Stern, C., Haidt, J., Jussim, L., & Tetlock, P. E. (2015). Political diversity will improve social psychological science. *Behavioral and Brain Sciences*, *38*, e130.

Dudo, A. (2015). Scientists, the media, and the public communication of science. *Sociology Compass*, *9*, 761–775. https://doi.org/10.1111/soc4.12298

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

Ellemers, N. (2013). Connecting the dots: Mobilizing theory to reveal the big picture in social psychology (and why we should do this). *European Journal of Social Psychology*, *43*, 1–8. https://doi.org/10.1002/ejsp.1932

Elms, A. C. (1975). The crisis of confidence in social psychology. *American Psychologist, 30*, 967–976.

Engber, D. (2017, 14 February). Hips do lie. Retrieved from http://www.slate.com/articles/health_and_science/science/2017/02/why_did_the_press_c over_a_dubious_study_on_what_makes_women_great_dancers.html

Eriksson, K., Coultas, J. C., & de Barra, M. (2016). Cross-cultural differences in emotional selection on transmission of information. *Journal of Cognition and Culture*, *16*, 122-143.

European Commission (2002). The sixth framework programme at a glance. Retrieved from https://ec.europa.eu/research/fp6/pdf/fp6-in-brief_en.pdf .

Fabrigar, L. R., & Wegener, D. T. (2016). Conceptualizing and evaluating the replication of research results. *Journal of Experimental Social Psychology*, *66*, 68-80.

Faye, C. (2012). American social psychology: Examining the contours of the 1970s crisis. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *43*, 514-521.

Fishman, D. B., & Neigher, W. D. (1982). American psychology in the eighties: Who will buy. *American Psychologist, 37,* 533-546.

Francis, G. (2017). Comment on: Conceptualizing and evaluating the replication of research results. *Journal of Experimental Social Psychology*, *69*, 237-240.

Gergen, K. J. (1973). Social psychology as history. *Journal of Personality and Social Psychology*, *26*, 309–320.

Gernsbacher, M. A. (2006). Reaching for relevance. *APS Observer*, *19*, 5–33.

Gilbert, D. (2016, January 30). Talking to humans: Is it a good idea? Retrieved from http://www.danielgilbert.com/SPSP2016.pdf

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

Giner-Sorolla, R. (2016). Approaching a fair deal for significance and other concerns. *Journal of Experimental Social Psychology*, *65*, 1–6. https://doi.org/10.1016/j.jesp.2016.01.010

Gladwell, M. (2000). *The tipping point: How little things can make a big difference.* New York: Little Brown.

Godin, B. (2006). The linear model of innovation: The historical construction of an analytical framework. *Science, Technology, & Human Values*, *31*, 639-667.

Goff, P. A., Jackson, M. C., Di Leone, B. A. L., Culotta, C. M., & DiTomasso, N. A. (2014). The essence of innocence: Consequences of dehumanizing black children. *Journal of Personality and Social Psychology*, *106*, 526-545.

Grassegger, H., & Krogerus, M. (2017) The data that turned the world upside down. Retrieved from https://motherboard.vice.com/en_us/article/how-our-likes-helped-trump-win

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4–27.

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

Grignolo, A., & Pretorius, S. (2016). Phase III trial failures: Costly, but preventable. *Applied Clinical Trials*, *25*, 36–42.

Hacker Hughes, J. (2016, April). Twelve ways to make an impact. Presidential Address delivered at the Annual Conference of the British Psychological Society, Nottingham. Retrieved from

https://www1.bps.org.uk/system/files/Public%20files/HOPC/J%20Hacker%20Hughes%2 0Pres%20Address.pdf

Harford, T. (2014, March 21). Behavioural economics and public policy. *Financial Times Online.* Retrieved from https://www.ft.com/content/9d7d31a4-aea8-11e3-aaa6-00144feab7de.

Harré, R., & Secord, P. (1972). *The explanation of social behaviour*. Oxford: Blackwell.

Haslam, S. A., & McGarty, C. (2001). A 100 years of certitude? Social psychology, the experimental method and the management of scientific uncertainty. *British Journal of Social Psychology*, *40*, 1–21.

Helmreich, R. (1975). Applied social psychology: The unfulfilled promise. *Personality and Social Psychology Bulletin*, *1*, 548-560.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*, 61–135.

Hertwig, R., & Ortmann, A. (2001). Experimental practices in economics: A methodological challenge for psychologists?. *Behavioral and Brain Sciences*, *24*, 383-403.

Hoffman, R. R., & Deffenbacher, K. A. (1993). An analysis of the relations between basic and applied psychology. *Ecological Psychology*, *5*, 315-352.

Israel, J., & Tajfel, H., Eds. (1972). *The context of social psychology: A critical assessment.* Oxford: Academic Press.

Issenberg, S. (2012). *The victory lab.* New York; Crown Publishers.

Jonas, K. J., & Cesario, J. (2015). How can preregistration contribute to research in our field?. *Comprehensive Results in Social Psychology*, *1*, 1-7.

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*, 54-69.

Kaslow, N. J. (2014). Translating our science for the public. *Monitor on Psychology,* 45, 5.

Kaslow, N. J. (2015). Translating psychological science to the public. *American Psychologist, 70*, 361-371.

Keller-Allen & Li (July, 2013). *Psychological Science and Behavioral Economics in the Service of Public Policy.* Retrieved from https://www.nia.nih.gov/sites/default/files/d7/psychological_science_and_behavioral_economics.pdf

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

Kinderman, N. (2017). 'Psychology is action, not thinking about oneself'. Retrieved from https://thepsychologist.bps.org.uk/volume-30/june-2017/psychology-action-not-thinking-about-oneself

Kitayama, S. (2017). Editorial. *Journal of Personality and Social Psychology*, *112*, 357–360. https://doi.org/10.1037/pspa0000077

Labonte, R., Feather, J., & Hills, M. (1999). A story/dialogue method for health promotion knowledge development and evaluation. *Health Education Research*, *14*, 39-50.

Landers, R. N., & Behrend, T. S. (2015). An inconvenient truth: Arbitrary distinctions between organizational, Mechanical Turk, and other convenience samples. *Industrial and Organizational Psychology*, *8*, 142-164.

Leary, M. (January 21, 2015). Enhancing the relevance of personality and social psychology. Retrieved from http://www.spsp.org/news-center/announcements/enhancing-relevance-personality-and-social-psychology.

LeBel, E., Borsboom, D., Giner-Sorolla, R., Hasselman, F., Peters, K., Ratliff, K., & Smith, C. (2013). Psychdisclosure.org: Grassroot support for reforming reporting standards in psychology. *Perspectives on Psychological Science*, *8*, 319–325.

LeBel, E. P., & Peters, K. R. (2011). Fearing the future of empirical psychology: Bem's (2011) evidence of psi as a case study of deficiencies in modal research practice. *Review of General Psychology*, *15*, 371–379.

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

Leung, A. K. Y., Kim, S., Polman, E., Ong, L. S., Qiu, L., Goncalo, J. A., & Sanchez-Burks, J. (2012). Embodied metaphors and creative "acts". *Psychological Science*, *23*, 502-509.

Lévy-Leboyer, C. (1988). Success and failure in applying psychology. *American Psychologist*, *43*, 779.

Lilienfeld, S. O. (2012). Public skepticism of psychology: Why many people perceive the study of human behavior as unscientific. *The American Psychologist*, *67*, 111-129.

Limbaugh, R. (Nov. 4, 2011). Dutch sociologist falsified data to codify liberalism as science. Retrieved from https://www.rushlimbaugh.com/daily/2011/11/04/dutch_sociologist_falsified_data_to_codify_liberalism_as_science/ .

Lindsay, D. S. (2015). Replication in psychological science. *Psychological Science, 26*, 1827–1832.

Loftus, G. R. (1996). Psychology will be a much better science when we change the way we
     analyze data. *Current Directions in Psychological Science*, *5*, 161-171.

Long, W., & Foster, D. (2013). The changing face of "relevance" in South African psychology.
     Psychology in Society, *45*, 3-16.

Major, B. (2006). Toto, we're not in Kansas any more. *Dialogue, 21,*, 22-23.

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

Markus, H. R. (2004a). A social psychological model of behavior. *Dialogue*, *19*, (Suppl. 1)1–4.

Markus, H. R. (2004b). Thinking outside the person. *Dialogue*, *19*, , 19-23.

Maton, K. I. (2017). *Influencing social policy: Applied psychology serving the public interest.*
     Oxford, UK: Oxford University Press.

McGuire, W. J. (2009). Theory-oriented research in natural settings: The best of both worlds for
     social psychology. In M. Sherif & C. M. Sherif (Eds.), *Interdisciplinary relationships in
     the social sciences* (pp. 21-51). New York: Routledge.

Medin, D. L. (2012). A dangerous dichotomy: Basic and applied research. *APS Observer*, *25*.

Meyer, A. D. (1991). What is strategy's distinctive competence?. *Journal of Management*, *17*,
     821-833.

Miller, G. (1969). Psychology as a means of promoting human welfare. *American Psychologist,
     24*, 1063–1075.

Mitchell, G. (2012). Revisiting truth or triviality: The external validity of research in the
     psychological laboratory. *Perspectives on Psychological Science*, *7*, 109–117.
     https://doi.org/10.1177/1745691611432343

Moghaddam, F. M. (1987). Psychology in the three worlds. *American Psychologist*, *42*, 912-920.

Moscovici, S. (1972). Society and theory in social psychology. In J. Israel, & H. Tajfel (Eds.), *The context of social psychology: A critical assessment* (pp. 17–68). London: Academic Press.

Moscovici, S. (2008). *Psychoanalysis: Its image and its public*. London: Polity.

Mullinix, K. J., Leeper, T. J., Druckman, J. N., & Freese, J. (2015). The generalizability of survey experiments. *Journal of Experimental Political Science*, *2*, 109-138.

Nairn, A., & Fine, C. (2008). Who's messing with my mind? *International Journal of Advertising*, *27*(3), 447–470. https://doi.org/10.2501/S0265048708080062

Neuliep, J. W. & Crandall, R. (1990). Editorial bias against replication research. *Journal of Social Behavior and Personality*, *5*, 85-90.

Neuliep, J. W., & Crandall, R. (1993). Reviewer bias against replication research. *Journal of Social Behavior and Personality*, *8*, 21-29.

Nicks, S. D., Korn, J. H., & Mainieri, T. (1997). The rise and fall of deception in social psychology and personality research, 1921 to 1994. *Ethics & Behavior*, *7*, 69–77. https://doi.org/10.1207/s15327019eb0701_6

Nishi, A., Christakis, N. A., & Rand, D. G. (2017). Cooperation, decision time, and culture: Online experiments with American and Indian participants. *PLOS ONE*, *12*, e0171252. https://doi.org/10.1371/journal.pone.0171252

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., ... & Contestabile, M. (2015). Promoting an open research culture. *Science*, *348*, 1422-1425.

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

Novelrank (2017). Web site. Retrieved from https://www.novelrank.com/asin/006124189X

O'Donnell, J. M. (1979). The crisis of experimentalism in the 1920s. *American Psychologist*, *34*, 289-295.

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist, 17*, 776-783.

Packard, V. (1957). *The hidden persuaders.* New York: McKay.

Paolacci, G., Chandler, J., & Ipeirotis, P. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, *5*, 411-419.

Peer, E., Samat, S., Brandimarte, L., & Acquisti, A. (2016). Beyond the Turk: An empirical comparison of alternative platforms for crowdsourcing online behavioral research.

Peters, M. (2016, October 1). "Power poses" Co-author: "I do not believe the effects are real." Retrieved March 31, 2017, from http://www.npr.org/2016/10/01/496093672/power-poses-co-author-i-do-not-believe-the-effects-are-real

Peterson, R. A. (2001). On the use of college students in social science research: Insights from a second-order meta-analysis. *Journal of Consumer Research*, *28*, 450-461.

Pettigrew, T. F. (2018). Summing up: Did social psychology answer Dr. King's call?. *Journal of Social Issues*, *74*, 377-385.

Pielke, R. (2012). Basic research as a political symbol. *Minerva*, *50*, 339-361.

Pratkanis, A. R. (2017). The (partial but) real crisis in social psychology. In *Psychological Science Under Scrutiny: Recent Challenges and Proposed Solutions*, (S. O. Lilienfled and I. D. Waldman, Eds.), 141-163. Oxford: Wiley Blackwell.

Preregistration Challenge (2017). Retrieved from https://cos.io/prereg/.

Rosenthal, R. (1963). On the social psychology of the psychological experiment: The

    experimenter's hypothesis as unintended determinant of experimental results. *American*

    Scientist, *51*, 268-283.

Quiñones-Vidal, E., López-García, J. J., Peñaranda-Ortega, M., & Tortosa-Gil, F. (2004). The

    nature of social and personality psychology as reflected in *JPSP*: 1965-2000. *Journal of*

    Personality and Social Psychology, 86, 435-452.

Ranehill E., Dreber A., Johannesson M., Leiberg S., Sul S., Weber R. A. (2015). Assessing the

    robustness of power posing: No effect on hormones and risk tolerance in a large sample

    of men and women. Psychological Science, *26*, 653–656.

Rijsman, J., & Stroebe, W. (1989). The two social psychologies or whatever happened to the

    crisis? *European Journal of Social Psychology*, *19*, 339-343.

Richard, F. D., Bond Jr, C. F., & Stokes-Zoota, J. J. (2003). One hundred years of social

    psychology quantitatively described. *Review of General Psychology*, *7*, 331–363.

Ring, K. (1967). Experimental social psychology: Some sober questions about some frivolous

    values. *Journal of Experimental Social Psychology*, *3*, 113-123.

Roll-Hansen, N. (2017). A historical perspective on the distinction between basic and applied

    science. *Journal for General Philosophy of Science*, *48*, 535-551.

Schaller, M. (2016). The empirical benefits of conceptual rigor: Systematic articulation of

    conceptual hypotheses can reduce the risk of non-replicable results (and facilitate novel

    discoveries too). *Journal of Experimental Social Psychology*, *66*, 107-115.

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize?.

    Journal of Research in Personality, *47*, 609-612.

Schwartz, B. (1997). Psychology, idea technology, and ideology. *Psychological Science*, *8*, 21–27.

Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology*, *51*, 515–530.

Sears, D. O., & Freedman, J. L. (1967). Selective exposure to information: A critical review. *Public Opinion Quarterly*, *31*, 194-213.

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

Sherif, M. (1970). On the relevance of social psychology. *American Psychologist*, *25*, 144–156. http://doi.org/10.1037/h0029429

Sherif, M. (1977). Crisis in social psychology: Some remarks towards breaking through the crisis. *Personality and Social Psychology Bulletin*, *3*, 368–382.

Silverman, I. (1971). Crisis in social psychology: The relevance of relevance. *American Psychologist*, *26*, 583.

Simmons, J. P., Nelson, L. D. & Simonsohn, U., (2012). A 21 word solution. Retrieved from https://ssrn.com/abstract=2160588. http://dx.doi.org/10.2139/ssrn.2160588

Smith, M. B. (1973). Is psychology relevant to new priorities? *American Psychologist*, *28*(6), 463–471. https://doi.org/10.1037/h0035096

Smith, N. C. (1970). Replication studies: A neglected aspect of psychological research. *American Psychologist*, *25*, 970-975.

Smith, N. C., Goldstein, D. G., & Johnson, E. J. (2013). Choice without awareness: Ethical and policy implications of defaults. *Journal of Public Policy & Marketing*, *32*, 159-172.

Snyder, M. (1993). Basic research and practical problems: The promise of a" functional" personality and social psychology. *Personality and Social Psychology Bulletin*, *19*, 251-264.

Social and Behavioral Sciences Team (2017). Retrieved from https://sbst.gov/.

Spellman, B. A. (2015). A short (personal) future history of Revolution 2.0. *Perspectives on Psychological Science*, *10*, 886–899.

Stewart, N., Ungemach, C., Harris, A. J., Bartels, D. M., Newell, B. R., Paolacci, G., & Chandler, J. (2015). The average laboratory samples a population of 7,300 Amazon Mechanical Turk workers. *Judgment and Decision Making*, *10*, 479.

Stillman, S. (January 23, 2017). Can behavioral science help in Flint? *The New Yorker.* Retrieved from http://www.newyorker.com/magazine/2017/01/23/can-behavioral-science-help-in-flint

Sumner, P., Vivian-Griffiths, S., Boivin, J., Williams, A., Venetis, C. A., Davies, A., … Chambers, C. D. (2014). The association between exaggeration in health related science news and academic press releases: retrospective observational study. *BMJ : British Medical Journal*, *349,* g7015.

Tate, C. U. (2015). On the overuse and misuse of mediation analysis: It may be a matter of timing. *Basic and Applied Social Psychology, 37,* 235–246.

Thoreson, R. W., Krauskopf, C. J., Mcaleer, C. A., & Wenger, H. D. (1972). The future for applied psychology. *American Psychologist*, *27*, 134-139.

Trafimow, D., & Marks, M. (2015). Editorial. *Basic and Applied Social Psychology, 37,* 1-2.

Tressoldi, P. E., Cumming, G., Fresc, L., & Giofré, D. (2016). *Effects of the Psychological Science New Statistical, Research Disclosure Statements and Open Practices Submission*

Guidelines from 2013 to 2015 (SSRN Scholarly Paper No. ID 2712603). Rochester, NY: Social Science Research Network. Retrieved from https://papers.ssrn.com/abstract=2712603

Turner, E. H., Matthews, A. M., Linardatos, E., Tell, R. A., & Rosenthal, R. (2008). Selective publication of antidepressant trials and its influence on apparent efficacy. *New England Journal of Medicine*, *358*, 252-260.

Vankov, I., Bowers, J., & Munafò, M. R. (2014). On the persistence of low power in psychological science. *The Quarterly Journal of Experimental Psychology*, *67*, 1037-1040.

Vazire, S. (2016). Editorial. *Social Psychological and Personality Science, 7,* 3–7.

Vitelli, R. (1988). The crisis issue assessed: An empirical analysis. *Basic and Applied Social Psychology*, *9*, 301–309.

Wang, M. T. M., Bolland, M. J., Gamble, G., & Grey, A. (2015). Media coverage, journal press releases and editorials associated with randomized and observational studies in high-impact medical journals: A cohort study. *PLOS ONE*, *10*, e0145294. https://doi.org/10.1371/journal.pone.0145294

Ward, L. M. (1992). Mind in psychophysics. In D. Algom (Ed.), *Psychophysical approaches to cognition* (pp. 187-250). London: Elsevier.

Webber, D., Babush, M., Schori-Eyal, N., Vazeou-Nieuwenhuis, A., Hettiarachchi, M., Bélanger, J. J., ... & Gelfand, M. J. (2018). The road to extremism: Field and experimental evidence that significance loss-induced need for closure fosters radicalization. *Journal of Personality and Social Psychology*, *114*, 270-285.

Wegner, D. M. (1992). The premature demise of the solo experiment. *Personality and Social Psychology Bulletin*, *18*, 504-508.

Wells, G. L., Memon, A., & Penrod, S. D. (2006). Eyewitness evidence: Improving its probative value. *Psychological Science in the Public Interest*, *7*, 45-75.

Wicker, A. W. (1969). Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects. *Journal of Social Issues*, *25*, 41-78.

Wiener, R. L., Krauss, D. A., & Lieberman, J. D. (2011). Mock jury research: Where do we go from here?. *Behavioral Sciences and the Law*, *29*, 467-479.

Williamson, V. (2016, January). On the ethics of crowdsourced research. Retrieved 6 April 2017, from /core/journals/ps-political-science-and-politics/article/on-the-ethics-of-crowdsourced-research/B1BDFB1111B416DD0B71540CD6E7D94F

Witte, E. H., & Brandt, V. (2011). Social psychological research: The comparison of four journals. Retrieved from http://psydok.sulb.uni-saarland.de/volltexte/2011/2743

Zhou, H., & Fishbach, A. (2016). The pitfall of experimenting on the web: How unattended selective attrition leads to surprising (yet false) research conclusions. *Journal of Personality and Social Psychology*, *111*, 493-504.