
INVITED REVIEW

Enhancing transparency of the research process to increase accuracy of findings: A guide for relationship researchers

LORNE CAMPBELL,^a TIMOTHY J. LOVING,^b AND ETIENNE P. LEBEL^c

^aThe University of Western Ontario, Ontario, Canada; ^bThe University of Texas at Austin; and ^cMontclair State University

Abstract

The purpose of this paper is to extend to the field of relationship science, recent discussions and suggested changes in open research practises. We demonstrate different ways that greater transparency of the research process in our field will accelerate scientific progress by increasing accuracy of reported research findings. Importantly, we make concrete recommendations for how relationship researchers can transition to greater disclosure of research practices in a manner that is sensitive to the unique design features of methodologies employed by relationship scientists. We discuss how to implement these recommendations for four different research designs regularly used in relationship research and practical limitations regarding implementing our recommendations and provide potential solutions to these problems.

The best single rule may be Feynman's [1986] principle of total scientific honesty. ... This is not nearly so easy as it seems since it is natural to become infatuated with one's own ideas, to become an advocate, to be a much gentler critic of one's own work than one is of others'. Many of us are able to tear other people's research limb from limb while we smile upon our own like an indulgent parent. In fact, I think one should be protective at first until the toddler at least can stand erect. But before one lets the little devil out into the neighborhood, one must learn to look at it as critically as others will. (Lykken, 1991, p. 37)

From January 2014, the journal *Psychological Science* has introduced several changes to its publication standards, including promoting

more open practices in the research process, using what are termed “new” statistics (Cumming, 2014), enhancing disclosure of study methods, and focusing more strongly on evaluating the quality of the methods used to obtain results (Eich, 2014). Similar standards have been put in place from June 2014 at the *Journal of Experimental Social Psychology* (2014). Likewise, the Society for Personality and Social Psychology (SPSP) Task Force on Publication and Research Practices made a number of suggestions regarding how to improve the dependability of research such as using power analyses to justify sample sizes, promoting open practices like disclosing all study materials and procedures, and publishing high-quality replication attempts (Funder et al., 2014). Furthermore, the open access journal *PLoS ONE* now asks authors, as a requirement for publication, to make available the raw data underlying the results reported in published papers. At the time of writing at least six other journals (*Attention, Perception, &*

Correspondence should be addressed to Lorne Campbell, University of Western Ontario, Social Science Centre, London, Ontario, Canada, N6A 5C2, e-mail: lcampb23@uwo.ca.

Psychophysics; AIMS Neuroscience; Cortex; Experimental Psychology; Human Movement Science; Perspective in Psychological Science) devote journal space to registered reports, or articles accepted for publication *ahead of data collection* based on the logic of the hypotheses and strength of the proposed research methods. Manuscripts will be published regardless of study results when authors follow the accepted protocols. Relatedly, the European Union (EU) recently approved the new Clinical Trials Regulation, or laws geared toward improving transparency in research conducted by pharmaceutical companies that are testing new drugs (e.g., all drug trials are to be registered prior to data collection, and reports of full results must be available publicly). Motivating all these changes and recommendations is the goal of increasing the accuracy of, and minimizing mistakes in, published research findings, that is, maximizing the probability of getting it right (Lykken, 1991; Nosek, Spies, & Motyl, 2012). Or as Meehl (1967) put it, to better determine whether the results in fact reflect reality.

The voices calling for improved accuracy of published research findings, and hence a more transparent research process, are undoubtedly growing both in number and volume across disciplines in the social sciences (e.g., economics, political science, sociology, public policy). In the field of psychology, this discussion to date has occurred largely in areas dominated by experimental approaches (i.e., cognitive and social psychology). All fields of scientific inquiry, however, share the goal of getting it right in terms of understanding how the world works. The purpose of this paper is to extend this discussion to the field of relationship science and to spark a sincere discourse about what should be best practices in our discipline. Specifically, we hope to demonstrate the benefits that greater transparency of the research process affords to scientific progress in our field. We also provide concrete recommendations for how relationship researchers can transition to greater disclosure of research practices, and therefore increase the accuracy of research findings, in a manner that is sensitive to the unique design features of methodologies employed by relationship scientists across disciplines (e.g., lab-based

experiments, large-scale correlational designs, diary studies, multiwave dyadic designs, etc.).

A Brief History of Concerns Regarding the Accuracy of Published Research Findings

In a survey of published research articles, Sterling (1959) found that the null hypothesis was rejected in 97% of papers that used tests of statistical significance. Thirty-six years later this tendency remained virtually unchanged (Sterling, Rosebaum, & Weinkam, 1995; see also Fanelli, 2010). Recently, O'Boyle Banks, and Gonzalez-Mule (2014) compared the results presented in dissertations with those presented in subsequent peer-reviewed publications that used the same dissertation data. They found a much higher ratio of significant to nonsignificant research findings in the peer-reviewed publications relative to what was observed in the dissertations. They used the term the "Chrysalis Effect" to refer to this rise in the predictive accuracy of results with the same data set, a transformation achieved by dropping statistically nonsignificant results and in some cases altering hypotheses to be consistent with the results that did emerge (i.e., hypothesizing after results are known, or HARKing; Kerr, 1998).

To put in context the negative consequences of publishing almost only statistically significant findings, we need to consider the typical levels of statistical power present in the literature. Statistical power is the probability of detecting an effect of a certain size or larger given it exists (Cohen, 1988). Having an appropriate level of statistical power (at least .80 or 80%, but ideally higher), therefore, is essential for accurately testing hypotheses (e.g., Schimmack, 2012). The average power of most published studies in many fields, however, is very low. In 1962, Cohen reported the mean power to detect medium effect sizes in published articles was .48, and he advocated strongly for increasing the power of studies (Cohen, 1962). Despite this call, little improvement in power was achieved in the intervening years (Cohen, 1992; Rossi, 1990). Recently, Bakker, van Dijk, and Wicherts (2012) estimated the average power of studies

in psychology overall to be 35%, and Button et al. (2013) estimated the average power of research in neuroscience, a field increasingly integrated with psychological research, to be even lower (21%). Taken together, the bias for publishing statistically significant effects, paired with the low power of most studies, leads to the conclusion that a non-trivial amount of published findings *must* be false positives (Ioannidis, 2005). Identifying false positive findings is made extremely difficult by the fact that direct replication attempts of prior findings are very rarely conducted and, until recently, seldom published (Makel, Plucker, & Hegarty, 2012; Neuliep & Crandall, 1990, 1993).

The bias toward publishing statistically significant findings is also problematic when we consider the fact that academics' professional success depends largely on publishing (Nosek et al., 2012). For example, evaluation of faculty for academic jobs, promotions, awards, and honors, as well as grant funding typically requires a strong record of publication. As noted by Nosek and colleagues (2012), a conflict of interest therefore exists between the personal interests of researchers to *get it right* and their interest to *get it published* to advance their careers. Publishing research papers customarily requires a neat package of statistically significant effects; the process of accumulating accurate knowledge, however, requires the application of rigorous research methods to test hypotheses, a process that is agnostic regarding the outcomes of a study. Researchers make a number of decisions regarding collection and analysis of data as well as what results to present in research reports. The bias toward publishing statistically significant results can influence this decision-making process by inadvertently encouraging the selective reporting of analyses involving particular variables that reached the threshold of statistical significance and the selective use of data reduction and statistical procedures that yield just statistically significant results (Ioannidis, Munafò, Fusar-Poli, Nosek, & David, 2014; O'Boyle et al., 2014; Simonsohn, Nelson, & Simmons, 2013). These practices enhance the statistical significance of the results in a study, but *not* the accuracy of these

results (Meehl, 1967; Simmons, Nelson, & Simonsohn, 2011).

The Need for Greater Transparency in the Research Process

The ultimate goal of science is to accumulate accurate explanations of how our world actually works. Knowledge accumulation has been achieved over hundreds of years via application of the scientific method, an approach that is the bedrock of the social sciences. Accordingly, the accuracy of scientific explanations rests on the strength of the methods used to test hypotheses involving associations between theoretical constructs via specific operationalizations of those constructs (Brewer, 2000). Importantly, the *only way* to accurately evaluate the quality of scientific explanations is to meticulously inspect the robustness of the research process used to provide empirical support for these explanations. Transparency in the research process, therefore, is an essential component of the scientific method because it is the only window through which we have access to this process.

The current approach to conducting research and presenting the outcomes of research in peer-reviewed publications, however, focuses attention primarily on the outcomes of the research process, particularly statistically significant results, at the expense of fully describing how researchers obtained those outcomes. Consider the following hypothetical trajectory of events for a typical research project. A researcher first generates some testable hypotheses, develops one or more ways to test these hypotheses, submits a detailed summary of the proposed research ahead of data collection to the relevant research ethics board for approval, and then collects data. The researcher may stop data collection for a number of reasons, including when the desired number of participants has been recruited, when the available participant pool closes, or when a desired level of statistical significance in key analyses has been achieved. Upon completion of data collection, the researcher conducts analyses to test study hypotheses, likely making adjustments to these analyses during this process (e.g., dropping

outliers, removing psychometrically poor items from a scale, substituting conceptually similar scales in the statistical models, combining different items together to create new scales, adding/removing covariates, testing for moderation or mediation, testing for gender differences, using different statistical analyses, etc.). When a set of statistically significant results are obtained that surpass a *subjective* threshold of explanatory coherence, the researcher then attempts to gauge the overall strength and uniqueness of the findings in order to select possible publication outlets to submit the manuscript for review. The manuscript is carefully crafted to present the results in a theoretical context while simultaneously highlighting the novelty of the findings in an attempt to convince others of the *substantial theoretical contribution* of the work (e.g., Giner-Sorolla, 2012). The researcher now waits for a decision from an editor. If the researcher is offered the opportunity to submit a revised manuscript for further review there are likely a number of concerns raised by the reviewers that need to be addressed, requiring substantial edits as well as additional data analyses. The final version of the paper accepted for publication in all likelihood differs from the initial submission in many ways, including the addition, and/or elimination, of analyses and results.

The amount of space devoted to the methods section in published papers, however, does not allow for a thorough discussion of *all* study details—from the generation of hypotheses to the selection of results to include in a paper. Furthermore, the results that are presented may only reflect a subset of the results of all the analyses conducted both initially and during the review process; the majority of the details of the research process are not included in the manuscript. Evaluating the quality of the scientific explanation proffered for the phenomena under study in published papers can therefore be a challenging or impossible task under current reporting standards (LeBel et al., 2013). Presently, “surplus” details of the research process (which we and others argue are anything but surplus) are simply not available publicly for the overwhelming majority of studies submitted for review and ultimately published

(but see PsychDisclosure.org for exceptions). Consequently, the current level of transparency in the research process is not sufficient to permit precise evaluations of the outcomes of the research (e.g., Meehl, 1990), a state of affairs that, as we have discussed, can directly undermine the accuracy of published research results.

How to Achieve Greater Transparency in the Research Process

A number of suggestions regarding how to achieve greater transparency in the research process have been made (e.g., Donnellan, Roisman, Fraley, & Lucas, 2013, Summer; Kashy, Donnellan, Ackerman, & Russell, 2009; LeBel et al., 2013; Miguel et al., 2014; Simmons et al., 2011). Many of these suggestions are reflected in Simmons, Nelson, and Simonsohn’s (2012) “21 Word Solution”: “We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.” Simmons et al. suggested authors should always include these 21 words in their research papers as well as appropriate methodological details to back up the claims made in this statement. This solution simply asks researchers to briefly disclose details of study design (e.g., how sample size was determined, all of the procedural details of the study), as well as decisions made when analyzing data (e.g., were some study conditions, or measures, found not to be important for predicting outcomes? were some participants excluded from analyses?). Whereas this 21-word statement itself would not take up much space in a research report, disclosing all the suggested details could add many pages and disrupt the flow of a paper. Thus, it has been suggested to include these details in supplementary material as needed.

How well do suggestions like the 21-word solution apply to relationship scientists that routinely employ varied research methods including, but not limited to, extensive longitudinal designs? For example, disclosing “all measures in the study” may be fairly straightforward for experiments with a few (manipulated) independent variables and dependent variables. Staying in line with

such a requirement becomes more complex to include in a research report, however, when a researcher is using a small subset of data, drawn from a multiwave longitudinal data set that includes self-report and behavioral measures. Where, and when, should the researcher disclose this voluminous amount of information? In such a study it is also the case that the data will be used to test a number of different hypotheses over a prolonged period of time, many that will be developed well after the original study was conceived (e.g., Simpson, Collins, Tran, & Haydon, 2007). In these instances the researcher is using existing data to test new hypotheses, but the data were not originally collected to specifically test these hypotheses (although, in most cases, an inkling of an idea existed to spur the researcher to include specific additional measures). This fact makes it impossible to make definitive disclosures regarding how sample sizes were determined to test current hypotheses given that the current hypotheses did not guide the design of the study.

Solutions such as the one discussed above, therefore, may not directly translate to more complex research designs that are typical in relationship science. Of course, there are other challenges associated with adopting these solutions in relationship science; we address the most obvious of these challenges when discussing our recommendations for how we can move toward greater transparency in the research process. Specifically, the goal of our recommendations is to transition toward routinely making more information available publicly that will assist others in appraising key aspects of the research process. As such, our recommendations emphasize disclosure, via preregistered protocols where appropriate, of (a) study hypotheses, (b) procedures and methods, (c) study materials, (d) data analytic plan(s), and (e) a postanalysis disclosure that addresses follow-up exploratory analyses. We also discuss the sharing of raw data and syntax. We recommend using extant online platforms, such as the OpenScienceFramework.org (OSF), to create the recommended preregistered protocols, and to register protocols for studies already conducted or in progress, given that such platforms maintain

a detailed record of when protocols were uploaded. Once “locked in,” a researcher is no longer able to make changes to uploaded files; note, however, that additional material can be added to any project at any time.

Recommendations for Improving Transparency in the Research Process for Relationship Scientists

In this section we make several recommendations regarding specific aspects of the research process that can be disclosed at different points along a project’s timeline. For ease of presentation we discuss the disclosure of different aspects of the research process separately, but researchers can combine some of these elements in their actual registered protocols as appropriate. In Table 1 we list each of the recommended disclosures and provide brief summaries of the material we recommend disclosing in registered protocols.

There are three terms we use frequently when discussing our recommendations and how to implement them: disclosure, preregistered protocol, and registered protocol. To clarify, by disclosure we are simply referring to the release or reporting of information. Each type of registered protocol refers to one or more disclosure statements made available publicly by posting them to an online platform such as the OSF and locking the files. A protocol is considered preregistered when it is publicly posted prior to data collection or data analysis.

Disclosure of hypotheses

Research studies are typically conducted to test particular hypotheses and we recommend that researchers disclose a priori hypotheses either in advance of data collection or prior to the analysis of existing data (e.g., Van ’t Veer, Stel, van Beest, & Gallucci, 2014). When a study is primarily exploratory in nature, the researcher can disclose the primary research questions, or goals of data collection (e.g., item analysis in the development of a new scale; to assess the base rate of a given set of behaviors in target populations; qualitative analysis of interviews).

Table 1. List of disclosure/transparency recommendations and summary of information to be disclosed

Transparency recommendations	Summary
1. Disclosure of study hypotheses ^a	A priori predictions as well as exploratory research questions.
2. Disclosure of procedures and methods ^a	All planned manipulations for studies that include experimental conditions for a target research question as well as all details of the procedures used to collect all data.
3. Disclosure of study materials ^a	All scales to be administered in a study, all items that are part of behavioral coding, as well as all data to be collected from other sources should be disclosed for a target research question.
4. Disclosure of analytic plans ^a	A discussion of the independent (e.g., manipulations, measured individual differences) and dependent variables from the study used to test hypotheses that are the primary focus of analyses. Also, a discussion of the statistical tests planned to be used to test study hypotheses.
5. Postanalysis disclosure of exploratory analyses	Discussion of differences between the analyses presented in the a priori data analytic plan and the analyses used to generate results presented in the research report, including a discussion of exploratory analyses guided by the results of the planned analyses.

^aDisclosures to be included in publicly preregistered protocols, except in instances where noted in the text (e.g., using existing data to test new hypotheses allow for preregistered protocols of new hypotheses and analytic plans, but not the existing procedures and methods or study materials).

Disclosure of procedures and methods

We recommend that researchers disclose all planned manipulations for studies that include experimental conditions for a target research question, as well as all details of the procedures used to collect all data, including a data collection timeline (e.g., timing between assessments or components of a laboratory study). Researchers should also indicate the population(s) from which participants will be recruited and provide a rationale for the targeted sample size. Ideally, the discussion of sample size will include a discussion of the estimated effect size of the primary effect(s) of interest, the desired statistical power for the study, and a calculation of the number of participants required to achieve this level of statistical power given the study design.

There will likely be instances when a priori power analyses to determine sample size for a hypothesized effect using a particular method of data collection may not be readily available. For example, in some instances the appropriate statistical tools for conducting power analyses may not yet exist (e.g., predicting a three-way dyadic interaction between continuous variables assessed daily for 21 consecutive days from both members of dating couples). In such instances, one productive approach is simply to recruit the largest sample of available participants that the researchers can afford, as well as acquire as many repeated measurements as is feasible given subject fatigue constraints from participants in longitudinal designs. One can then estimate the minimum effect size that is reliably detectable (e.g., at a 95%

power level) for those sample sizes. From this perspective, observed null results become much more interpretable because one can nonetheless conclude (assuming integrity of manipulations and measurement instruments; LeBel & Peters, 2011) that the effect under study may exist but would require substantially larger sample sizes to reliably detect (LeBel, 2014; Simonsohn, 2014). Ultimately, the fundamental point here is that specific factors used to determine the target sample size are disclosed in detail, thus making fully transparent the researchers' decision-making process.

Ideally, these disclosures should be made prior to data collection via a publicly pre-registered protocol, but researchers can also register a protocol for studies that they are currently running or for which they have already completed data collection. In instances where researchers use existing data sets to test new hypotheses, the researcher should register a protocol, or provide a link to an existing protocol, discussing the rationale for the target population and sample size in the original study. In such cases, we recommend that researchers also note specifically where overlap does versus does not occur with prior analyses and publications (including those under review) as well any deviations between prior and current analyses (e.g., use of different statistical controls when predicting the same dependent variable).

Disclosure of study materials

All scales to be administered in a study, all items that are part of behavioral coding, as well as all data to be collected from other sources (e.g., daily weather conditions as part of a diary study on mood and relationship perceptions) should be disclosed for a target research question. A list of scales should be provided with citation information, and researchers should explicitly note any modifications of existing scales including, but not limited to, rewording of instructions, items, and scale anchors. When an abbreviated scale is employed, the specific items included in the abbreviated measure should be noted. Further, all items should be provided for scales

created by the author(s) for a study. Additionally, the order that scales are presented should be clearly indicated, and specific a priori decision rules for scale creation detailed (e.g., minimum factor-loading requirements and/or cross-loading cutoffs). Ideally, these disclosures should be made prior to data collection via a preregistered protocol. Again, however, researchers can register a protocol to disclose these details for studies that they are currently running or for which they have already completed data collection.

Disclosure of data analytic plan

Researchers should disclose the specific analyses that are planned to test study hypotheses (for an example "preanalysis plan" from economics, see Casey, Glennerster, & Miguel, 2012). This disclosure includes a discussion of the independent (e.g., manipulations, measured individual differences) and dependent variables from the study used to test hypotheses that are the primary focus of analyses. Also, researchers should include a discussion of the statistical tests they plan to use to test study hypotheses (e.g., analysis of variance, multiple regression, hierarchical linear modeling, structural equation modeling). Planned exploratory analyses, including the variables to be included in these analyses, should also be discussed. Regardless of whether planned analyses use data from a new study or from a portion of existing data, disclosures should be made via preregistered protocols prior to conducting the analyses in order to clearly distinguish exploratory from confirmatory analyses reported in the published article (e.g., Wagenmakers, Wetzels, Borsboom, van der Maas, & Kievit, 2012).

Postanalysis disclosure

We recommend that researchers disclose any differences between the analyses presented in the a priori data analytic plan and the analyses used to generate results presented in the research report. This disclosure may include a discussion of items dropped (or added) to scales, study conditions that are dropped from analyses, as well as participants that are

excluded from analyses for any reason (e.g., outliers, noncompliance, incomplete data; ideally, exclusion criteria will be noted in advance of data collection). This section should also include a discussion of exploratory analyses guided by the results of the planned analyses. The results that are discussed in the research report might sometimes include a mixture of confirmatory and exploratory analyses (i.e., a hybrid approach), but whether specific analyses are confirmatory or exploratory should be clearly stated for all findings.

Sharing of raw data and metadata

We full-heartedly agree with new data-sharing policies recently enacted by leading journals. *PLoS ONE* now requires all researchers to make publicly available all raw data required to reproduce the results of every analysis reported in a submitted article. Likewise, the SPSP now requires all authors to share raw data with other competent researchers seeking to verify the empirical claims underlying the reported findings.¹ When adopting open data practices, researchers also need to consider specific rules of local research ethics boards regarding data sharing (Burnham, 2014).

Examples of How to Implement these Disclosure Statements with Different Study Designs

In this section we discuss four prototypical examples of the varied research methods routinely used by relationship researchers, and how researchers can approach the recommended disclosures when using these types of study designs. We also want to stress that disclosure statements included in all registered protocols should be informative but concise as possible, getting directly to the point when providing the suggested details of the study.

Laboratory-based experiment

As we note above, much of the discussion regarding accuracy of research findings has focused on experimental psychology. Relationship researchers are not immune to this discussion; experimental designs are

regularly employed in relationship science (e.g., Karremans, Dotsch, & Corneille, 2011; Murray, Holmes, MacDonald, & Ellsworth, 1998; Smith, Loving, Crockett, & Campbell, 2009). When experimental designs are employed, researchers can disclose via a preregistered protocol the hypotheses to be tested, the proposed study design, all study procedures, a list of all scales to be administered (including any items created by the researchers to serve as the primary independent and/or dependent variables), as well as the data analytic plan. Many of these details will have already been put together as part of the application submitted to the local ethics board and if so can be used when putting together the disclosure statements for the protocol to be registered. Following data collection and data analyses, the researchers could then create the postanalysis disclosure statement to point out differences between planned and exploratory analyses (if any), and make available this statement on the OSF. The manuscript submitted for submission should contain a link to these files (e.g., LeBel & Campbell, 2013; see also LeBel & Wilbur, 2014) to allow editors and reviewers (and ultimately future consumers of the research) to be in a better position to accurately evaluate the reported results.

Single-sample correlational design

In many respects, the most commonly employed design in our field falls under this label. Relationship researchers regularly test hypotheses regarding associations between relationship-relevant constructs (e.g., Campbell, Simpson, Kashy, & Fletcher, 2001; Murray, Holmes & Griffin, 1996; Slotter & Luchies, 2014). The same process for putting together and making available the preregistered protocol containing the recommended disclosures outlined in the prior example can be followed in this case.

It is possible, however, that researchers may develop hypotheses at a later date and utilize a preexisting data set to test these hypotheses given that responses were collected for a number of questionnaires that were not the focus of the original study hypotheses. In this instance,

the researchers should, to the best of their abilities and available study information, create and register protocols disclosing procedures, methods, and materials for the original study. Additionally, the new hypotheses, as well as the data analytic plan to test these hypotheses, should be preregistered in advance of analyzing the data from this existing data set, indicating the variables from the prior data set selected for analysis. The preregistered protocol of new hypotheses and the data analytic plan should also provide a link to the registered protocol of the original study, something that will be made easier going forward as researchers more routinely post preregistered protocols disclosing details of new studies.

Diary study

Many relationship researchers conduct diary studies (Bolger, Davis & Rafaeli, 2003), collecting data from individuals in relationships, or both partners, over a set period of time (e.g., 14–28 consecutive days; e.g., Birnbaum, Reis, Mikulincer, Gillath & Orpaz, 2006; Campbell, Simpson, Boldry & Kashy, 2005). Diary studies typically involve the collection of data across several measures at a prediary intake session, collection of responses during the diary period in a manner suited to testing the primary study hypotheses (e.g., event, signal, or interval contingent responses), and sometimes collection of additional data following the completion of the diary period. These studies are rarely designed to test only one set of hypotheses given the time and resources that are involved in designing and running them. Additionally, not all hypotheses involving the data are established at the time of data collection. That is, researchers may develop hypotheses at a later date and utilize the preexisting data set to test these hypotheses.

The same process for putting together and making available the preregistered protocol outlined in the prior examples can be followed in this example. The initial data analytic plan will include the first set of hypotheses to be tested, along with a discussion of the measures that will be used out of the basket of measures available to test these hypotheses. If the researchers would like to use other data

collected as part of a diary project to test additional hypotheses, the new hypotheses as well as the data analytic plan to test these hypotheses can be preregistered in advance of analyzing the existing data, indicating the variables from the prior data set selected for analysis.

Multiwave longitudinal study

A number of novel and important findings have resulted from large-scale multiwave longitudinal studies that have collected massive amounts of data over prolonged periods of time. For example, Simpson et al. (2007) tested hypotheses generated from an attachment theoretical perspective using data collected from a sample of individuals beginning when these participants were infants and ending when they were young adults. Also, Ted Huston and colleagues have followed a large sample of married couples since 1981 beginning when the couples first married, culminating in numerous publications on the importance of relationship processes at different times in marriage (e.g., Huston, Caughlin, Houts, Smith, & George, 2001).

The same process for putting together and making available the preregistered protocols outlined in the diary study case can be followed in this example. The initial data analytic plan will include the first set of hypotheses to be tested, along with a discussion of the measures that will be used out of the basket of measures available to test these hypotheses. It is very likely, however, that some study details, particularly at later stages of data collection, will not be known at the launch of the study. For example, the researcher may add new design features to later stages of the study (e.g., to test additional hypotheses, to follow up on unexpected effects emerging with data collected in the earlier stages). Additionally, it is possible that a researcher will decide to continue data collection beyond the expected time frame of the study because new funding becomes available. As new features are added to the study, the researcher can create new registered protocols as appropriate and add, or link them, to the original registered project.

Practical Challenges and Possible Solutions

In this section, we consider a number of practical challenges to adopt the suggested recommendations and we provide some possible solutions to these challenges.

A burden on researchers

It could be argued that disclosing the recommended information via the suggested registered protocols during the research process would create a cumbersome bureaucratic process that takes time away from actually conducting research. However, this potential burden is offset, in part or in full, by (a) greater confidence in the research that forms the basis for new ideas (vs., e.g., conducting a study that builds on false positive results) and (b) improving the quality of the research process as to increase the probability of getting it right. Many of the suggested disclosures already form the basis of the research ethics applications that researchers are required to submit in advance of data collection, so it should not take an unreasonable amount of time to edit this material for inclusion in the protocols to be registered. Adopting the recommendations will admittedly require more time and attention of the researcher in the earlier stages of the research process compared to current practices that typically do not involve public disclosure of details of the research process prior to submitting an article for peer review, but this material can potentially save time for the researcher at the time of writing articles given that the bulk of the methodological details of the study have already been carefully thought through and written. Over time, the recommendations will, ideally, simply become part of the research process itself, but it will take a generation of researchers to achieve this shift.

A handicap for researchers

More fully disclosing details of the research process via the suggested registered protocols may handicap researchers in the review process relative to researchers that do not engage in these practices. Stated differently, providing

much more information about the research process to reviewers may open the door for them to find more reasons to recommend rejection of a paper. Moreover, the analyses disclosures we suggest will likely result in the reporting of results that are not as consistently statistically significant as those presented in manuscripts from researchers choosing not to disclose as many details of their research process (albeit the results will be more accurate than is customarily reported). This could give reviewers additional ammunition to recommend rejection in light of the bias for statistically significant findings.

Additionally, disclosing data analytic plans in advance of data collection and/or analysis can result in the presentation of results derived from a mix of confirmatory and exploratory analyses, perhaps leading reviewers to request the replication of the exploratory results as a condition for publication (as suggested by Wagenmakers et al., 2012). This may be an unreasonable request, however, in many instances. For example, if researchers present results from a several-decade-long longitudinal study, with data collection beginning when participants were infants and ending when participants were young adults (e.g., Simpson et al., 2007), it would be virtually impossible to directly replicate any exploratory or unexpected findings with a new sample (see also Meehl, 1990). Preventing publication of results from this research would essentially signal to the research community to cease collecting longitudinal data over such time periods, a situation that is arguably as problematic as publishing results relying to some degree on exploratory data analyses.

One possible solution to this challenge is that editors will have to police the process a bit more until the field fully embraces more transparent reporting. For example, just as the field frowns upon researchers' cherry-picking results or measures to increase chances of publication, editors should equally frown upon reviewers cherry-picking from study details to create an unfair standard by which research is evaluated (Maner, 2014). When researchers are presenting results from longitudinal data sets, preregistering study hypotheses, for example,

is a mechanism to demonstrate the confirmatory nature of the research in advance.

Disclosures for existing data sets

As we have discussed in some of our examples of different research designs used by relationship researchers, some researchers will have very large data sets composed of a variety of methods of data collection, perhaps over very long periods of time. These researchers may reasonably wonder about the challenges involved in disclosing details of the research process via the suggested registered protocols for large and expensive preexisting data sets. At minimum, we suggest that all new hypotheses to be tested with this data, as well as data analytic plans, be preregistered. Additionally, disclosure of all relevant procedures and materials should now be registered. Ideally, the researcher could register all the procedural details, as well as materials, for existing data sets that will be used to provide results for future research reports, and then the preregistered protocols of new hypotheses and data analytic plans could simply provide a link to this registered protocol.

Researchers may also have access to similar types of data sets that have been collected by others. Is it appropriate for them to disclose all of the methodological information outlined above in registered protocols? Or is it the responsibility of the original researcher? We recommend that when researchers publish results from such existing data sets, they should preregister hypotheses and data analytic plans as well as disclose the relevant procedures, materials, and measures that pertain to the target research question at hand in a registered protocol.

The rights of researchers

The suggested preregistered and registered protocols are meant to increase transparency of the research process, but it could be argued that disclosures within these protocols inadvertently provide a means for other researchers to use the disclosed methods and procedures for their own research. If this were to occur, the researchers using the methods and procedures

disclosed in the registered protocols of others would need to properly cite the source of their methodological inspiration. This task is made easy given that the protocols would have a dedicated web link—which is actually citable for projects on the OSF—containing a detailed history of when all materials were posted and disclosed. Also, if the methods and procedures represent an advance in studying a given relational process, then having others use and properly cite their use of these methods and procedures may represent a significant advance to the field (e.g., by accelerating the spread of ideas that is often severely hindered by the current publication lag in our discipline).

There are a few different approaches, however, that a researcher could adopt when disclosing details of the research process. First, the researcher could choose to only disclose in registered protocols the procedures, methods, and materials that are relevant for the results presented in a given research report, and then post registered protocols for the same data set as additional research reports are written (i.e., a piecemeal approach). To be clear, all future hypotheses and data analytic plans using the same data set would need to be preregistered and clearly list all procedures, methods, and materials the researcher plans to use to test hypotheses. The postanalytic disclosure would also need to be posted.

A second approach, one that we advocate, is to preregister all known study procedures, methods, and materials up front prior to data collection. Additional details (e.g., measures added after the study began) can be added in separate registered protocols over time as appropriate. All this information can also be registered for studies where data collection is already complete but data analyses are still planned. The researcher would then create separate preregistered protocols for hypotheses and data analytic plans using the data set, clearly detailing the materials to be used in the analyses and at what point in the study these measures were collected.

It is worth noting that openly disclosing all study details as recommended may actually reduce rather than increase “idea stealing” (or “scooping”), as there is a written record of when a study’s details were posted.

As a result, for example, if a researcher uses an unpublished measure disclosed by another researcher, this fact will need to be cited by the borrowing author given the public record of the disclosure of this scale in the relevant protocol. The same logic applies to hypotheses disclosed in advance of data collection or data analysis. Failure to properly cite the source of borrowed measures, procedures, or hypotheses would be akin to plagiarism.

Why Should Relationship Researchers Embrace these Recommendations?

As suggested by Lykken (1991) in the quote appearing at the beginning of this article, cumulative science requires scientific honesty. Ideally we should not need to persuade researchers of the benefits of disclosing details of the research process; instead, researchers should need to provide solid rationale for not openly sharing these details. That said, we want to emphasize that we are not advocating that the recommended disclosure practices be rendered mandatory and, thus, a regulatory committee govern the research process. Instead, it is our belief that the benefit of these practices to individual researchers and, by extension, the *field* of relationship science, are self-evident in that they improve the transparency of the research process and thereby increase the accuracy of the results presented in research papers (Lykken, 1991). It is our contention that results emanating from more, compared to less, transparent research practices will over time be proven to be more robust (e.g., Hanson, 1958), thereby earning higher confidence from the research community. If so, in due course there should be a growing trend toward published research coming from researchers that have adopted more open practices given the higher confidence earned for this research, further incentivizing more transparent research practices.

We also want to acknowledge that our proposed recommendations may have limitations that we have not yet considered or anticipated, but we advance the recommendations in the spirit of stimulating a constructive debate among relationship scientists about how best to achieve greater transparency in our research

reporting practices. We also acknowledge that additional challenges may exist in embracing greater transparency in the research process in relationship science that we did not discuss and hence need to be considered. It is very likely, therefore, that important amendments will be made to these recommendations over time, and that new recommendations will be put forward. Overall, it is our view that the practical problems facing relationship researchers during the transition to adopting more transparent research practices represent challenges that can, and indeed must, be overcome with a bit of will and ingenuity. In this spirit, we welcome an open discussion of our recommendations as well as the introduction of new ideas from other relationship researchers.

Conclusions

In September 2011 researchers using the Oscillation Project with Emulsion-tRacking Apparatus (OPERA) instrument observed neutrinos apparently traveling faster than the speed of light (Adam et al., 2011). News of the observation spread quickly all over the world, but the researchers did not automatically conclude that Einstein's special theory of relativity had been disconfirmed. Rather, researchers shared sufficient details of their experiment and results so that independent labs could accurately evaluate and help explain what they had observed. Other labs using better instruments conducted direct replication attempts but failed to measure neutrinos traveling faster than the speed of light (e.g., Brumfiel, 2012). The discrepant findings were later explained when the original researchers noticed a faulty cable connection with their equipment, a problem that was subsequently shown to result in an inaccurate recording of the speed that neutrinos were traveling.

When presented with extraordinary results, it perhaps seems obvious that details of the research process *should* have been made readily available by the original researchers, and that independent direct replications of the experiment *should* have been conducted to determine if these results could be reproduced. When dealing with such serious matters as exceeding the speed of light and potential time

travel, the expectation is that the accuracy of the findings *should* be rigorously verified, and the only way to engage in such verification is by knowing exactly what was done originally.

Relationship researchers are of course not collecting data that could potentially disconfirm Einstein's special theory of relativity. However, we argue that the contributions that relationship researchers make to the knowledge base of relationship processes are no less important than those physicists make to the knowledge base of natural science. We should therefore hold our results to the same high level of scrutiny as those of the original OPERA measurements. Doing so necessitates embracing transparency practices, like the ones we have recommended, to ensure that the research process is sufficiently transparent so that we can collectively arrive at an accurate understanding of relationship phenomena. To do otherwise—for example, to argue that transparency of the research process for relationship science is not essential for evaluating the accuracy of reported findings—would be a tacit admission that what we do is not science, polemics perhaps, but not science.

References

- Adam, T., Agafonova, N., Aleksandrov, A., Altinok, O., Alvarez Sanchez, P., Anokhina, A. . . . Zghiche, A. (2011). *Measurement of the neutrino velocity with the OPERA detector in the CNGS beam*. Retrieved from <http://arxiv.org/abs/1109.4897>
- Bakker, M., van Dijk, A., & Wicherts, J. M. (2012). The rules of the game called psychological science. *Perspectives on Psychological Science*, 7, 543–554.
- Birnbaum, G. E., Reis, H. T., Mikulincer, M., Gillath, O., & Orpaz, A. (2006). When is sex more than just sex: Attachment orientations, sexual experience, and relationship quality. *Journal of Personality and Social Psychology*, 91, 929–943.
- Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology*, 54, 579–616.
- Brewer, M. G. (2000). Validity issues and research design. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 3–16). Cambridge, England: Cambridge University Press.
- Brumfiel, G. (2012). Neutrinos not faster than light. *Nature News*. Retrieved from <http://www.nature.com/news/neutrinos-not-faster-than-light-1.10249> doi:10.1038/nature.2012.10249
- Burnham, B. (2014, February 5). *Open data and IRBs* [Blog post]. Retrieved from <http://osc.centerforopencscience.org/2014/02/05/open-data-and-IRBs/>
- Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafo, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14, 365–376.
- Campbell, L., Simpson, J. A., Boldry, J. G., & Kashy, D. (2005). Perceptions of conflict and support in romantic relationships: The role of attachment anxiety. *Journal of Personality and Social Psychology*, 88, 510–531.
- Campbell, L., Simpson, J. A., Kashy, D. A., & Fletcher, G. J. O. (2001). Ideal standards, the self, and flexibility of ideals in close relationships. *Personality and Social Psychology Bulletin*, 27, 447–462.
- Casey, K., Glennerster, R., & Miguel, E. (2012). Reshaping institutions: Evidence on aid impacts using a pre-analysis plan. *Quarterly Journal of Economics*, 127, 1755–1812.
- Cohen, J. (1962). The statistical power of abnormal-social psychological research: A review. *Journal of Abnormal and Social Psychology*, 65, 145–153.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159.
- Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, 25, 7–29.
- Donnellan, M. B., Roisman, G. L., Fraley, R. C., & Lucas, R. E. (2013, Summer). Methodological reform 101 for developmental researchers. *Developmental Psychologist Division 7 Newsletter*, 34–40.
- Eich, E. (2014). Business not as usual. *Psychological Science*, 25, 3–6.
- Fanelli, D. (2010). “Positive” results increase down the hierarchy of Science. *PLoS ONE*, 5(4), e10068. doi:10.1371/journal.pone.0010068
- Feynman, R. P. (1986). *Surely you're joking. Mr. Feynman!*. New York, NY: Bantam Books.
- Funder, D. C., Levine, J. M., Mackie, D. M., Morf, C. C., Vazire, S., & West, S. G. (2014). Improving the dependability of research in personality and social psychology: Recommendations for research and educational practice. *Personality and Social Psychology Review*, 18, 3–12.
- 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.
- Hanson, R. (1958). Evidence and procedural characteristics of “reliable” propositions in social science. *American Journal of Sociology*, 63, 357–370.
- Huston, T. L., Caughlin, J. P., Houts, R. M., Smith, S. E., & George, L. J. (2001). The connubial crucible: Newlywed years as predictors of marital delight, distress, and divorce. *Journal of Personality and Social Psychology*, 80, 237–252.
- Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, 2, e124.

- Ioannidis, J. P., Munafo, M. R., Fusar-Poli, P., Nosek, B. A., & David, S. P. (2014). Publication and other reporting biases in cognitive sciences: Detection, prevalence, and prevention. *Trends in Cognitive Science, 18*, 235–241.
- Journal of Experimental Social Psychology*. (2014). *JESP editorial guidelines*. Retrieved from <http://www.journals.elsevier.com/journal-of-experimental-social-psychology/news/jesp-editorial-guidelines/>
- Karremans, J. C., Dotsch, R., & Corneille, O. (2011). Romantic relationship status biases memory of faces of attractive opposite-sex others: Evidence from a reverse-correlation paradigm. *Cognition, 121*, 422–426.
- Kashy, D. A., Donnellan, M. B., Ackerman, R. A., & Russell, D. W. (2009). Reporting and interpreting research in PSPB: Practices, principles, and pragmatics. *Personality and Social Psychology Bulletin, 35*, 1131–1142.
- Kerr, N. L. (1998). HARKing: Hypothesizing after results are known. *Personality and Social Psychology Review, 2*, 196–217.
- LeBel, E. P. (2014, April 15). *A simpler and more intuitive publication bias index?* [Blog post]. Retrieved from <http://proveyourselfwrong.wordpress.com/2014/04/15/a-simpler-publication-bias-index/>
- LeBel, E. P., Borsboom, D., Giner-Sorolla, R., Hasselmann, F., Peters, K. R., Ratliff, K. A., & Tucker Smith, C. (2013). PsychDisclosure.org: Grassroots support for reforming reporting standards in psychology. *Perspectives on Psychological Science, 8*, 424–432.
- LeBel, E. P., & Campbell, L. (2013). Heightened sensitivity to temperature cues in highly anxiously attached individuals: Real or elusive phenomenon? *Psychological Science, 24*, 2128–2120.
- 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.
- LeBel, E. P., & Wilbur, C. J. (2014). Big secrets do not necessarily cause hills to appear steeper. *Psychonomic Bulletin & Review, 21*, 696–700.
- Lykken, D. T. (1991). What's wrong with psychology anyway? In D. Cicchetti & W. M. Grove (Eds.), *Thinking clearly about psychology: Vol. 1. Matters of public interest* (pp. 3–39). Minneapolis: University of Minnesota Press.
- 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.
- Maner, J. K. (2014). Let's put our money where our mouth is: If authors are to change their ways, reviewers (and editors) must change with them. *Psychological Science, 9*, 343–351.
- Meehl, P. E. (1967). Theory-testing in psychology and physics: A methodological paradox. *Philosophy of Science, 34*, 103–115.
- Meehl, P. E. (1990). Why summaries of research on psychological theories are often uninterpretable. *Psychological Reports, 66*, 195–244.
- Miguel, E., Camerer, C., Casey, K., Cohen, J., Esterling, K. M., Gerber, A. . . . Van der Laan, M. (2014). Promoting transparency in social science research. *Science, 343*(6166), 30–31.
- Murray, S. L., Holmes, J. G., & Griffin, D. W. (1996). The benefits of positive illusions: Idealization and the construction of satisfaction in close relationships. *Journal of Personality and Social Psychology Bulletin, 70*, 79–98.
- Murray, S. L., Holmes, J. G., MacDonald, G., & Ellsworth, P. (1998). Through the looking glass darkly? When self-doubt turns into relationship insecurities. *Journal of Personality and Social Psychology, 75*, 1459–1480.
- 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.
- Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific utopia II: Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science, 7*, 615–631.
- O'Boyle, E. H., Jr., Banks, G. C., & Gonzalez-Mule, E. (2014). The Chrysalis effect: How ugly initial results metamorphosize into beautiful articles. *Journal of Management*. Advance online publication. doi:10.1177/0149206314527133
- Rossi, J. S. (1990). Statistical power of psychological research: What have we gained in 20 years. *Journal of Consulting and Clinical Psychology, 58*, 646–656.
- Schimmack, U. (2012). The ironic effect of significant results on the credibility of multiple-study articles. *Psychological Methods, 17*, 551–566.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*, 1359–1366.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). *A 21 word solution*. Available at SSRN: <http://ssrn.com/abstract=2160588> or <http://dx.doi.org/10.2139/ssrn.2160588>
- Simonsohn, U. (2014). *Go big or go home: A practical alternative to impractical power analysis*. Presented at the annual conference of the Society of Personality and Social Psychology, Austin, TX.
- Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2013). *P-curve: A key to the file drawer*. *Journal of Experimental Psychology: General, 143*, 534–547.
- Simpson, J. A., Collins, W. A., Tran, S., & Haydon, K. D. (2007). Attachment and the experience and expression of emotions in adult romantic relationships: A development perspective. *Journal of Personality and Social Psychology, 92*, 355–367.

- Slotter, E. B., & Luchies, L. B. (2014). Relationship quality promotes the desire for closeness among distressed avoidantly attached individuals. *Personal Relationships, 21*, 22–34.
- Smith, A., Loving, T. J., Crockett, E. E., & Campbell, L. (2009). What's closeness got to do with it? Men's and women's cortisol responses when providing and receiving support. *Psychosomatic Medicine, 71*, 843–851.
- Sterling, T. D. (1959). Publication decisions and their possible effects on inferences drawn from tests of significance—Or vice versa. *Journal of the American Statistical Association, 54*, 30–34.
- Sterling, T. D., Rosebaum, W. L., & Weinkam, J. J. (1995). Publication decisions revisited: The effect of the outcome of statistical tests on the decision to publish and vice versa. *The American Statistician, 49*, 108–112.
- Van 't Veer, A. E., Stel, M., van Beest, I., & Gallucci, M. (2014). Registered report: Measuring unconscious deception detection by skin temperature. *Frontiers in Psychology, 5*, 442. doi:10.3389/fpsyg.2014.00442
- Wagenmakers, E. J., Wetzels, R., Borsboom, D., van der Maas, H. L. J., & Kievit, R. A. (2012). An agenda for purely confirmatory research. *Perspectives on Psychological Science, 7*, 632–638.