THIS ARTICLE HAS BEEN CORRECTED. SEE LAST PAGE

American Psychologist

0003-066X/18/$12.00

2018, Vol. 73, No. 1, 3–25
http://dx.doi.org/10.1037/amp0000191

Journal Article Reporting Standards for Quantitative Research in Psychology: The APA Publications and Communications Board Task Force Report

Mark Appelbaum
University of California, San Diego

Harris Cooper
Duke University

Rex B. Kline
Concordia University, Montréal

Evan Mayo-Wilson
Johns Hopkins University

Arthur M. Nezu
Drexel University

Stephen M. Rao
Cleveland Clinic, Cleveland, Ohio

Following a review of extant reporting standards for scientific publication, and reviewing 10 years of experience since publication of the first set of reporting standards by the American Psychological Association (APA; APA Publications and Communications Board Working Group on Journal Article Reporting Standards, 2008), the APA Working Group on Quantitative Research Reporting Standards recommended some modifications to the original standards. Examples of modifications include division of hypotheses, analyses, and conclusions into 3 groupings (primary, secondary, and exploratory) and some changes to the section on meta-analysis. Several new modules are included that report standards for observational studies, clinical trials, longitudinal studies, replication studies, and *N*-of-1 studies. In addition, standards for analytic methods with unique characteristics and output (structural equation modeling and Bayesian analysis) are included. These proposals were accepted by the Publications and Communications Board of APA and supersede the standards included in the 6th edition of the *Publication Manual of the American Psychological Association* (APA, 2010).

Keywords: reporting standards, research methods, meta-analysis, APA Style

The involvement of the American Psychological Association (APA) in the establishment of journal article reporting standards began as part of a mounting concern with transparency in science. The effort of the APA was contemporaneous with the development of reporting standards in other fields, such as the Consolidated Standards of Report-

Mark Appelbaum, Department of Psychology, University of California, San Diego; Harris Cooper, Department of Psychology and Neuroscience, Duke University; Rex B. Kline, Department of Psychology, Concordia University, Montréal, Quebec, Canada; Evan Mayo-Wilson, Bloomberg School of Public Health, Johns Hopkins University; Arthur M. Nezu, Department of Psychology, Drexel University; Stephen M. Rao, Neurological Institute, Cleveland Clinic, Cleveland, Ohio.

The authors comprise the APA Publications and Communications Board Working Group on Quantitative Research Reporting Standards (Mark Appelbaum, Chair). The order of authorship is alphabetical. This report was prepared with assistance from David Kofalt, Emily Leonard Ayubi, and Anne Woodworth. The group thanks Scott Maxwell, Arthur Stone, and Kenneth J. Sher for their contributions to the 2008 Working Group on Journal Article Reporting Standards that served as the foundation for this article. We also thank David Rindskopf, who authored the reporting standards for the use of Bayesian statistical approaches. We also acknowledge Rick Hoyle, whose 2013 paper with Jennifer Isherwood serves as the basis for the structural equation modeling reporting standards; Robyn Tate (Tate, Perdices, Rosenkoetter, McDonald, et al., 2016), lead author of "The Single-Case Reporting Guideline In BEhavioural Interventions (SCRIBE) 2016 Statement: Explanation and Elaboration" that is the basis of the *N*-of-1 reporting standards; Pamela Cole and others from the Society for Research in Child Development, who advised on the reporting standards for longitudinal designs; Graeme Porte for his comments on the reporting standards for replication studies; Bruce Thompson, Cecil Reynolds, and Ronald Livingston for their comments on reporting standards about psychometrics; and members of the Society for Research Synthesis Methodology for advice on the meta-analysis reporting standards.

Correspondence concerning this article should be addressed to Mark Appelbaum, Department of Psychology, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0109. E-mail: mappelbaum@ucsd.edu

ing Trials (CONSORT; see http://www.consort-statement.org/) in the medical sciences. Work on the APA standards began with the appointment of the first Working Group on Journal Article Reporting Standards (JARS) by the Publications and Communications (P&C) Board of APA in 2006. The report of that committee was received by the P&C Board and subsequently published in the *American Psychologist* (APA Publications and Communications Board Working Group on Journal Article Reporting Standards, 2008). The content of that report and the article was also incorporated into the sixth edition of the *Publication Manual of the American Psychological Association* (hereinafter referred to as the *Publication Manual*; APA, 2010).

In May 2015, the P&C Board of APA authorized the appointment of two working groups: one to revisit and expand the work of the original JARS (JARS–Quant Working Group or Working Group) and the other to establish new standards for reporting qualitative research (JARS–Qual Working Group). This report is the result of the deliberations of the JARS–Quant Working Group and both updates the 2008 article and extends its scope.

Developing New Reporting Standards

The development of reporting standards is an ongoing process. In selecting the reporting standards to include in this report, the Working Group tried to balance several factors. These included perceptions of the frequency of research involving a particular research strategy, experimental design, or analytic strategy; the extent to which an approach needed a separate set of reporting standards; and the state of technical development in the publishing–archiving domains that would allow for the recommended standards. The Working Group made judgments that a different group of individuals may not have made and expects that future groups will continue to develop new standards and modify some that are in the current document. The list of uncovered topics is long, and the next JARS–Quant Working Group may venture into domains that are now just being developed. One example is the development of reporting standards for secondary data analysis. Changes in attitudes about data sharing, new technologies for data sharing, and emerging ideas about the responsible conduct of data-sharing ventures make it likely that reporting standards for secondary data analysis may appear in future versions of reporting standards. The development of reporting standards spans many fields and is an international undertaking. In the process of developing the new standards, the Working Group took into account standards that had been developed in many areas and aspired to utilize features of the existing standards that could be adapted into the scientific needs of the field.

Between Then and Now

Since about the year 2000, many organizations have created or further refined their own sets of reporting standards. Where work on these reporting standards overlap with the work often done by those in the behavioral, social, and psychological sciences, the Working Group has chosen to cite (and, on occasion, incorporate) those standards into JARS–Quant rather than try to develop a complementary set of reporting standards. For example, a few words from the Animal Research: Reporting of In Vivo Experiments (ARRIVE; Kilkenny et al., 2010) guidelines have been incorporated into JARS–Quant to make the two sets of standards for reporting studies using nonhuman living organisms consistent. In the case of reporting standards when neuropsychological measurements are used, recent work on standards for reporting work that includes functional MRI (fMRI; e.g., Nichols et al., 2017), event-related potential (ERP; e.g., Picton et al., 2000), and other neuropsychological measures was cited. On other occasions, sections of other published standards have been adapted into the tables, as in the case of reporting standards for structural equation modeling (SEM) and *N*-of-1 studies. Finally, other groups, including the Society for Research in Child Development and the Society for Research Synthesis Methodology on longitudinal studies and meta-analysis, respectively, provided input, as did individuals with special expertise and insights into a particular issue, including David Rindskopf for standards for reporting the results of studies using Bayesian analyses. For those looking for guidelines for study types not covered in JARS–Quant but with health outcomes, the Enhancing the QUAlity and Transparency Of health Research (EQUATOR) network (http://www.equator-network.org/) currently lists more than 300 different sets of guidelines, including some similar to JARS. The EQUATOR set contains some guidelines that are general (e.g., CONSORT) and some that are very narrowly construed, such as guidelines for reporting studies specific to disease types.

During the same period, there has been a gathering movement to register or preregister randomized control trials and randomized clinical trials (Cybulski, Mayo-Wilson, & Grant, 2016). Although these registrations are most commonly found in the medical domain, increasingly they are appearing for trials with psychoeducational, psychotherapeutic, or related studies. In JARS–Quant, guidance on where to report the registration information for studies that are registered is provided. Some APA journals, at the discretion of their editors, are now requiring registration of some kinds of clinical trials to qualify for publication. Routine registration of psychological studies that involve controlled trials is encouraged in JARS–Quant. There are several ways that studies can be registered, particularly using ClinicalTrials.gov (http://www.clinicaltrials.gov), a registry and results database of publicly and privately supported clinical studies using human participants conducted around the world.

The Structure of JARS–Quant

Recommendations in both the JARS–Quant and the original JARS follow the same basic structure. These recommendations are stated in a series of tables that apply either singly or in combination to cover varying designs of empirical studies. Over time, additional tables (and modules within tables) may be added as new reporting standards emerge.

In the current version, there are three general groups of tables: Tables 1–6, the uses of which are determined by the nature of the inquiry being reported; Tables 7 and 8, the uses of which are dictated by specific statistical–quantitative analyses being reported; and Table 9, which contains reporting standards for research syntheses and meta-analysis. Tables have been designed to be comprehensive and to apply widely. For any individual report, the authors would be expected to select those items that apply to the particular study. Efforts were made to minimize overlap among tables; however, in some cases, this was not always possible or even desirable. Certain items, such as reporting the flow of participants and participant attrition from studies, appear in multiple tables. This was done because the implications of reporting this information may vary over different kinds of studies (e.g., clinical trials vs. longitudinal studies). Figure 1 provides a flowchart that shows the decisions that are made in determining which tables, among Tables 1–6, apply for a particular study. All tables presume that Table 1 has been completed by the reporters of the research. The structure for reporting the flow of participants through each stage of an experiment or quasi-experiment can be found in the appendix of the *Publication Manual of the American Psychological Association* (6th ed.; APA, 2010).

The JARS–Quant tables do not specify where this information should be reported. The intent is for the information to be presented without compromising the readability of the paper. Information that is needed by the reader to understand the content of the report and evaluate the credibility of the results and conclusions should be immediately available to the reader (i.e., in the print version or online main text of the article). When possible, well-constructed tables can be used to present this information without disturbing the flow of the text. More detailed information that would be needed to allow replication of the empirical data collection or fine-grained understanding of the content of the article can be successfully provided in the supplemental materials provided by publishers. These supplemental materials however, should be ones freely open to all readers of the journal article, not just for subscribers.

Providing the information specified in the JARS–Quant tables is expected to become routine and minimally burdensome because these data are (or should be) regularly collected in the process of conducting empirical research; thus, JARS–Quant only represents guidelines for presenting the already-available data.

Table 1 remains the master table in the JARS–Quant hierarchy. All other tables involve detailed reporting expec-

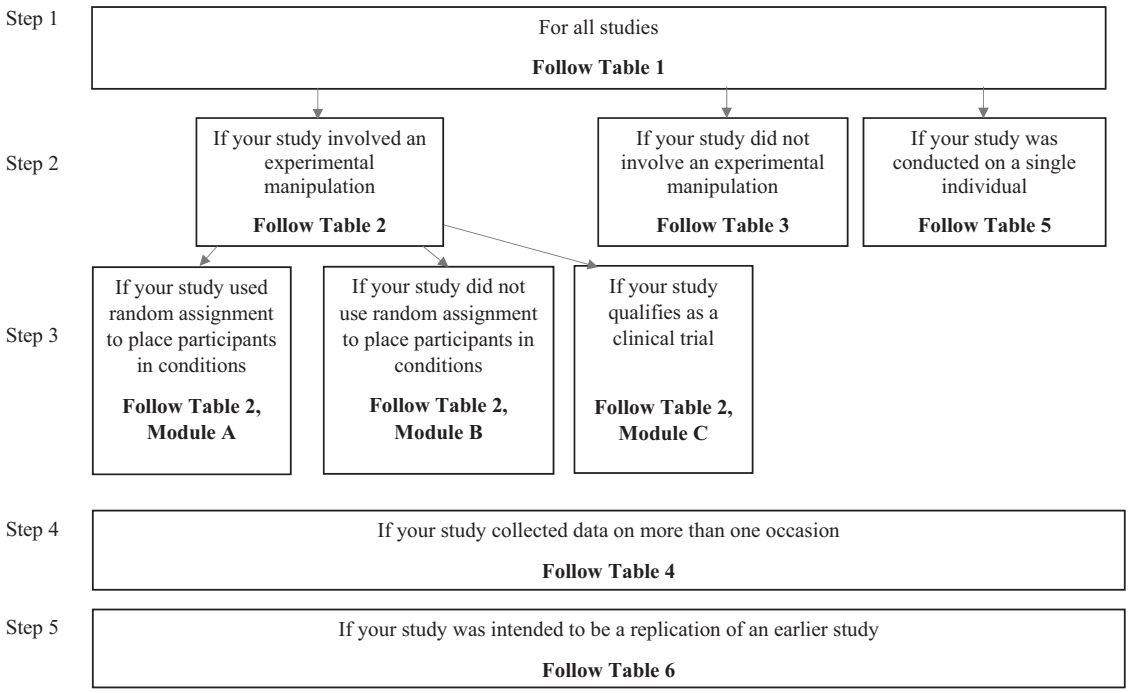

Figure 1. A flowchart describing the steps in choosing the JARS–Quant tables to complete depending on research design.

Table 1
Journal Article Reporting Standards (JARS): Information Recommended for Inclusion in Manuscripts That Report New Data Collections Regardless of Research Design

Paper section and topic	Description
Title and title page	
Title	• Identify main variables and theoretical issues under investigation, the relationships between them. Identify the populations studied.
Author note	• Provide, in the author note, acknowledgment and explanation of any special circumstances, including • Registration information if the study has been registered • Use of data also appearing in previous publications • Prior reporting of the fundamental data in dissertations or conference papers • Sources of funding or other support • Relationships or affiliations that may be perceived as conflicts of interest • Previous (or current affiliation of authors) if different from location where study was conducted • Contact information for the corresponding author • Additional information of importance to the reader that may not be appropriately included in other sections of the paper
Abstract	
Objectives	• State the problem under investigation. • Main hypotheses
Participants	• Describe subjects (animal research) or participants (human research), specifying their pertinent characteristics for this study; in animal research, include genus and species. Participants will be described in greater detail in the body of the paper.
Study method	• Describe the study method, including • Research design (e.g., experiment, observational study) • Sample size • Materials used (e.g., instruments, apparatus) • Outcome measures • Data-gathering procedures, including a brief description of the source of any secondary data. If the study is a secondary data analysis, so indicate.
Findings	• Report findings, including effect sizes and confidence intervals or statistical significance levels.
Conclusions	• State conclusions, beyond just results, and report the implications or applications.
Introduction	
Problem	• State the importance of the problem, including theoretical or practical implications.
Review of relevant scholarship	• Provide a succinct review of relevant scholarship, including • Relation to previous work • Differences between the current report and earlier reports if some aspects of this study have been reported on previously
Hypothesis, aims, and objectives	• State specific hypotheses, aims, and objectives, including • Theories or other means used to derive hypotheses • Primary and secondary hypotheses; other planned analyses • State how hypotheses and research design relate to one another.
Method	
Inclusion and exclusion	• Report inclusion and exclusion criteria, including any restrictions based on demographic characteristics.
Participant characteristics	• Report major demographic characteristics (e.g., age, sex, ethnicity, socioeconomic status) as well as important topic-specific characteristics (e.g., achievement level in studies of educational interventions). • In the case of animal research, report the genus, species, and strain number or other specific identification, such as the name and location of the supplier and the stock designation. Give the number of animals and the animals' sex, age, weight, physiological condition, genetic modification status, genotype, health–immune status; if known, drug- or test-naïve, and previous procedures to which the animal may have been subjected.
Sampling procedures	• Describe procedures for selecting participants, including • Sampling method if a systematic sampling plan was implemented • Percentage of sample approached that actually participated • Whether self-selection into the study occurred (either by individuals or by units, such as schools or clinics) • Settings and locations where data were collected as well as dates of data collection. • Agreements and payments made to participants • Institutional Review Board agreements, ethical standards met, and safety monitoring

Table 1 (*continued*)

Paper section and topic	Description
Sample size, power, and precision	• Describe the sample size, power, and precision, including • Intended sample size • Achieved sample size, if different from intended sample size • Determination of sample size, including • Power analysis, or methods used to determine precision of parameter estimates • Explanation of any interim analyses and stopping rules employed
Measures and covariates	• Define all primary and secondary measures and covariates, including measures collected but not included in this report.
Data collection	• Describe methods used to collect data.
Quality of measurements	• Describe methods used to enhance the quality of measurements, including • Training and reliability of data collectors • Use of multiple observations
Instrumentation	• Provide information on validated or ad hoc instruments created for individual studies, for example, psychometric and biometric properties.
Masking	• Report whether participants, those administering the experimental manipulations, and those assessing the outcomes were aware of condition assignments. • If masking took place, provide statement regarding how it was accomplished and if and how the success of masking was evaluated.
Psychometrics	• Estimate and report values of reliability coefficients for the scores analyzed (i.e., the researcher's sample), if possible. Provide estimates of convergent and discriminant validity where relevant. • Report estimates related to the reliability of measures, including • Interrater reliability for subjectively scored measures and ratings • Test–retest coefficients in longitudinal studies in which the retest interval corresponds to the measurement schedule used in the study • Internal consistency coefficients for composite scales in which these indices are appropriate for understanding the nature of the instruments being employed in the study • Report the basic demographic characteristics of other samples if reporting reliability or validity coefficients from those sample(s), such as those described in test manuals or in the norming information about the instrument.
Conditions and design	• State whether conditions were manipulated or naturally observed. Report the type of design as per the JARS–Quant tables: • Experimental manipulation with participants randomized • Table 2 and Module A • Experimental manipulation without randomization • Table 2 and Module B • Clinical trial with randomization • Table 2 and Modules A and C • Clinical trial without randomization • Table 2 and Modules B and C • Nonexperimental design (i.e., no experimental manipulation): observational design, epidemiological design, natural history, and so forth (single-group designs or multiple-group comparisons) • Table 3 • Longitudinal design • Table 4 • *N*-of-1 studies • Table 5 • Replications • Table 6 • Report the common name given to designs not currently covered in JARS–Quant.
Data diagnostics	• Describe planned data diagnostics, including • Criteria for post-data collection exclusion of participants, if any • Criteria for deciding when to infer missing data and methods used for imputation of missing data • Defining and processing of statistical outliers • Analyses of data distributions • Data transformations to be used, if any
Analytic strategy	• Describe the analytic strategy for inferential statistics and protection against experiment-wise error for • Primary hypotheses • Secondary hypotheses • Exploratory hypotheses
Results	
Participant flow	• Report the flow of participants, including • Total number of participants in each group at each stage of the study • Flow of participants through each stage of the study (include figure depicting flow when possible; see Figure 2)

(*table continues*)

Table 1 (*continued*)

Paper section and topic	Description
Recruitment	• Provide dates defining the periods of recruitment and repeated measures or follow-up.
Statistics and data analysis	• Provide information detailing the statistical and data-analytic methods employed, including • Missing data • Frequency or percentages of missing data • Empirical evidence and/or theoretical arguments for the causes of data that are missing, for example, missing completely at random (MCAR), missing at random (MAR), or missing not at random (MNAR) • Methods actually employed for addressing missing data, if any • Descriptions of each primary and secondary outcome, including the total sample and each subgroup that includes the number of cases, cell means, standard deviations, and other measures that characterize the data employed. • Inferential statistics, including • Results of all inferential tests conducted, including exact *p* values if null hypothesis statistical testing (NHST) methods were employed, including reporting the minimally sufficient set of statistics (e.g., *df*s, mean square [*MS*] effect, *MS* error) needed to construct the tests • Effect-size estimates and confidence intervals on those estimates that correspond to each inferential test conducted, when possible • Clear differentiation between primary hypotheses and their tests–estimates, secondary hypotheses and their tests–estimates, and exploratory hypotheses and their test–estimates • Complex data analyses, for example, structural equation modeling analyses (see also Table 8), hierarchical linear models, factor analysis, and multivariate analyses, and so forth, including • Details of the models estimated • Associated variance–covariance (or correlation) matrix or matrices • Identification of the statistical software used to run the analyses (e.g., SAS PROC GLM, or the particular R library program) • Estimation problems (e.g., failure to converge, bad solution spaces), regression diagnostics, or analytic anomalies that were detected and solutions to those problems. • Other data analyses performed, including adjusted analyses, if performed, indicating those that were planned and those that were not planned (though not necessarily in the level of detail of primary analyses). • Report any problems with statistical assumptions and/or data distributions that could affect the validity of findings.
Discussion	
Support of original hypotheses	• Provide a statement of support or nonsupport for all hypotheses whether primary or secondary, including • Distinction by primary and secondary hypotheses • Discussion of the implications of exploratory analyses in terms of both substantive findings and error rates that may be uncontrolled
Similarity of results	• Discuss similarities and differences between reported results and work of others.
Interpretation	• Provide an interpretation of the results, taking into account • Sources of potential bias and threats to internal and statistical validity • Imprecision of measurement protocols • Overall number of tests or overlap among tests • Adequacy of sample sizes and sampling validity
Generalizability	• Discuss generalizability (external validity) of the findings, taking into account • Target population (sampling validity) • Other contextual issues (setting, measurement, time; ecological validity)
Implications	• Discuss implications for future research, program, or policy.

Note. Tables have been designed to be comprehensive and to apply widely. For any individual report, the author would be expected to select those items that apply to the particular study.

tations for specific kinds of designs and for use with specific statistical approaches. The reporting standards for research synthesis and meta-analyses are self-contained in Table 9. In essence, Table 1 covers the basic features for reporting all forms of quantitative empirical studies. It is organized around the usual structure of a journal article found in the behavioral and psychological sciences literature. Whether individual items are reported in the text of the article or in archived supplemental materials depends on the flow of the article. Much of Table 1 is similar to that of the original JARS Table 1, but some important changes are noted next.

Changes in Table 1

The Method section of the JARS–Quant Table 1 contains subsections on Data Diagnostics and Analytic Strategy. These subsections were added for two reasons. First, they highlight the importance of including in reports descriptions of ways in which, if any, a data set has been modified prior to data analysis. These modifications could include, for example, the exclusion of data, imputation of missing data, identification and adjustment of statistical outliers, and the application of data transformations to alter the distribution of data points. These subsections are included in the Method

section of Table 1 because the criteria and methods used to make such modifications should be established prior to data analysis. If such modifications occur after data analysis has begun, this should be mentioned in the report and a clear rationale for the post hoc modifications should be provided. In addition, the unmodified data set should be retained and made available for verification purposes. Researchers making post hoc modifications should describe results based on both the modified and the unmodified data sets (however, one of the two sets of analyses could be included in supplemental materials).

Second, the subsections of Method in Table 1 emphasize the designation of hypotheses as of primary, secondary, and exploratory interest. These designations are meant to help convey to readers how experiment-wise results, of both null hypothesis significance tests and effect-size estimations, might be influenced by chance. This distinction in hypotheses is also reflected in changes to the JARS–Quant that appear in the subsections on Statistical and Data Analysis and Discussion in Table 1.

Animal research. References to animal research in Table 1 have been slightly modified to include the words *genetic modification status*, *genotype*, *health–immune status*, *drug-* or *test-naïve*, and *previous procedures* to make JARS–Quant consistent with the reporting standards included in ARRIVE (Kilkenny et al., 2010).

Psychometrics. When describing psychometric characteristics, authors should use language that is consistent with the recommendations of the most recent standards for educational and psychological testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014). Specifically, the term *reliability* should refer to test scores in particular samples and not to tests or testing instruments. Likewise, the term *validity* should refer not to a test but to the proposed interpretations of test scores (Reynolds & Livingston, 2012). That is, reliability and validity are not properties of tests that are invariant across all samples or proposed score interpretations (Thompson, 2003). Best practice is to estimate both reliability and validity, when possible, within the researcher's sample or samples (i.e., the scores analyzed). If the report includes values of reliability or validity coefficients from other published or unpublished sources, then these should be directly compared with the characteristics of the researcher's sample. Finally, the report should contain the appropriate types of score reliability coefficients, given the characteristics of the test, design, or analysis (see Slaney, Tkatchouk, Gabriel, & Maraun, 2009, for more information).

Neural measurement techniques. Although there are no changes in the content of Table 1 concerning neural measurement techniques, the variety of such techniques (including electroencephalogram [EEG], fMRI, magnetoencephalography [MEG], ERP, etc.) and their use have increased greatly in the past 2 decades. Alongside these developments in the techniques, sets of reporting standards for them have emerged. The JARS–Quant Working Group does not specifically endorse these independent standards but recognizes their value for reporting details of data collection and processing that are outside of the purview of this report.

When reporting the results of studies using this class of measurements, researchers must make full information about the technique accessible, including advanced data-processing information, either in the text of the paper or in the supplemental materials (usually the latter). These reporting expectations are in addition to those that would be expected in studies that do not employ this set of measurements. Other than the changes noted above, all of the other elements in the original Table 1 remain in full application, including the use of subject flow diagrams as illustrated in Figure 2. No items in the original Table 1 were eliminated.

Reporting Standards for Studies With an Experimental Manipulation

Listed in Table 2 are reporting standards for studies in which there is an experimental manipulation. Presented in Table 2 are additional modules that then further refine reporting standards for cases when assignment of participants to manipulation is done by a random process (Module A), when assignment is nonrandom (Module B), or when the study is a randomized clinical trial or a randomized control trial (Module C). Table 2 with Modules A and B were all present in the original JARS but have been slightly revised. Module C on reporting standards for clinical trials is new to JARS–Quant.

Clinical Trials

Listed in Table 2 Module C are the new additional reporting standards for formal clinical trials. There are two similar terms used to describe a wide class of studies with experimental manipulations: *randomized control trials* and *randomized clinical trials* (which are a subset of randomized control trials). In the literature, there is a tendency to use both terms interchangeably—however, a randomized clinical trial is a subset of the larger universe of randomized controlled trials. Module C includes reporting standards for clinical trials because certain investigations do require researchers to use these reporting standards. It would not be required for the reporting or registration of a more general randomized control trial, for example, in a university setting aimed at evaluating the efficacy of a new approach to teach calculus skills among college students; however, registration of studies with nonhealth outcomes may nevertheless be desirable.

In this context, a clinical trial is a research investigation that evaluates the effects of one or more health-related interventions (e.g., psychotherapy, medication, or a diet

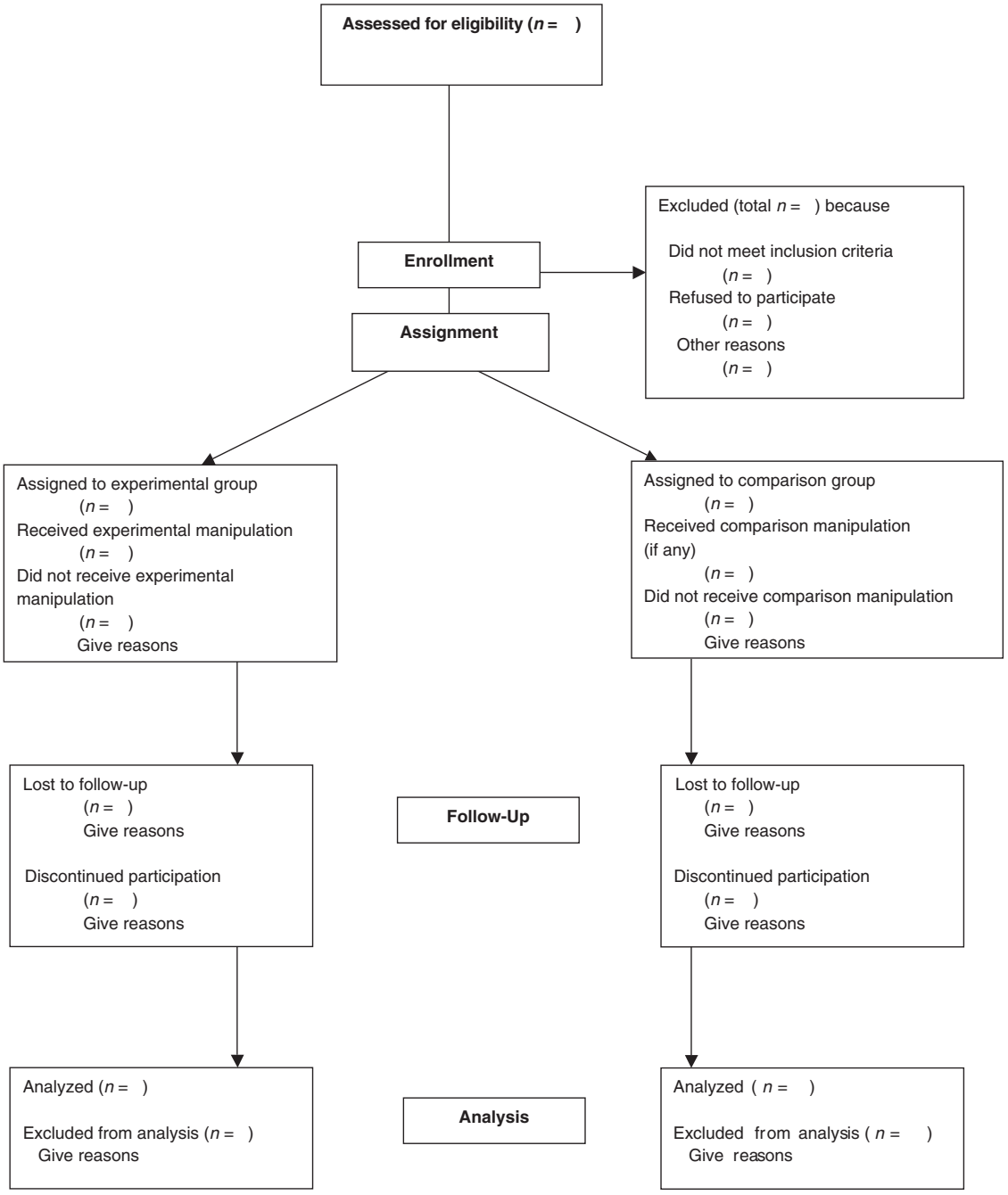

Figure 2. Flow of participants through each stage of an experiment or quasi-experiment. This flowchart is an adaptation of the flowchart offered by the CONSORT Group (Schulz, Altman, Moher, & the CONSORT Group, 2010).

intervention) on health outcomes (e.g., depression or diabetes) by prospectively assigning humans or groups of people to various experimental conditions. Assignment or allocation can be accomplished either randomly or nonrandomly (see Table 2, Modules A and B). Although the original JARS Table 1 included reporting standards that covered many aspects of a clinical trial, there are additional requirements for modern clinical trials. Thus, Module C was added to JARS–Quant.

One important difference concerns increased calls, both nationally and internationally, for *clinical trial registration*. This involves providing information to a registry about the

Table 2
Reporting Standards for Studies With an Experimental Manipulation (in Addition to Material Presented in Table 1)

Paper section and topic	Description
	General principles
Method	
Experimental manipulations	• Provide details of the experimental manipulation(s) intended for each study condition, including comparison conditions, and how and when experimental manipulations were actually administered, including • Content of the specific experimental manipulations (if experimental manipulation is part of a clinical trial, address Module C) • Summary or paraphrasing of instructions, unless they are unusual or compose the experimental manipulation, in which case they may be presented verbatim • Method of experimental manipulation delivery • Description of apparatus and materials used and their function in the experiment • Specialized equipment by model and supplier • Deliverer: who delivered the experimental manipulations • Level of professional training • Level of training in specific experimental manipulations • Number of deliverers, and in the case of experimental manipulations, the *M*, *SD*, and range of number of individuals–units treated by each • Setting: where the manipulations or experimental manipulations occurred • Exposure quantity and duration: how many sessions, episodes, or events were intended to be delivered and how long they were intended to last • Time span: how long it took to deliver the experimental manipulation to each unit • Activities to increase compliance or adherence (e.g., incentives) • Use of language other than English and the translation method • Sufficient detail to allow for replication, including reference to or a copy of the manual of procedures. If the manual of procedures is available, and how others may obtain it
Units of delivery and analysis	• State the unit of delivery (how participants were grouped during delivery). • Describe the smallest unit that was analyzed (and in the case of experiments, that was randomly assigned to conditions) to assess experimental manipulation effects (e.g., individuals, work groups, classes). • Describe the analytical method used to account for this (e.g., adjusting the standard error estimates by the design effect or using multilevel analysis) if the unit of analysis differed from the unit of deliver.
Results	
Participant flow	• Report the total number of groups (if experimental manipulation was administered at the group level) and the number of participants assigned to each group, including • Number of participants approached for inclusion • Number of participants who began the experiment • Number of participants who did not complete the experiment or crossed over to other conditions, with reasons • Number of participants included in primary analyses • Include a figure describing the flow of participants through each stage of the study (see Figure 2).
Treatment fidelity	• Provide evidence on whether the experimental manipulation was implemented as intended.
Baseline data	• Describe baseline demographic and clinical characteristics of each group.
Adverse events and side effects	• Report all important adverse events or side effects in each experimental condition. If none, state so.
Discussion	• Discuss results, taking into account the mechanism by which the experimental manipulation was intended to work (causal pathways) or alternative mechanisms. • Discuss the success of, and barriers to, implementing the experimental manipulation; fidelity of implementation if an experimental manipulation is involved. • Discuss generalizability (external validity and construct validity) of the findings, taking into account • Characteristics of the experimental manipulation • How, what outcomes were measured • Length of follow-up • Incentives • Compliance rates • Describe the theoretical or practical significance of outcomes and the basis for these interpretations.
	Module A: Reporting standards for studies using random assignment
Method	
Random assignment method	• Describe the unit of randomization and the procedure used to generate the random assignment sequence, including details of any restriction (e.g., blocking, stratification).
Random assignment implementation and concealment	• State whether and how the sequence was concealed until experimental manipulations were assigned, including who • Generated the assignment sequence • Enrolled participants

(*table continues*)

Table 2 (*continued*)

Paper section and topic	Description
	• Assigned participants to groups
Masking	• Report whether participants, those administering the experimental manipulations, and those assessing the outcomes were aware of condition assignments. • Provide a statement regarding how any masking (if it took place) was accomplished and whether and how the success of masking was evaluated.
Statistical methods	• Describe statistical methods used to compare groups on primary outcome(s). • Describe statistical methods used for additional analyses, such as subgroup comparisons and adjusted analysis. • Describe statistical methods used for mediation or moderation analyses, if conducted.
Module B: Reporting standards for studies using nonrandom assignment	
Method	
Assignment method	• Report the unit of assignment (i.e., the unit being assigned to study conditions; e.g., individual, group, community). • Describe the method used to assign units to study conditions, including details of any restriction (e.g., blocking, stratification, minimization). • State procedures employed to help minimize selection bias (e.g., matching, propensity score matching).
Masking	• Report whether participants, those administering the experimental manipulation, and those assessing the outcomes were aware of condition assignments. • Report whether masking took place. Provide a statement regarding how it was accomplished and how the success of masking was evaluated, if it was evaluated.
Statistical methods	• Describe statistical methods used to compare study groups on primary outcome(s), including complex methods for correlated data. • Describe statistical methods used for any additional analyses conducted, such as subgroup analyses and adjusted analysis (e.g., methods for modeling pretest differences and adjusting for them). • Describe statistical methods used for mediation or moderation analyses, if these analyses were used.
Module C: Reporting standards for studies involving clinical trials	
Title and title page	• State whether trial was registered prior to implementation.
Abstract	• State whether the trial was registered. If the trial was registered, state where and include the registration number. • Describe public health implications of trial results.
Introduction	• State the rationale for evaluating specific intervention(s) for a given clinical problem, disorder, or variable. • Describe the approach, if any, to assess mediators and moderators of treatment effects. • Describe potential public health implications of study. • State how results from current study can advance knowledge in this area.
Method	
Participant characteristics	• State the method(s) of ascertaining how participants met all inclusion and exclusion criteria, especially if assessing clinical diagnosis(es).
Sampling procedures	• Provide details regarding similarities and differences of data collection locations if multisite study.

Measures	• State whether clinical assessors were
	  • Involved in providing treatment for studies involving clinical assessments
	  • Aware or unaware of assignment to condition at post-treatment and follow-up assessment(s); (if unaware, how was this accomplished?)

Experimental interventions	• Report whether the study protocol was publicly available (e.g., published) prior to enrolling participants; if so, where and when. • Describe how intervention in this study differed from the "standard" approach in order to tailor it to a new population (e.g., differing age, ethnicity, comorbidity). • Describe any materials (e.g., clinical handouts, data recorders) provided to participants and how information about them can obtained (e.g., URL address). • Describe any changes to the protocol during the course of the study, including all changes to the intervention, outcomes, and methods of analysis. • Describe the Data and Safety Monitoring Board. • Describe any stopping rules.
Treatment fidelity	• Describe method and results regarding treatment deliverers' (e.g., therapists) adherence to the planned intervention protocol (e.g., therapy manual). • Describe method and results of treatment deliverers' (e.g., therapists) competence in implementing the planned intervention protocol (e.g., therapy manual). • Describe (if relevant) method and results regarding whether participants (i.e., treatment recipients) understood and/or followed treatment recommendations (e.g., did they comprehend what the treatment was intended to do, complete homework assignments if given, and/or practice activities assigned outside of the treatment setting?).

Table 2 (*continued*)

Paper section and topic	Description
	• Describe any additional methods used to enhance treatment fidelity.
Research design	• Provide rationale for length of follow-up assessment.
Results	• Describe how treatment fidelity (i.e., therapist adherence and competence ratings) and participant adherence was related to intervention outcome. • Describe method of assessing clinical significance, including if the threshold for clinical significance was prespecified (e.g., as part of a publicly available protocol). • Identify possible differences in treatment effects due to intervention deliverer. • Describe possible differences in treatment effects due to data collection site if multisite study. • Describe results of analyses of moderation–mediation effects, if tested. • Explain why study was discontinued, if appropriate. • Describe frequency and type of adverse effects that occurred (or state that none occurred).
Discussion	• Describe how this study advances knowledge about the intervention, clinical problem, and/or population.

study prior to its implementation as well as a summary of results upon its completion. Trial registration can enhance transparency by providing a complete description of the trial to both the scientific community and the general public. From an ethical perspective, the Declaration of Helsinki, which is the set of ethical principles regarding human experimentation developed by the World Medical Association (2013), stated that "every clinical trial must be registered in a publicly accessible database before recruitment of the first subject" (p. 2193). Trial registration also helps minimize publication bias and selective reporting of results. As of January 18, 2017, all clinical trials, funded in whole or in part by the National Institutes of Health (NIH), must be registered in ClinicalTrials.gov. A *clinical trial* is defined by NIH as a "research study in which one or more human subjects are prospectively assigned to one or more interventions (which may include placebo or other control) to evaluate the effects of those interventions on health-related biomedical or behavioral outcomes" (NIH, 2014, para. 4). Relevant to the majority of clinical trials conducted by psychologists, this definition includes various types of psychotherapy and psychosocial interventions (e.g., cognitive therapy, diet, exercise, problem-solving training) as well as delivery systems (e.g., telemedicine, face-to-face interviews). Additional information can be found in FAQs on the NIH website (http://www.grants.nih.gov/clinicaltrials_fdaaa/faq.htm#5052).

On an international basis, the World Health Organization (WHO) manages the International Clinical Trials Registry Platform (http://www.who.int/ictrp/trial_reg/en/), which provides a way to search ClinicalTrials.gov and other registries. Information about where a trial is registered should be reported on the title page, in the abstract, and in the reporting of the experimental manipulation. A second issue involves the difference in the amount of information necessary to adequately describe the experimental manipulation implemented in a clinical trial. This can include details regarding one or more psychotherapy treatment conditions as well as any comparators and control conditions.

In addition, because of the potential variability in performance among both assessors or data gatherers of clinical information (e.g., those conducting complex clinical interviews) and psychotherapists or interventionists, more details are usually requested. One issue involves taking steps to monitor how the intervention was delivered. This is often referred to as *treatment integrity* or *fidelity*, and includes the degree to which the planned intervention (e.g., as described in a treatment manual) was delivered by a therapist (e.g., did the individuals implementing the experimental manipulation follow the protocol?) and taken-up by participants (e.g., did the clients attend all sessions?; see Borrelli, 2011; Montgomery, Underhill, Gardner, Operario, & Mayo-Wilson, 2013; Nezu & Nezu, 2008). This information would be reported in the Results section.

Of particular importance, highlighted in this new module is the need to report mild to severe adverse events, or occurrences more likely to happen when evaluating interventions meant to affect health outcomes compared with other types of research investigations. Recent research has indicated that few behavioral health intervention studies monitor and report adverse events other than serious occurrences, such as suicide or hospitalization (see Peterson, Roache, Raj, & Young-McCaughan, 2012; for the STRONG STAR Consortium). Increased distress symptomatology or the negative effects of treatment on others are rarely reported (Duggan, Parry, McMurran, Davidson, & Dennis, 2014). Without such information, patients are unable to ascertain the full array of possible risks or benefits of psychological interventions, clinicians are unable to determine the valence and direction of a benefit–harm analysis, and the ability of policymakers and professional organizations to develop valid clinical practice guidelines is severely hampered. Such information, even including a statement that no adverse effects occurred, would be reported in the Results section.

Nonexperimental Research

Table 3 is new to JARS–Quant and deals with reporting standards for studies in which no variables are manipulated. Instead, the main goal of such studies is to observe, describe, classify, or analyze the naturally occurring relations between variables of interest. These studies are sometimes

Table 3
Reporting Standards for Studies Using No Experimental Manipulation (Single-Group Designs, Natural-Group Comparisons, etc.; in Addition to Material Presented in Table 1)

Paper section and topic	Description
Title/Abstract	
Study design	• Describe the design of the study.
Data use	• State the type of data used.
Method	
Participant selection	• Describe the method(s) of selecting participants (i.e., the units to be observed, classified, etc.), including • Method(s) of selecting participants for each group (e.g., methods of sampling, place of recruitment) and the number of cases in each group • Matching criteria (e.g., propensity score), if matching was used • Identify data sources used (e.g., sources of observations, archival records), and if relevant, include codes or algorithms to select participants or link records.
Variables	• Define all variables clearly, including • Exposure • Potential predictors, confounders, and effect modifiers • State how each variable was measured.
Comparability of assessment	• Describe comparability of assessment across groups (e.g., the likelihood of observing or recording an outcome in each group for reasons unrelated to the effect of the intervention).
Analysis	• Describe how predictors, confounders, and effect modifiers were included in the analysis.
Discussion	
Limitations	• Describe potential limitations of the study. As relevant, describe the possibility of misclassification, unmeasured confounding, and changing eligibility criteria over time.

called *observational*, *correlational*, or *natural history studies*. Given the nature of the research question, such studies may have different designs or sampling plans (e.g., prospective, retrospective, case-control, cohort, cohort-sequential). They include single-group studies, in which relations among attributes in a naturally occurring group are analyzed as well as studies in which comparisons are made across two or more naturally occurring groups on variables of interest. Reporting guidelines in Table 3 were informed, in part, by the STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) reporting standards (http://www.strobe-statement.org/index.php). As with other tables, Table 3 is intended to be used along with Table 1.

Longitudinal Research

In almost all cases, a *longitudinal study* (Table 4) employs one of the basic study designs but the same

Table 4
Reporting Standards for Longitudinal Studies (in Addition to Material Presented in Table 1)

Paper section and topic	Description
General reporting expectation	
Sample characteristics (when appropriate)	• Describe reporting (sampling or randomization) unit—individual, dyad, family, classroom: • *N* per group, age, and sex distribution • Ethnic composition • Socioeconomic status, home language, immigrant status, education level, and family characteristics • Country, region, city, and geographic characteristics
Sample recruitment and retention methods	
Attrition	• Report attrition at each wave, breaking down reasons for attrition. • Report any differential attrition by major sociodemographic and experimental condition.
Additional sample description	• Report any contextual changes for participants (units) as the study progressed (school closures–mergers, major economic changes; for long-term studies, major social changes that may need explanation for contemporary readers to understand the context of the study during its early years).
Method and measurement	• Specify independent variables and dependent variables at each wave of data collection. • Report the years in which each wave of the data collection occurred.
Missing data	• Report the amount of missing data and how issues of missing data where handled analytically.
Analysis	• Specify analytic approaches utilized and assumptions made in performing these analyses.
Multiple publication	• Provide information on where any portions of the data have been previously published and the degree of overlap with current report.

experimental unit or units are observed on the same response variables on more than one occasion. In these studies, the objective is usually to understand the occasion-changes either in and of themselves or as functions of other influences. As used here, longitudinal designs are distinct from three other similar designs. Rather than the passage of time (or some other metric), *occasion* might also refer to experimental manipulations such as dosage level, experimental condition, and so forth, and may occur within a single session or in different sessions. The latter kinds of designs are often called *within-subject designs* and generally have different reporting standards than for what is being referred to here as longitudinal studies. Although all longitudinal studies are within-subject, not all within-subject designs are longitudinal. Also, there are designs in which the same experimental units are measured on several different dependent variables at various points during a single session, but no attribute is measured on more than one occasion. These are multivariate outcome studies, and reporting standards for longitudinal studies do not generally apply to such studies. Finally, there are time-series experiments that generally have their own reporting standards.

Longitudinal studies come in many different shapes and forms and are traditionally, but not uniquely, seen in developmental, geriatric, and educational studies as well as in clinical trials. These studies may typically involve some preintervention measures, an intervention, and then one or more postintervention measures. Other longitudinal studies may involve a selection of cases at a particular time or event and then repeated observations of these participants on a prespecified schedule (which may or may not be the actual achieved schedule). The prespecified schedule may be time or event based.

Some work has been published on reporting standards for longitudinal studies (Tooth, Ware, Bain, Purdie, & Dobson, 2005), but most of these are designed to report studies that arise from epidemiology. To develop reporting standards that would reflect longitudinal studies in the behavioral sciences, several organizations were consulted, including the Governing Council of the Society for Research in Child Development. The Working Group received assistance from the Governing Council of the Society for Research in Child Development in creating standards for reporting longitudinal studies.

As with other reporting standards, those for longitudinal studies can be divided into two general classes: (a) those items that are required for a well-trained reader to be able to make decisions concerning the validity and scope of application of the findings as the article is being read, and (b) those details that might be necessary for a fine-grained understanding of the work and its possible replication. Information of the first type should appear in the body of the paper, whereas the more fine-grained information can be made available in the supplemental materials. Because most longitudinal studies are, at any one measuring instance, a traditional experimental or observational design, the reporting standards for that class of design also would be expected to be followed in the report. In addition, any materials that pertain to the entire study (e.g., a study registration number) would be expected to appear as it would in a nonlongitudinal study.

It is expected that when important conditions change from observational period to observational period (e.g., a test form is changed, a new measure that is purportedly measuring a similar or the same construct is substituted for an earlier measure, when major life events occur such as family structure changes for some participants), these will be clearly noted. In such cases, the dates of data collection (e.g., years 2010–2015) should be given.

N-of-1

A class of studies (see Table 5) known as *N-of-1 experimental designs* (also known as *single-case studies*) are commonly employed in many areas of behavioral and educational research, but there has been limited specification of reporting standards for this class of design. In addition, there is ample evidence of incomplete reporting in single-case intervention research (e.g., Barker, Mellalieu, McCarthy, Jones, & Moran, 2013; Didden, Korzilius, van Oorsouw, & Sturmey, 2006; Maggin, Chafouleas, Goddard, & Johnson, 2011; Smith, 2012; Tate, Perdices, McDonald, Togher, & Rosenkoetter, 2014).

Although reporting guidelines for *N*-of-1 trials currently exist (CONSORT Extension for *N*-of-1 Trials [CENT]; Shamseer et al., 2015; Vohra et al., 2015), no reporting guidelines have been available specifically for behavioral science research. The SCRIBE 2016 guideline (Single-Case Reporting Guideline In BEhavioural interventions) was developed to address this need (Tate, Perdices, Rosenkoetter, McDonald, et al., 2016; Tate, Perdices, Rosenkoetter, Shadish, et al., 2016). The SCRIBE 2016 guideline provides researchers who conduct single-case experiments with a minimum standard, in the form of a 26-item checklist, with which they can write their reports clearly and accurately. The SCRIBE 2016 guideline is intended to be used in the four prototypical designs often used in single-case experiments: withdrawal–reversal, multiple baseline, alternating–simultaneous treatments, and changing criterion as well as combinations and variants of the designs. Two primary articles on the SCRIBE 2016 are available: (a) a SCRIBE Statement (Tate, Perdices, Rosenkoetter, Shadish, et al., 2016) that describes the methodology of their development, and (b) an explanation and elaboration article (Tate, Perdices, Rosenkoetter, McDonald, et al., 2016) that provides a

Table 5
Reporting Standards for N-of-1 Studies (in Addition to Material Presented in Table 1)

Paper section and topic	Description
Design	• Describe the design, including • Design type (e.g., withdrawal-reversal, multiple-baseline, alternating-treatments, changing-criterion, some combination thereof, or adaptive design) • Phases and phase sequence (whether determined *a priori* or data-driven) and, if applicable, criteria for phase change
Type of design	
Procedural changes	• Describe any procedural changes that occurred during the course of the investigation after the start of the study.
Replication	• Describe any planned replication.
Randomization	• State whether randomization was used, and if so, describe the randomization method and the elements of the study that were randomized (e.g., during which phases treatment and control conditions were instituted).
Analysis	
Sequence completed	• Report for each participant the sequence actually completed, including the number of trials for each session for each case. • State when participant(s) who did not complete the sequence stopped and the reason for stopping.
Outcomes and estimation	• Report results for each participant, including raw data for each target behavior and other outcomes.

rationale for each of the 26 items, along with examples from the literature of adequate reporting of the items.

The SCRIBE 2016 checklist will provide researchers, authors, reviewers, and editors involved in the publication of results of single-case trials with a tool to measure the clarity and accuracy of reporting. The checklist is also expected to facilitate the replication of these studies. In Table 5, a portion of the SCRIBE 2016 formulation is summarized. In this table, which is intended to be used in conjunction with Table 1 and the Table 2 modules, those elements of SCRIBE 2016 that are unique to the single-case design are summarized.

Studies Reporting Replications

Reproducibility is a core scientific principle. Increasingly, there have been efforts to make replication more likely, including the adoption of reporting standards such as these, policy changes in journals to include publication of replication studies as part of their primary mission, and so forth (Begley & Ellis, 2012; Nosek & Lakens, 2013; Open Science Collaboration, 2015). At the same time, there is a necessity to ensure that the replication studies themselves are reported in such a way that readers can easily understand what was done and how to evaluate the claims made in those replication studies (See Table 6). These standards concern *external replication*, which occurs when researchers state that a study being reported is a repetition of one or more specific, previously published, or archived studies. They do not apply to *internal replication*, which involves cross-validation of analyses within the same sample or the use of resampling or randomization procedures, such as *bootstrapping*, that recombine or generate cases to estimate the statistical precision of specific estimators.

Reporting should highlight the comparisons between the original and replication studies such that sufficient detail is provided to permit evaluation of whether any differences in outcomes between the original and the replication study are due to differences such as in participants, conditions, measures, methods of analysis, or other factors induced into the replication study. More information about terms and concepts mentioned in these standards is available in Mackey (2012) and Maxwell, Lau, and Howard (2015).

Reporting Standards for Some Quantitative Procedures

Although reporting standards are generally associated with entire research designs, some quantitative procedures are of sufficient complexity and open to such internal variation that additional information (beyond just the name of the technique and a few parameters) needs to be reported for the reader to be able to fully comprehend the analysis. Researchers may need additional information to evaluate the conclusions that the authors have drawn or replicate the analysis with their own data.

Structural Equation Modeling

Structural equation modeling (SEM) is a family of statistical techniques that involve the specification of a structural or measurement model, given relevant theory and previous empirical results. These statistical techniques will include a series of analytic steps that estimate effects represented in the model (parameters) and evaluate the extent of correspondence between the model and data. Hoyle and Isherwood (2013) developed standards for studies in which results of SEM analyses are reported (see Table 7). These standards take the form of a comprehensive description of data preparation, specification of the initial model(s), estimation, model fit assessment, respecification of the model(s), and the reporting of the results. Hoyle

Table 6
Reporting Standards for Replication Studies (in Addition to Material Presented in Table 1)

Paper section and topic	Description
Study type	• Report sufficient information both in the study title and, more important, in the text that allows the reader to determine whether the study is a direct (exact, literal) replication, approximate replication, or conceptual (construct) replication. • Indicate whether a replication study has conditions, materials, or procedures that were not part of the original study. • Describe these new features, where in the study they occur, and their potential impact on the results. • Report for both the original study and the replication study indications of treatment fidelity.
Participants	• Compare the recruitment procedures in the original and replication studies. Note and explain major variations in how the participants were selected, such as whether the replication study was conducted in a different setting (e.g., country or culture) or whether the allocation of participants to groups or conditions is different. Describe implications of these variations on the results. • Compare the demographic characteristics of the participants in both studies. If the units of analysis are not people (cases), such as classrooms, then report the appropriate descriptors of their characteristics.
Instrumentation	• Report instrumentation that includes both hardware (apparatus) and "soft" measures used to collect data, including questionnaires, structured interviews, or psychological tests. Clarify in appropriate subsections of the Method section any major differences between the original and replication studies. • Indicate whether questionnaires or psychological tests were translated to another language, and specify the method(s) used, such as back-translation, to verify that the translation was accurate. • Report psychometric characteristics of the scores analyzed in the replication study and compare these properties with those in the original study. • Specify and compare the informant(s) and method(s) of administration across the two studies. The latter includes the setting for testing, such as individual versus group administration, and the method of administration, such as paper-and-pencil versus online.
Analysis	• Report results of the same analytical methods (statistical or other quantitative manipulations) used. Results from additional or different analyses may also be reported. State the statistical criteria for deciding whether the original results were replicated in the new study. Examples of criteria include statistical significance testing, effect sizes, confidence intervals, and Bayes factors in Bayesian methods. Explain decision rules when multiple criteria, such as significance testing with effect size estimation, are employed. State whether the effect size in a power analysis was specified to equal that reported in the original study (conditional power) or whether power was averaged over plausible values of effect size based on an estimated standard error (predictive power), which takes account of sampling error.

and Isherwood's questionnaire was adapted with permission for inclusion in these revised JARS–Quant reporting standards.

The standards for SEM studies outlined next are organized by the sections of the manuscript. Those for the title, abstract, introduction, Method, and Discussion sections elaborate on certain points from Table 1 by adding information more specific to SEM. Standards for the Results section of manuscripts in which SEM results are reported concern estimation, evaluation of model fit, and the reporting of statistical findings. These standards call on authors to state their justification for choices made when alternative statistical methods (e.g., maximum likelihood vs. a different estimation method) or model testing strategies (e.g., trimming vs. building) are available. For respecification, authors should disclose the theoretical or statistical bases for modifying an initial model. For more information about best practices in SEM, see Kline (2016, Chapter 18), Mueller and Hancock (2008), and Schumaker and Lomax (2016, Chapter 18).

Bayesian Statistics

Bayesian statistical analysis has become a more commonly used statistical procedure in behavioral research. Relatively little has been published to guide authors regarding what information to report when using this class of analysis. To that end, the JARS–Quant Working Group invited David Rindskopf to develop a set of reporting standards for use with Bayesian analysis. These standards are summarized in Table 8.

Meta-Analysis Reporting Standards

Revisions to the meta-analysis reporting standards (MARS) were developed in three steps. First, two recent revisions to other reporting standards for research syntheses and meta-analysis in the health professions were examined (Montgomery, Underhill, et al., 2013; Stroup et al., 2000), as was a published recommendation regarding the reporting of literature search strategies (Atkinson, Koenka, Sanchez, Moshontz, & Cooper, 2015). Second, items not represented on the original MARS were added to the revised MARS after terminology was changed to reflect that used in the social sciences. Third, the members of the Society for Research Synthesis Methodology were asked to examine the original MARS and to suggest any changes for the revision. Two members made suggestion that were incorporated into the revision. Finally, the revision was vetted with the JARS–Quant Working Group.

Table 7
Reporting Standards for Studies Using Structural Equation Modeling

Paper section and topic	Description
Title	• Mention the basic mechanism or process reflected in the primary model to which the data are fit. (Note: The complexity of the multivariate data analyzed in many structural equation modeling (SEM) studies makes it unlikely that, in most cases, the variables under investigation and the relations between them could be concisely stated in the title.)
Abstract	• Report values for at least two global fit statistics, each from a different class, and include a brief statement about local fit (residuals). State whether the interpreted model (if any model is retained) is the originally specified model.
Introduction	• Describe the primary model to be fitted to the data, and include an explanation of theory or results from previous empirical studies that support the primary model. • Point out paths that are especially important, and justify directionality assumptions, such as the claim that *X* causes *Y* instead of the reverse. Do the same for paths of secondary importance. • State whether respecification is planned, if the primary model is rejected.

Method	• State whether the data were collected from research participants or generated by computer simulation.
	• Report whether indicators of latent variables were drawn from one questionnaire or from multiple questionnaires.
	• Describe, for each questionnaire, whether the indicators are items or total scores across homogeneous sets of items (scales, parcels), stating how
	  • Scales were constructed, reporting their psychometrics
	  • Items were treated in the analysis as continuous or categorical
	• Report how the target sample size was determined, including
	  • Rule of thumb
	  • Availability of resource constraints
	  • Results of a priori power analysis
	  • Estimates of parameter precision used to plan the number of cases with appropriate explanation
	• For a power analysis, state
	  • Target level of power
	  • Null and alternative hypotheses
	  • Significance of key parameters
	  • Fit statistics that figured in the analysis
	  • Expected population effect sizes
	• Report the computer software or algorithm used if the data were generated by simulation, state and justify the sizes of generated samples, and disclose whether samples were lost because of nonconvergence or inadmissible estimates.

Results	• Report data diagnostics, including
	  • Percentage of missingness (if some data are missing) and how it is distributed across cases and variables
	  • Empirical evidence or theoretical arguments about causes of missing data (i.e., missing completely at random [MCAR], missing at random [MAR], or missing not at random [MNAR])
	  • Evidence that distributional or other assumptions of estimation methods are plausible

Missing data	• Indicate the statistical method used to address missingness, such as multiple imputation, full information maximum likelihood (FIML), substitution of values, or deletion of cases. For multiple imputation or FIML estimates, state whether variables not included in the model were specified as auxiliary variables.

Distributions	• State whether the data were evaluated for estimation methods that assume multivariate normality.
	  • Report values of statistics that measure univariate or multivariate skewness and kurtosis that support the assumption of normal distributions.
	  • If the data were not multivariate normal, state the strategy used to address nonnormality, such as use of a different estimation method that does not assume normality or use of normalizing transformations of the scores.

Data summary	• Report in the manuscript—or make available in the supplemental materials—sufficient summary statistics that allow secondary analysis, including
	  • Covariance matrix with means, or a correlation matrix with standard deviations and means for continuous variables
	  • Polychoric correlation matrix, items thresholds, and the asymptotic covariance matrix for categorical variables
	• Indicate whether the case-level data are archived, and provide information about how these data can be accessed by interested readers.

Specification	• Indicate the general approach that best describes the application of SEM, strictly confirmatory, comparison of alternative models, or model generation. • Provide the diagram for each model fitted to the data. If the diagram would be overly complex, such as when large numbers of variables are analyzed, then clearly describe the models in text. A reader should be able to translate the text description of a model into a diagram. • Give a full account of the specification for all models to be evaluated, including observed variables, latent variables, fixed or free parameters, and constrained parameters. • Report sufficient information, such as tabulations of the numbers of observations versus free parameters, so that the model degrees of freedom can be derived by the reader. • Verify that models to be analyzed are actually identified. State the basis for this claim, including the method, rules, or heuristics used to establish identification. • State the basis in theory or results of previous empirical studies if a measurement model is part of a larger model.

Table 7 (*continued*)

Paper section and topic	Description
	• Describe fully the specification of the mean structure if the model has a means component. • Explain the rationale for including error correlations in the model if correlated error terms are specified. • Explain how the effects are specified if the model includes interaction effects. • Explain how nonindependence is accounted for in the model for nested data (e.g., occasions within persons, students within classrooms). • Describe any comparisons of parameters to be made between groups or occasions, and indicate which parameters are to be compared if models are fitted to data from multiple groups or occasions.
Estimation	• State the software (including version) used in the analysis. Also state the estimation method used and justify its use (i.e., whether its assumptions are supported by the data). • Disclose any default criteria in the software, such as the maximum number of iterations or level of tolerance, that were adjusted in order to achieve a converged and admissible solution. • Report any evidence of an inadmissible solution (e.g., error variances less than zero or constrained by the computer at zero; estimated absolute correlations or proportions of explained variance that exceed 1.0). Explain what was done to deal with the problem.
Model fit	• Report fit statistics or indices about global (omnibus) fit interpreted using criteria justified by citation of most recent evidence-based recommendations for all models to be interpreted. • Report information about local fit, such as covariance, standardized, normalized, or correlation residuals, that justify retaining the model at the level of pairs of observed variables for all interpreted models. • State the strategy or criteria used to select one model over another if alternative models were compared. Report results of difference tests for comparisons between alternative models. • State the test and criterion for testing estimates of individual parameters. If parameter estimated were compared over groups or occasions, indicate how those comparisons were made.
Respecification	• Indicate whether one or more interpreted models was a product of respecification. If so, then describe the method used to search for misspecified parameters. • State which parameters were fixed or freed to produce the interpreted model. Also provide a theoretical or conceptual rationale for parameters that were fixed or freed after specification searching. • Indicate whether models for which results are presented were specified before or after fitting other models or otherwise examining the data.
Estimates	• Report both unstandardized and standardized estimates for all estimated parameters. • Report the corresponding standard errors, especially if outcomes of significance testing for individual parameters are reported. State the cutoffs for levels of statistical significance, if such cutoffs were used. • Report estimates of indirect effects, both unstandardized and standardized. Also report values of standard errors for indirect effects, if possible. State and justify the strategy for testing indirect effects. • Report estimates of interaction effects and also results of follow-up analyses that clarify the underlying pattern for interpreted interactions. Also report values of standard errors for such interactions.
Discussion	• Summarize the modifications to the original model and the bases, theoretical or statistical, for doing so. • Address the issue of equivalent models that fit the same data as well as retained models or alternative-but-nonequivalent models that explain the data nearly as well as retained models. Justify the preference for retained models over equivalent or near-equivalent versions.

The revised MARS (see Table 9) includes six groups of changes. First, the wording has been revised to clarify that many sections of the MARS should be completed by research synthesists whether or not the evidence in their report is amenable to conducting a meta-analysis. Second, additional detail has been added to the description of the title page and author note. Most important, authors are now asked to (a) explicitly state whether a possible or perceived conflict of interest may exist, and (b) provide the name and document entry number if the work had been placed in a research register prior to being conducted (e.g., University of York, n.d.). Third, in recognition of the growing number of distinct research questions with unique techniques to which research syntheses are now being applied, both the abstract and the introduction sections now ask that authors specify the type of synthesis being conducted. Fourth, the Search Strategy section has been expanded to deal directly with the five types of strategies used most frequently in literature searches. Details of these searches are now asked for to allow more precise replication of searches and the evaluation of whether biases might exist in the retrieved literature. Fifth, the section on Coding Procedures has been expanded to include more information on code development. The section has also been changed with regard to reporting of coder reliability. Sixth, a subsection has been added to the Statistical Methods section asking for the reporting of recently developed statistical outcomes of meta-analysis.

Some Final Thoughts

The JARS–Quant formulations, like those of the original JARS, were developed specifically for use in the social, behavioral, and educational sciences. They incorporate numerous ideas from other reporting standards, of which there

Table 8
Reporting Standards for Studies Using Bayesian Techniques

Paper section and topic	Description
Model	• Completely specify both the systematic and the stochastic parts of the analyzed model, and give the rationale for choices of functional forms and distributions.
Distributions	• Describe the prior distribution(s) for model parameters of interest. If the priors are informative, state the rationale for that choice, and conduct a sensitivity analysis to check the dependence of the results on the prior distribution. • Describe the posterior distribution(s) for substantive model parameters and important functions of the parameters. If feasible, report the highest posterior density (HPD) interval for each parameter or function. • Plot or describe the joint distribution if substantive parameters are correlated. • If predictions are made for observable quantities, make available either the actual predictive distribution and parameter estimates, report summary statistics that describe the distribution, or a graphical summary.
Likelihood	• Describe the unnormalized or normalized likelihood if the prior distribution is informative.
Plots	• Include the prior distribution, likelihood, and posterior distribution in a single plot (i.e., a triplot) if the prior distribution is informative and plots are to be presented.
Decisions	• Report the utilities, or costs and benefits, and explain how they were derived if the data are used for decision making about possible actions. Also provide a sensitivity analysis for various prior distributions or assumptions about utilities for the decision.
Special cases	• Explain the rationale for assuming exchangeability (or conditional exchangeability if there are covariates) for multilevel analyses. If relevant to the research context, present plots or tables of shrinkage-adjusted estimates and their confidence intervals. • Report forest plots or caterpillar plots that include original and shrinkage-corrected estimates of effect sizes for each study with confidence intervals for meta-analytic summaries. If feasible for the analytic method, provide a parameter trace plot where shrinkage-adjusted estimates are shown against the standard deviation of the residual effects, combined with the posterior distribution of the residual variance. • Describe the details of all decision rules, if these rules were decided (before or during the study), and the consequences (results) of each decision in adaptive designs.
Computations	• Describe in detail, including the number of chains, the number of burn-in iterations for each chain, and thinning if Markov chain Monte Carlo (MCMC) or another sampling procedure is used. Specify the methods used to check for convergence and their results.
Model fit	• Describe the procedures used to check the fit of the model, and the results of those checks.
Bayes factors	• Specify the models being compared if Bayes Factors are calculated. • Report the Bayes Factors and how they were interpreted. • Test the sensitivity of the Bayes Factors to assumptions about prior distributions.
Bayesian model averaging	• State the parameter or function of parameters being estimated in Bayesian model averaging. Either plot the distribution or list the mean and standard deviation if it is near normal; otherwise, list a number of percentiles for the distribution if it is not near normal. • Describe how the models were generated and if a reduced set was used for averaging, how the selection was made and which models were used in the averaging.

are many. As with the original JARS, they share many features of other widely used systems such as CONSORT. In many cases, the Working Group has fairly directly adapted the schema of other systems (e.g., *N*-of-1 reporting standards); in other cases, other reporting systems are referred to without any details (e.g., standards for reporting measures of neural activities); and in some cases, completely new standards were developed, sometimes aided by individuals or groups of individuals who were not members of the Working Group (e.g., standards for reporting results using Bayesian analyses, longitudinal studies, and studies purporting to be replication studies). As in the original JARS, the ways in which the tables might be used (e.g., as a formal checklist, as a reference list for editors, authors, and reviewers) are not specified. Some editors may wish to develop formal checklists; some educators who teach scientific writing for psychology (see Cooper, 2011) might wish to utilize the tables as part of classroom handouts. The purpose of the tables in this article is as a communication device to organize otherwise complex and numerous ideas.

No matter the approach, the goal was to provide reporting standards that would be appropriate for most of the empirical, quantitative research conducted by those individuals who identify their work, at least in part, as behavioral, social, or educational science. The intent was to span as wide a range of work as possible, including randomized clinical trials, single-case designs, observational studies, longitudinal studies, research synthesis, and other forms of empirical study—including qualitative studies and mixed method studies that were the province of the JARS–Qual Working Group.

Although the focus for the development of JARS–Quant was the social, behavioral, and educational sciences, the Working Group stayed keenly aware that many fields use some of these domains in their work—medicine, nursing, law, and

Table 9
Information Recommended for Inclusion in Manuscripts Reporting Meta-Analyses

Paper section and topic	Description
Title	• State the research question and type of research synthesis (e.g., narrative synthesis, meta-analysis).
Author note	• List all sources of monetary and in-kind funding support; state the role of funders in conducting the synthesis and deciding to publish the results, if any. • Describe possible conflicts of interest, including financial and other nonfinancial interests. • Give the place where the synthesis is registered and its registry number, if registered. • Provide name, affiliation, and e-mail address of corresponding author.
Abstract	
Objectives	• State the research problems, questions, or hypotheses under investigation.
Eligibility criteria	• Describe the characteristics for inclusion of studies, including independent variables (treatments, interventions), dependent variables (outcomes, criteria), and eligible study designs.
Methods of synthesis	• Describe the methods for synthesizing study results, including • Statistical and other methods used to summarize and to compare studies • Specific methods used to integrate studies if a meta-analysis was conducted (e.g., effect-size metric, averaging method, the model used in homogeneity analysis)
Results	• State the results of the synthesis, including • Number of included studies and participants, and their important characteristics • Results for the primary outcome(s) and moderator analyses • Effect size(s) and confidence interval(s) associated with each analysis if aa meta-analysis was conducted
Conclusions	• Describe strengths and limitations of the evidence, including evidence of inconsistency, imprecision, risk of bias in the included studies and risk of reporting biases.
Introduction	
Problem	• State the question or relation(s) under investigation, including • Historical background, including previous syntheses and meta-analyses related to the topic • Theoretical, policy, and/or practical issues related to the question or relation(s) of interest • Populations and settings to which the question or relation(s) is relevant • Rationale for (a) choice of study designs, (b) the selection and coding of outcomes, (c) the selection and coding potential moderators or mediators of results • Psychometric characteristics of outcome measures and other variables
Objectives	• State the hypotheses examined, indicating which were prespecified, including • Question in terms of relevant participant characteristics (including animal populations), independent variables (experimental manipulations, treatments, or interventions), ruling out of possible confounding variables, dependent variables (outcomes, criterion), and other features of study designs • Method(s) of synthesis and if meta-analysis was used, the specific methods used to integrate studies (e.g., effect-size metric, averaging method, the model used in homogeneity analysis)
Protocol	• List where the full protocol can be found (e.g., a supplement), or state that there was no protocol. State that the full protocol was published (or archived in a public registry) or that it was not published before the review was conducted.
Method	
Inclusion and exclusion criteria	• Describe the criteria for selecting studies, including • Independent variables (e.g., experimental manipulations, types of treatments or interventions or predictor variables) • Dependent variable (e.g., outcomes, in syntheses of clinical research including both potential benefits and potential adverse effects) • Eligible study designs (e.g., methods of sampling or treatment assignment) • Handling of multiple reports about the same study or sample, describing which are primary and handling of multiple measures using the same participants • Restrictions on study inclusion (e.g., by study age, language, location, or report type) • Changes to the prespecified inclusion and exclusion criteria, and when these changes were made • Handling of reports that did not contain sufficient information to judge eligibility (e.g., lacking information about study design) and reports that did not include sufficient information for analysis (e.g., did not report numerical data about those outcomes)
Information sources	• Describe all information sources: • Search strategies of electronic searches, such that they could be repeated (e.g., include the search terms used, Boolean connectors, fields searched, explosion of terms) • Databases searched (e.g., PsycINFO, ClinicalTrials.gov), including dates of coverage (i.e., earliest and latest records included in the search), and software and search platforms used • Names of specific journals that were searched and the volumes checked • Explanation of rationale for choosing reference lists if examined (e.g., other relevant articles, previous research syntheses) • Documents for which forward (citation) searches were conducted, stating why these documents were chosen • Number of researchers contacted if study authors or individual researchers were contacted to find studies or to obtain more information about included studies, as well as criteria for making contact (e.g., previous relevant publications), and response rate

(*table continues*)

Table 9 (*continued*)

Paper section and topic	Description
	• Dates of contact if other direct contact searches were conducted such as contacting corporate sponsors or mailings to distribution lists • Search strategies in addition to those above and the results of these searches
Study selection	• Describe the process for deciding which studies would be included in the syntheses and/or included in the meta-analysis, including • Document elements (e.g., title, abstract, full text) used to make decisions about inclusion or exclusion from the synthesis at each step of the screening process • Qualifications (e.g., training, educational or professional status) of those who conducted each step in the study selection process, stating whether each step was conducted by a single person or in duplicate as well as an explanation of how reliability was assessed if one screener was used and how disagreements were resolved if multiple were used
Data collection	• Describe methods of extracting data from reports, including • Variables for which data were sought and the variable categories • Qualifications of those who conducted each step in the data extraction process, stating whether each step was conducted by a single person or in duplicate and an explanation of how reliability was assessed if one screener was used and how disagreements were resolved if multiple screeners were used as well as whether data coding forms, instructions for completion, and the data (including metadata) are available, stating where they can be found (e.g., public registry, supplemental materials)
Methods for assessing risk to internal validity	• Describe any methods used to assess risk to internal validity in individual study results, including • Risks assessed and criteria for concluding risk exists or does not exist • Methods for including risk to internal validity in the decisions to synthesize of the data and the interpretation of results
Summary measures	• Describe the statistical methods for calculating effect sizes, including the metric(s) used (e.g., correlation coefficients, differences in means, risk ratios) and formula(s) used to calculate effect sizes.
Methods of synthesis	• Describe narrative and statistical methods used to compare studies. If meta-analysis was conducted, describe the methods used to combine effects across studies and the model used to estimate the heterogeneity of the effects sizes (e.g., a fixed-effect, random-effects model robust variance estimation), including • Rationale for the method of synthesis • Methods for weighting study results • Methods to estimate imprecision (e.g., confidence or credibility intervals) both within and between studies • Description of all transformations or corrections (e.g., to account for small samples or unequal group numbers) and adjustments (e.g., for clustering, missing data, measurement artifacts, or construct-level relationships) made to the data and justification for these • Additional analyses (e.g., subgroup analyses, meta-regression), including whether each analysis was prespecified or post hoc • Selection of prior distributions and assessment of model fit if Bayesian analyses were conducted • Name and version number of computer programs used for the analysis • Statistical code and where it can be found (e.g., a supplement)
Publication bias and selective reporting	• Address the adequacy of methods used (e.g., contacting authors for unreported outcomes to identify unpublished studies and unreported data). Describe any statistical methods used to test for publication bias and selective reporting or address the potential limitations of the synthesis's results if no such methods were used.
Results	
Study selection	• Describe the selection of studies, ideally with a flowchart, including • Number of citations assessed for eligibility • Number of citations and number of unique studies included in the syntheses • Reasons for excluding studies at each stage of screening • Table with complete citations for studies that met many but not all inclusion criteria with reasons for exclusion (e.g., effect size was not calculable)
Study characteristics	• Summarize the characteristics of included studies. Provide a table showing, for each included study, the principle variables for which data were sought, including • Characteristics of the independent and outcome or dependent variables and main moderator variables • Important characteristics of participants (e.g., age, sex, ethnicity) • Important contextual variables (e.g., setting, date) • Study design (e.g., methods of sampling or treatment assignment). • Report where the full data set is available (e.g., from the authors, supplemental materials, registry)
Results of individual studies	• Report the results for each study or comparison (e.g., the effect size with confidence intervals for each independent variable). If possible, present this information in a figure (e.g., forest plot).
Synthesis of results	• Report a synthesis (e.g., meta-analysis) for each study result (e.g., weighted average effect sizes, confidence intervals, estimates of heterogeneity of results).

Table 9 (*continued*)

Paper section and topic	Description
Assessment of internal validity of individual studies	• Describe risks of bias different design features might introduce into the synthesis results.
Publication and reporting bias	• Describe risk of bias across studies, including • Statement about whether (a) unpublished studies and unreported data, or (b) only published data were included in the synthesis and the rationale if only published data were used • Assessments of the impact of publication bias (e.g., modeling of data censoring, trim-and-fill analysis) • Results of any statistical analyses looking for selective reporting of results within studies
Adverse and harmful effects	• Report any adverse or harmful effects identified in individual studies.
Discussion	
Summary of the evidence	• Summarize the main findings, including • Main results of the synthesis, including all results of prespecified analyses • Overall quality of the evidence • Strengths and limitations (e.g., inconsistency, imprecision, risk of bias, and publication bias or selective outcome reporting) of findings • Alternative explanations for observed results (e.g., confounding, statistical power) • Similarities and differences with previous syntheses
Generalizability	• Describe the generalizability (external validity) of conclusions, including • Implications for related populations, intervention variations, dependent (outcome) variables
Implications	• Interpret the results in light of previous evidence. • Address the implications for further research, theory, policy, and/or practice.

social work, to name but a few. In developing the JARS–Quant recommendations, the Working Group tried to ensure that these reporting standards would be usable for scholarly work in other fields but ones that incorporate aspects of the social, behavioral, and educational sciences into them.

The implementation of reporting standards is a slow process, with many individuals contributing to realizing the goals set forth in the JARS–Quant. Certainly, journal editors, associate editors, and manuscript reviewers will be prime movers of the adoption of any set of standards. This fact, however, implies that scholars who have the most contact with the authors of journal articles must be encouraged to be aware of these reporting standards and encourage their use. It is also incumbent upon publishers (such as APA) to provide training materials that present these standards in convenient formats. This is essential if those entering the field—indeed anyone committed to the advancement of transparent social sciences—are to acquire the habit of utilizing reporting standards as a part of their formulation of how scholarly research is reported.

Much can be said about the value of adopting reporting standards and developing ways to provide creative and forward thinking tools for communicating and training those who use these standards. However, most basic is the realization that these reporting standards are important ways of systematically communicating the work that scholars have done in the process of doing their science. These standards do not specify how the science should be done but rather what about the science needs to be reported so that (a) scientific claims can be clearly understood, assessed, and evaluated by the reader, and (b) the work can be replicated with reasonable accuracy such that a replication would reflect the science being reported. The Working Group does, at the same time, recognize that specifying what needs to be reported also influences what data need to be gathered (one cannot report the reasons for drop-out if one does not collect that data). Therefore, a command of the reporting standards is a critical part of the initial planning of an empirical study. This initial understanding may not only ease the task of complete and transparent reporting but also improve the implementation of the research.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: Author.

American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.

APA Publications and Communications Board Working Group on Journal Article Reporting Standards. (2008). Reporting standards for research in psychology: Why do we need them? What might they be? *American Psychologist, 63,* 839–851. http://dx.doi.org/10.1037/0003-066X.63.9.839

Atkinson, K. M., Koenka, A. C., Sanchez, C. E., Moshontz, H., & Cooper, H. (2015). Reporting standards for literature searches and report inclusion criteria: Making research syntheses more transparent and easy to replicate. *Research Synthesis Methods, 6,* 87–95. http://dx.doi.org/10.1002/jrsm.1127

Barker, J. B., Mellalieu, S. D., McCarthy, P. J., Jones, M. V., & Moran, A. (2013). A review of single-case research in sport psychology 1997–2012: Research trends and future directions. *Journal of Applied Sport Psychology, 25,* 4–32. http://dx.doi.org/10.1080/10413200.2012.709579

Begley, C. G., & Ellis, L. M. (2012, March 28). Drug development: Raise standards for preclinical cancer research. *Nature, 483,* 531–533. http://dx.doi.org/10.1038/483531a

Borrelli, B. (2011). The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *Journal of Public Health Dentistry, 71*(Suppl. 1), S52–S63. http://dx.doi.org/10.1111/j.1752-7325.2011.00233.x

Cooper, H. (2011). *Reporting research in psychology: How to meet journal article standards*. Washington, DC: American Psychological Association.

Cybulski, L., Mayo-Wilson, E., & Grant, S. (2016). Improving transparency and reproducibility through registration: The status of intervention trials published in clinical psychology journals. *Journal of Consulting and Clinical Psychology, 84,* 753–767. http://dx.doi.org/10.1037/ccp0000115

Didden, R., Korzilius, H., van Oorsouw, W., & Sturmey, P. (2006). Behavioral treatment of challenging behaviors in individuals with mild mental retardation: Meta-analysis of single-subject research. *American Journal on Mental Retardation, 111,* 290–298. http://dx.doi.org/10.1352/0895-8017(2006)111

Duggan, C., Parry, G., McMurran, M., Davidson, K., & Dennis, J. (2014). The recording of adverse events from psychological treatments in clinical trials: Evidence from a review of NIHR-funded trials. *Trials, 15,* 335. http://dx.doi.org/10.1186/1745-6215-15-335

Hoyle, R. H., & Isherwood, J. C. (2013). Reporting results from structural equation modeling analyses in Archives of Scientific Psychology. *Archives of Scientific Psychology, 1,* 14–22. http://dx.doi.org/10.1037/arc0000004

Kilkenny, C., Browne, W. J., Browne, W. J., Cuthill, I. C., Emerson, M., & Aktman, D. G. (2010). Improving bioscience research reporting: The ARRIVE guidelines for reporting animal research. *PLoS Biology, 8*(6), e1000412. http://dx.doi.org/10.1371/journal.pbio.1000412

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). New York, NY: Guilford Press.

Mackey, A. (2012). Why (or why not), when, and how to replicate research. In G. Porte (Ed.), *Replication research in applied linguistics* (pp. 21–46). New York, NY: Cambridge University Press.

Maggin, D. M., Chafouleas, S. M., Goddard, K. M., & Johnson, A. H. (2011). A systematic evaluation of token economies as a classroom management tool for students with challenging behavior. *Journal of School Psychology, 49,* 529–554. http://dx.doi.org/10.1016/j.jsp.2011.05.001

Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does "failure to replicate" really mean? *American Psychologist, 70,* 487–498. http://dx.doi.org/10.1037/a0039400

Montgomery, P., Grant, S. P., Hopewell, S., Macdonald, G., Moher, D., Michie, S., & Mayo-Wilson, E. (2013, September 2). Protocol for CONSORT-SPI: An extension for social and psychological interventions. *Implementation Science, 8,* 1–7. http://dx.doi.org/10.1186/1748-5908-8-99

Montgomery, P., Underhill, K., Gardner, F., Operario, D., & Mayo-Wilson, E. (2013). The Oxford Implementation Index: A new tool for incorporating implementation data into systematic reviews and meta-analyses. *Journal of Clinical Epidemiology, 66,* 874–882. http://dx.doi.org/10.1016/j.jclinepi.2013.03.006

Mueller, R. O., & Hancock, G. R. (2008). Best practices in structural equation modeling. In J. W. Osborne (Ed.), *Best practices in quantitative methods* (pp. 488–508). Thousand Oaks, CA: Sage.

National Institutes of Health. (2014, October). *Notice of revised NIH definition of "clinical trial"* (Notice No. NOT-OD-15-015). Retrieved from https://grants.nih.gov/grants/guide/noticefiles/

Nezu, A. M., & Nezu, C. M. (2008). Ensuring treatment integrity. In A. M. Nezu & C. M. Nezu (Eds.), *Evidence-based outcome research: A practical guide to conducting randomized clinical trials for psychosocial interventions* (pp. 263–281). New York, NY: Oxford University Press.

Nichols, T. E., Das, S., Eickhoff, S. B., Evans, A. C., Glatard, T., Hanke, M., . . . Yeo, B. T. (2017). Best practices in data analysis and sharing in neuroimaging using MRI. *Nature Neuroscience, 20,* 299–303. http://dx.doi.org/10.1038/nn.4500

Nosek, B. A., & Lakens, D. (Eds.). (2013). Replications of important results in social psychology [Special issue]. *Social Psychology, 44*.

Open Science Collaboration. (2015, August 28). Estimating the reproducibility of psychological science. *Science, 349,* aac4716. http://dx.doi.org/10.1126/science.aac4716

Peterson, A. L., Roache, J. D., Raj, J., & Young-McCaughan, S. (2012). The need for expanded monitoring of adverse events in behavioral health clinical trials. *Contemporary Clinical Trials, 34,* 152–154. http://dx.doi.org/10.1016/j.cct.2012.10.009

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., Jr, . . . Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology, 37,* 127–152. http://dx.doi.org/10.1111/1469-8986.3720127

Reynolds, C. R., & Livingston, R. B. (2012). *Mastering modern psychological testing: Theory & methods*. Boston, MA: Pearson.

Schulz, K. F., Altman, D. G., Moher, D., & the CONSORT Group. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomized trials. *Annals of Internal Medicine, 152,* 726–732. http://dx.doi.org/10.7326/0003-4819-152-11-201006010-00232

Schumaker, R. E., & Lomax, R. G. (2016). *A beginner's guide to structural equation modeling* (4th ed.). New York, NY: Routledge.

Shamseer, L., Sampson, M., Bukutu, C., Schmid, C. H., Nikles, J., Tate, R., . . . CENT group. (2015). CONSORT extension for reporting N-of-1 Trials (CENT) 2015: Explanation and elaboration. *British Medical Journal, 350,* h1793. http://dx.doi.org/10.1136/bmj.h1793

Slaney, K. L., Tkatchouk, M., Gabriel, S. M., & Maraun, M. D. (2009). Psychometric assessment and reporting practices: Incongruence between theory and practice. *Journal of Psychoeducational Assessment, 27,* 465–476. http://dx.doi.org/10.1177/0734282909335781

Smith, J. D. (2012). Single-case experimental designs: A systematic review of published research and current standards. *Psychological Methods, 17,* 510–550. http://dx.doi.org/10.1037/a0029312

Stroup, D. F., Berlin, J. A., Morton, S. C., Olkin, I., Williamson, G. D., Rennie, D., . . . Thacker, S. B. (2000, April 19). Meta-analysis of observational studies in epidemiology: A proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *Journal of the American Medical Association, 283,* 2008–2012.

Tate, R. L., Perdices, M., McDonald, S., Togher, L., & Rosenkoetter, U. (2014). The design, conduct and report of single-case research: Resources to improve the quality of the neurorehabilitation literature. *Neuropsychological Rehabilitation, 24,* 315–331. http://dx.doi.org/10.1080/09602011.2013.875043

Tate, R. L., Perdices, M., Rosenkoetter, U., McDonald, S., Togher, L., Shadish, W., . . . Vohra, S. (2016). The Single-Case Reporting guideline In BEhavioural interventions (SCRIBE) 2016: Explanation and elaboration. *Archives of Scientific Psychology, 4,* 10–31. http://dx.doi.org/10.1037/arc0000027

Tate, R. L., Perdices, M., Rosenkoetter, U., Shadish, W., Vohra, S., Barlow, D. H., . . . Wilson, B. (2016). The Single-Case Reporting guideline In BEhavioural interventions (SCRIBE) 2016 statement. *Archives of Scientific Psychology, 4,* 1–9. http://dx.doi.org/10.1037/arc0000026

Thompson, B. (Ed.), (2003). *Score reliability*. Thousand Oaks, CA: Sage.

Tooth, L., Ware, R., Bain, C., Purdie, D. M., & Dobson, A. (2005). Quality of reporting of observational longitudinal research. *American Journal of Epidemiology, 161,* 280–288. http://dx.doi.org/10.1093/aje/kwi042

University of York. (n.d.). *Welcome to PROSPERO*. Retrieved from http://www.crd.york.ac.uk/PROSPERO/

Vohra, S., Shamseer, L., Sampson, M., Bukutu, C., Schmid, C. H., Tate, R., . . . CENT group. (2015). CONSORT extension for reporting *N*-of-1 trials (CENT) 2015 statement. *British Medical Journal, 350,* h1738. http://dx.doi.org/10.1136/bmj.h1738

World Medical Association. (2013, November 27). World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. *Journal of the American Medical Association, 310,* 2191–2194. http://dx.doi.org/10.1001/jama.2013.281053

Received September 6, 2016
Revision received June 27, 2017
Accepted June 29, 2017 ■

Correction to Appelbaum et al. (2018)

The article "Journal Article Reporting Standards for Quantitative Research in Psychology: The APA Publications and Communications Board Task Force Report" by Mark Appelbaum, Harris Cooper, Rex B. Kline, Evan Mayo-Wilson, Arthur M. Nezu, and Stephen M. Rao (*American Psychologist*, 2018, Vol. 73, No. 1, pp. 3–25. http://dx.doi.org/10.1037/amp0000191) contained a citation error. In the "Clinical Trials" subsection in the section, "Reporting Standards for Studies With an Experimental Manipulation" (p. 13), the reference to the World Medical Association's Declaration of Helsinki should be to the 2008 version of the declaration. The sentence should read as follows:

> From an ethical perspective, the Declaration of Helsinki, which is the set of ethical principles regarding human experimentation developed by the World Medical Association, Inc. (2008), stated that "every clinical trial must be registered in a publicly accessible database before recruitment of the first subject" (Item 19, p. 3).

The corrected reference, on p. 25, is the following:

> World Medical Association, Inc. (2008). Declaration of Helsinki: Ethical principles for medical research involving human subjects. Retrieved from https://www.wma.net/wp-content/uploads/2018/07/DoH-Oct2008.pdf

http://dx.doi.org/10.1037/amp0000389