Desrosiers, E., Sherony, K., Barros, E., Ballinger, G. A., Senol, S., & Campion, M. A. (2002). Writing research articles: Update to the Article Review Checklist. In S. G. Rogelberg (Ed.), Handbook of research methods in Industrial and Organizational Psychology. London: Blackwell.

CHAPTER TWENTY-TWO

Writing Research Articles: Update on the Article Review Checklist

Erica I. Desrosiers, Kathryn Sherony, Eduardo Barros, Gary A. Ballinger, Sinem Senol, and Michael A. Campion

In 1993, Michael Campion published the "Article review checklist: A criterion checklist for reviewing research articles in applied psychology."[1] As a good deal of time has passed since the publication of the checklist, we felt it appropriate to investigate if there have been any shifts in thinking or developments in the ways articles are reviewed and evaluated. For the purposes of this Handbook, we interviewed a sample of recent editors and associate editors of top-tier journals.[2] Twelve out of twenty-one individuals contacted (57 percent) responded to this inquiry. Specifically, each was asked to identify what he or she believed to be the five characteristics of a publishable journal article. There were a total of 60 characteristics identified by the editors. We content-analyzed this group of comments to identify the predominant themes (see table 22.1).

There are two primary purposes of this chapter. The first is to reprint Campion's original article review checklist as a methodological tool for evaluating the quality of research articles (see appendix 22.1). The second purpose is to update the checklist by presenting the results of interviews with a sample of editors. In the original checklist a two-step Delphi-like procedure was followed. In the first step, a list of 93 criteria for reviewing research articles was developed and circulated among journal reviewers, who provided detailed feedback and added several new criteria. The new list of criteria was categorized and sent back to the reviewers. In this second step, they were asked to rate each criterion according to the importance it should have when reviewing an article. After processing this information, the final checklist contained 223 criteria, 44 clusters, and 15 categories (for more detail about the process followed to create the final checklist, see appendix 22.1).

We had several reasons for expecting to find changes in article review criteria since the publication of Campion's checklist. For example, advances in statistical techniques in the past decade may have pushed editors toward a stronger focus on analytic technique.

Table 22.1 Summary of editors' responses

1. **Quality of methods and analysis** (17 of 60 items; 11 of 12 editors)
   1A. *Appropriate research design* (11/17)
      - *design is constructed so as to rule out alternative explanations*
      - *method (including sample) appropriate for the question*
      - *data collected can address the question*

   1B. *Quality of measures* (3/17)
      - *measures used are reliable, valid, and interpretable*
      - *rigor in terms of measurement of the variables*

   1C. *Appropriate data analysis* (3/17)
      - *proper use of data analytic approaches and techniques*
      - *minimally sufficient statistics*

2. **Importance of the research question** (15 of 60 items; 12 of 12 editors)
   2A. *Unique contribution* (11/15)
      - *question is compelling, important*
      - *answering question would contribute to our knowledge, close gaps, advance theory*
      - *make case for why we should be interested, how we would benefit from knowing the answer to the question*

   2B. *Practical significance* (4/15)
      - *should have clear value to practitioners*
      - *practical relevance/usefulness of ideas should be stated*

3. **Conceptual development and definition** (11 of 60 items; 8 of 12 editors)
      - *sound grounding in relevant literature*
      - *constructs are well defined within the relevant literature*
      - *logical conceptual development*

4. **Writing style** (8 of 60 items; 6 of 12 editors)
      - *well organized, structured, easy to follow*
      - *tells a coherent, straightforward story*
      - *clear writing*

5. **Defensible and appropriate conclusions** (3 of 60 items; 3 of 12 editors)
      - *possible to make sense of the findings*
      - *conclusions are valid, appropriate, and defensible*

6. **Miscellaneous** (6 of 60 items; 3 of 12 editors)
      - *acknowledge limitations*
      - *suggest directions for future research*
      - *conduct multiple lab studies (as opposed to single study)*
      - *collect field rather than lab data for I-O psychology*
      - *stress external validity*
      - *borrow theory from other areas*

Note. Some comments counted in multiple categories, and some editors made multiple comments on the same category.

Also, we suspected it possible that industry trends may have influenced differential importance of topic areas. We also looked for any changes in emphasis on the part of editors, for instance to see if there have been increases in rigor requirements. We interviewed editors rather than reviewers because we felt it to be more expeditious to go to the final source of editorial decisions.

The characteristics mentioned by the editors comprise six major categories (see table 22.1): methods and analysis, importance of question, conceptual development and definition, writing style, defensible and appropriate conclusions, and miscellaneous. The largest category, consisting of 17 of the 60 comments, is "Quality of Methods and Analysis." Although all comments in this category relate to methods or analysis, we felt this category was broad enough to warrant separating the comments into three subgroups: appropriateness of the research design (11 of 17 comments), quality of the measures (3 of 17), and the use of appropriate data analysis methods (3 of 17).

Comments regarding research design expressed concern that the design should be solid and allow the researcher to rule out confounds or alternative explanations for results, as well as address the question in an appropriate methodological sense. Several editors commented that many papers pose fascinating questions and then design a study that is unable to answer the question. Whereas many editors noted the importance of the rigor of the methodology, one editor commented that the rigor should be qualified by the relevance of the question. This editor observed: "If the question is important enough, but cannot be addressed with a rigorous methodology, a weaker method could be acceptable. The same could be true in the case of new questions or areas of research." Another editor noted the importance of the research situation being a reasonable one in terms of fully activating the processes the investigator wishes to study. This editor gave the following example: "If one is looking at motivation, are the critical processes likely to be engaged by the research situation, and are these processes likely to generalize?" An additional design factor raised by several editors was use of an appropriate sample for the question. One editor mentioned that student samples are acceptable for some purposes but inappropriate for others, such as when the research question relates to real managerial decisions.

Comments regarding quality of measures and data analysis were fairly straightforward – editors want to see that the measures used by researchers are reliable, valid, and interpretable and that appropriate data analytic techniques were used. One editor pointed out that when and how the measures are taken must be carefully detailed for the reader. Another editor expressed concern for what he sees as a "growing trend toward fancier statistical methods, often at the expense of telling a clear story that can reach a wide audience." This editor promoted the use of "minimally sufficient statistics," not inappropriately simple, but also not unduly fancy and complex. If the question can be adequately answered with a t-test, then only use a t-test. Along the same lines, this editor added that the paper should be "written for the general reader, not the methodological specialist."

We speculated that this category of items (methods and analysis) was mentioned most often not only because the quality of data analysis methods speaks directly to the quality of the overall work but also because there are more objective standards for research methods than for most of the other factors, such as writing style or importance of question. That this subject is a high priority for editors highlights the importance of advance planning in research design. Improper design and low-quality measurement are

frequently unrecoverable. An inappropriate research design or the use of low-quality measures are things that ordinarily cannot be remedied by revision. These problems are often perceived as fatal flaws. If a submission is not solid with regard to these foundation issues of research methods and data analysis, it does not stand a high chance for publication.

With 15 of 60 mentions, "Importance of Question" is the second largest category. Comments in this category also fell neatly into two sub-groups: unique contribution (11 of 15 mentions), and practical significance (4 of 15 mentions). Typical comments in the "unique contribution" subgroup expressed the view that the research question should be compelling and address an important issue, and that answering the question should advance theory and add substantially to our knowledge. Several editors specifically stated that authors need to make a better case in the introduction for why we should be interested, why the work is important, and how we would benefit from knowing the answer to their research question. Essentially, authors need to provide a good answer to the "so what?" question. As one editor explained: "Researcher time and resources are scarce; journal reader time is scarce; journal space is scarce. As such, I value highly a paper that has the potential to change the way we think and am less enamored of 'small increment in knowledge' papers." As another editor phrased it more simply: "Unless the paper adds substantially to what we already know in the topic area, it simply reduces the signal/noise ratio and contributes to deforestation." Apparently, a good question is necessary but not sufficient. You must also convincingly sell the reader on the importance of the question. As one editor noted: "The importance of the question presented is not always self-evident. The way people present their research questions is essential. You have to motivate the problem . . . in order to do that, you have to understand your audience, its critical view in particular . . . It is important to know the typical reader of your work. You have to make them say, 'Hey, this is worth my attention, it deals with something important.' Most of the time the importance is essentially in the mind of the reader."

To some extent, we were surprised that the answers of the editors did not allow us to include "replication" as an important element of the category "Importance of Question." On one side, we think that this could reflect contradictory perspectives that journals adopt regarding replication. For example, the *Academy of Management Journal*'s "Information for Authors" explicitly discourages replications, while the instructions of the *Journal of Applied Psychology* mention "important replications" as potential short articles. This lack of focus on replication can simply be a product of the type of question asked to the editors. It is possible that making the editors rank the five most important aspects of publishable articles did not lead them to emphasize replication as a priority. However, it does not seem fair to conclude that they would discourage replication as an important characteristic of a study. Quality replications are always needed, in any science, and are particularly important in organizational sciences where abstract constructs can be operationalized in many different ways, potentially leading to incongruent results. In those situations, which are not uncommon in organizational sciences, sound replications are highly valuable. In this vein, the original checklist can serve as a good guide. In different parts of the checklist the importance of clearly presenting methods and procedures is emphasized, because accuracy in that regard is a prerequisite to good replications. In

addition, the original checklist recognizes the contribution of articles that provide a "constructive replication" (e.g., replicates but extends in an important way). Nevertheless, based on the results of the present survey, we think that replications will clearly have to bring interesting, new ideas to light if they hope to stand a chance of publication.

As for practical significance, four of the twelve editors responded that this is one of the five most important characteristics of a publishable article. These editors felt that published works should have clear value to practitioners and should "make a contribution of practical significance for the change and enhancement of organizational practice." One editor stated that "I look for papers that show me how to make a better practice." As with the issue of unique contribution, it is preferable for the author to clearly state the practical relevance or usefulness of their ideas rather than assume that the reader will appreciate the practical significance on their own.

Whereas newness or unique contribution was not the most frequently mentioned factor overall, it was often the *first* factor mentioned by responding editors and was mentioned by all twelve of the editors. Several respondents commented that the ability of a manuscript to make a novel contribution to the literature was by far the most critical factor in reviewing a work. Those who are publishing in our top journals should be consistently on the cutting edge in theoretical development. A focus on the newness of the work to be published ensures that it is fresh as well as relevant to academics and practitioners.

Comprising the third category of factors is the conceptual and theoretical development of the work, mentioned 11 of 60 times by eight different editors. Responding editors felt strongly that all manuscripts should be soundly grounded in relevant theory and literature, include well-defined questions and constructs, and be characterized by good logical development. This emphasis places a burden on authors to ensure that they perform comprehensive literature reviews in order to properly place the work in the context of the field of research. Editors commented that authors must "build off and recognize past research and show familiarity with the history of the literature," and "tie their perspective to the extant literature and show that they are not rediscovering the wheel." The stress on theory development requires that authors fully develop and articulate the bases for their research propositions, and not focus only on the research methods. One editor commented: "Too often you can see methodologically correct work, but without theory behind it. Therefore, the contribution coming from that type of work is very limited." Authors need to present a "defensible, coherent, literature-based rationale."

The fourth category was writing style, to which 8 of the 60 comments related. The editors felt that manuscripts should be well organized and tell a coherent, straightforward story. Whereas writing style can often be polished upon revision, there is great advantage in a first submission telling an easy-to-follow and yet compelling story. It helps predispose a positive initial review. As one editor noted: "Many manuscripts meander, introducing lots of extraneous citations and reviewing peripheral literature. I can't say enough for the tremendous appeal of a manuscript which tells a straightforward story clearly and directly." Authors need to take time to organize their work and should stick to an outline in writing their manuscripts. Writing should be clear and concise, and ideas should logically follow from one another. One editor added that good tables, figures, and examples can help, although another editor clarified that examples should be

supplemental in nature: "they are not substitutes for good writing." Although obvious, several editors also mentioned that manuscripts should be free of grammatical errors and consistent with stylistic requirements. The fact that several editors mentioned this seemingly obvious characteristic suggests that many researchers submit work that is careless in this regard. Such errors may inadvertently send the message that there are also errors in other unseen aspects of the research. One editor stated: "It is amazing to me how many authors fail even to consult the standard 'style manuals' and 'information for contributors' that most journals clearly identify."

Comments from three editors comprised category five, "Defensibility and appropriateness of conclusions." The editors expressed concern that conclusions are valid and follow logically from the results. It is important for authors to ensure that their conclusions are based on their findings and that the data match the conclusions drawn. Authors should not extrapolate from their data to make conclusions that are not warranted. It is also critical that authors ensure that their research methods support the conclusions made. For example, it is improper to make causal conclusions from a non-experimental or correlational study.

A few unrelated comments (6 of 60) were grouped into a miscellaneous category. One editor commented on the value of appropriately acknowledging limitations in a manuscript. Another commented on the value of suggesting directions for future research. A third editor stressed the value of augmenting a literature stream by borrowing theory from other areas. Finally, preferences were expressed for strong external validity, field data over lab data in the field of I-O psychology, and multiple studies as opposed to single studies when reporting lab research.

While the ideas captured here are not particularly novel, we hope that they reveal the areas which are most heavily valued by editors – the areas which could be perceived as "fatal flaws." The viewpoints of these editors may be used to evaluate research articles before submission and perhaps even throughout the planning process. It can be easy to overlook some of the key points made by them when one is deeply involved in a research project, but to do so may cost the chance for publication.

As stated earlier, the purpose of this appendix was to update Campion's original article checklist and to find out whether there have been any changes in article reviewing criteria. It should be clear to the reader that the comments of editors outlined above overlap significantly with the original checklist. In spite of the enhancement of our data analysis techniques and the constant changes in the operation of organizations, editors still value similar aspects of good research: appropriate methods of analysis and important questions. Although the original checklist is more extensive and has a broader scope (i.e., it presents almost all criteria that reviewers could think of at that time), this later investigation provides a guide to aspects of journal articles that seem to be receiving special attention on the part of editors. Indeed, it does not seem likely that all of the criteria presented in the original checklist merit equal attention of editors and reviewers when dealing with a particular manuscript. As stated before, some aspects of the original checklist seem to be at the top of editors' minds and could in fact receive more weight when publication decisions are made. Therefore, when writing a research manuscript it is advisable to take into account both the original article review checklist and the key issues of concern to editors as outlined here and in table 22.1.

Notes

1. The checklist is reprinted by permission from *Personnel Psychology*. Campion, M. A. (1993). Article review checklist: A criterion checklist for reviewing research articles in applied psychology. *Personnel Psychology*, 46, 705–18.
2. The journals sampled were: *Academy of Management Journal*, *Academy of Management Review*, *Journal of Applied Psychology*, *Personnel Psychology*, and *Organizational Behavior and Human Decision Processes*. We wish to thank all the editors who participated in this survey for their valuable time and generous contribution that has made this work possible.

Appendix 22.1

Article Review Checklist: A Criterion Checklist for Reviewing Research Articles in Applied Psychology

Michael A. Campion

Over the past couple of years, a large group of reviewers and I have been developing a comprehensive checklist of criteria for reviewing research articles. The purpose of this effort was to provide a heuristic device of issues to think about when reviewing an article. As such, we hoped that the checklist might be a useful tool for reviewers, authors, and students.

A two-part Delphi-like procedure was followed. In the first part, a preliminary checklist of criteria was developed and circulated to a large number of reviewers. It contained 93 criteria and was divided into 14 categories (e.g., literature review, sample, measures, procedures, analyses, conclusions, etc.). Reviewers were asked to examine the list and think about what they look for when reviewing an article, and then to modify the criteria on the list or add criteria to it. They suggested 860 additional items and 429 modifications to the original criteria. They were also asked to send in any existing unpublished checklists they might have, and several were obtained containing 135 more items. As the criteria were edited and condensed, special effort was made to incorporate all the content and much of the specific wording of the reviewers' suggestions. The resulting checklist had 246 criteria divided into 16 categories.

In the second part of the study, the revised checklist was again circulated to the group of reviewers. This time they were asked to rate each criterion in terms of the weighting it should receive when reviewing an article, and they were asked about their background and experience. The information was used to develop the final version of the checklist. First, approximately 9 percent of the criteria were eliminated because they were relatively unimportant (e.g., received low ratings) or were ambiguous (e.g., many ratings left blank). Second, the checklist was simplified by grouping the criteria within each category into clusters of similar criteria. Finally, within each cluster, the criteria were listed in a very gross rank ordering of importance based on the ratings. The final checklist contained 223 criteria, 44 clusters, and 15 categories.

The participating reviewers consisted of the editorial board and ad hoc reviewers for *Personnel Psychology*, and the editorial boards of *Journal of Applied Psychology* and

Academy of Management Journal. A total of 156 reviewers participated in the first part of the study (65 percent response rate), and 227 in the second part (66 percent response rate).

Several obvious limitations of this checklist should be noted. First, these criteria are self-report and may reflect social desirability and not actual reviewer behavior. In that sense, these data are more prescriptive than descriptive. Second, the most important caveat is that the checklist is not meant to replace reviewer judgment in any way, but only to provide a memory aid to remind reviewers of some potentially important criteria to consider. By analogy, it is like a preflight checklist for a pilot. It is not meant to replace flying skill, but only to remind pilots not to forget anything. Furthermore, the article review checklist is not meant to be applied in some mechanical fashion. Reviewers should not lose sight of the "big picture" when judging an article. Third, these criteria have not been validated against any external measures of scientific quality or contribution.

Several possible uses could be made of the checklist. As noted, reviewers might use it as a memory aid in reviewing manuscripts. In this regard, it is obviously unrealistic to expect articles to meet all the criteria, and only some criteria may be applicable to any given article. It may be especially useful to new and inexperienced reviewers, but more seasoned reviewers might also appreciate the extensive listing of criteria in order to reduce the memory demands of conducting reviews in an increasingly complex science. Authors might also find the checklist useful for evaluating planned research studies. In that role, it could be used to make improvements in the studies before they are conducted. It could also be used to evaluate the articles before submission in order to determine whether all the important topics are addressed, and it can be used at this stage to help anticipate possible criticisms and bolster the article accordingly. Finally, the checklist might be useful for training graduate students by helping them learn how to critically evaluate research.

Note

Special thanks to the reviewers who not only provided the ideas and data for this checklist, but who also labor tirelessly with little recognition to make the review process work.

Article Review Checklist

A. Importance of Topic

1. Theoretical importance.
   - Is theoretically important.
   - Can take field in new direction or change future research.
   - Justifies claims of importance on valid and clearly stated assumptions.
2. Practical importance.
   - Is practically important.
   - Links theory and practice in an important way.
3. Appropriateness.
   - Is appropriate to the journal and readership.
   - Is a new, emerging, or under-researched topic.
   - Is timely in terms of current trends in the field.

B. Literature Review

1. Linkage to most important literature.
   - References key (i.e., highly relevant) previous studies.
   - Considers recent literature.
   - Recognizes all relevant and important areas of literature.
2. Framing within the literature.
   - Uses literature to develop the arguments (i.e., not just a review).
   - Fits the study into the logical development of the research area.
   - Justifies direction taken by the study.

3. Thoroughness and accuracy.
   - Demonstrates understanding of the literature.
   - Draws proper inferences from previous studies, without overstating, misinterpreting, misapplying, or selectively reporting.
   - Identifies the major issues and themes in the literature that are relevant to the article.
   - Reviews literature critically, pointing out limitations, conflicts, and ambiguities in a fair manner (i.e., not too harsh or lenient).
   - Organizes literature properly to facilitate review.
   - References properly (e.g., recognizes seminal and definitive works, recognizes original research rather than over reliance on reviews and textbooks, minimizes non-scholarly citations, etc.).
   - Avoids tangents, marginally relevant citations, exhaustive listings of literature if not needed, and excessive self-citations.
   - Integrates multiple literatures when they are used.
   - Educates unfamiliar readers enough to evaluate the subsequent research.
   - Considers wide range of areas of literature.

C. Conceptual Development

1. Adequacy of scope and complexity.
   - Uses correct levels/units of analysis (e.g., behavior, person, job, group, organization, etc.).
   - Focuses on most critical variables (i.e., those known to be potentially influential), and explains rationale for inclusion and exclusion of variables.
   - Specifies relationships among variables clearly (including importance, direction, and size), in multivariate terms where needed (e.g., addition, interaction, all else equal, etc.), with special clarity regarding complicated relationships (e.g., form of interactions, mediation/moderation, causal models, etc.).
   - Has falsifiable hypotheses.
   - Has appropriate depth of conceptual development.
   - States antecedents and consequences of each focal construct clearly, and directions of causation, if relevant to research purpose.
   - Has hypotheses or research questions that are appropriate to level of knowledge and state of research on the topic (e.g., exploratory versus specific/complex).
   - Considers all relevant theories, or range of theories, and uses or acknowledges competing theories as necessary.
   - Explains processes underlying the constructs.
   - Specifies boundary conditions or limits of the theory or conceptual domain (e.g., in terms of units, context, and time).
   - Does not force a theoretical framework when the study is essentially exploratory.
2. Clarity and logical coherence.
   - Defines constructs/variables clearly and differentiates them from similar constructs/variables.

- Uses theory and arguments that are internally consistent.
- Uses clear and logical conceptual and theoretical development, leading from literature review to hypotheses or theses.
- States purposes, hypotheses, research questions, and intended unique contribution clearly.
- Reaches logical and clear deductions about the theory or conceptual development.
- States assumptions clearly and justifies them based on logic or evidence.
- Explains basic ideas and arguments clearly enough to be grasped by those outside the immediate topic area.

D. Additional Criteria for Literature Reviews and Conceptual Papers

1. Thoroughness.
   - Uses suitable approaches to analyzing, synthesizing, integrating, and evaluating the studies.
   - Summarizes a large and diverse literature, including all the information in the domain of interest.
   - Pulls together diverse findings from literatures that would be unfamiliar to researchers, yet pertinent to the topic.
   - Defines the domain and rules for including and excluding articles clearly and justifiably.
2. Uniqueness and incremental value.
   - Goes beyond previous reviews in the area.
   - Provides new insight, calls attention to a new problem, suggests new solutions, or otherwise adds value to current thinking.
   - Analyzes the literature critically (e.g., methods, findings, contradictions, etc.) and suggests improvements for future studies.
   - Goes beyond simply applying theory, and instead improves theory in some manner.
   - Organizes and explains previous findings, including anomalous findings and differences across studies.
   - Develops propositions, hypotheses, or questions for future research (i.e., tries to influence future research in some way).
   - Reframes the problem with, and integrates diverse issues into, a single theoretical framework.
   - Has appropriate timing (e.g., sufficient studies accumulated to be reviewed, sufficient time since last review, etc.).

E. Sample and Setting

1. Appropriateness.
   - Uses a sample (e.g., people, jobs, etc.) that is appropriate for the research question and adequately generalizable.

- Uses a setting (e.g., lab, field, archival, etc.) that is appropriate for the research question and adequately generalizable.
- Uses a context (e.g., situation, job, organization, etc.) that is appropriate for the research question and adequately generalizable.

2. Justifications.
   - Uses acceptable sampling strategy (e.g., random, representative, convenience, etc.) of people, jobs, or other important units of study.
   - Recognizes proper units of analysis (including nesting) and considers multiple units of analysis if needed.
   - Has adequate statistical power, reports power analyses, and interprets non-significant results accordingly.
   - Justifies the generalizability of student samples when used.
   - Considers timing of the study relevant to events which could influence results.
3. Sufficiency of description.
   - Explains sampling strategy clearly enough to determine degree to which statistical inferences can be made.
   - Has acceptable return rates and attrition rates (e.g., has adequate efforts to increase return rates, addresses the influence of non-respondents and drop-outs, has no obvious biases, etc.).
   - Has acceptable explanations for loss of sample, differing sample sizes, and so forth.
   - Describes population and sampling plan and size clearly, and population parameters appear likely to be accurately estimated.
   - Has adequately detailed demographics. Compares to known populations, previous studies, and theories if possible.

F. Measurement

1. Operationalization.
   - Operationalizes constructs correctly (e.g., consistent with literature, theory, or conceptualization). Defines constructs and theory well enough so this judgment can be made.
   - Justifies all measures based on purpose, theory, or previous research, and measures all critical variables.
2. Reliability.
   - Has adequate types and levels of reliability (e.g., internal consistency, inter-rater, test-retest, alternative forms, etc.).
   - Avoids inappropriate single-item measures.
   - Considers agreement (i.e., absolute level differences) as well as reliability (i.e., covariation) as needed.
3. Validity.
   - Avoids obvious criterion contamination, or assesses contamination adequately.
   - Uses measures that are free from bias (e.g., halo, social desirability, knowledge of predictor, etc.), are non-reactive, are likely to be accurate (e.g., asks questions respondents can answer), and have adequate range and variation.

- Avoids obvious criterion deficiency (e.g., samples content domain fully, uses multiple measures, uses proper criterion development procedures, etc.), or assesses deficiency adequately.
- Presents evidence of construct validity (e.g., convergent and discriminant validity) as needed.
- Uses multiple measures and sources if possible.
- Has adequate independence between measures.
- Addresses dimensionality of measures properly in development or analysis.

4. Availability.
   - Uses standardized, readily available, and well-researched instruments, when available. Explains fully when existing, accepted measures are not used.
   - Includes new measures or examples of new measures in text or appendix, and provides references for measures available elsewhere.
   - Uses existing organization-collected measures (e.g., turnover, absenteeism, performance, etc.) as needed, and explains and evaluates them fully.
5. Procedural adequacy.
   - Distinguishes clearly between measuring perceptions and intentions versus actual behaviors and outcomes.
   - Addresses levels of analysis and issues of aggregation correctly (and avoids ecological fallacy).
   - Forms scales correctly (e.g., weighting scheme logical) and describes them fully.
   - Uses adequate scaling and anchoring methodology (e.g., Likert, Thurstone, behaviorally anchored, etc.).
   - Uses highest level of measurement reasonably possible (e.g., nominal, ordinal, interval, or ratio).

G. Design – Experimental and Quasi-Experimental

1. Appropriateness.
   - Uses a high quality experimental design considering the constraints of the topic and setting.
   - Examines questions that are amenable to experimental/quasi-experimental research.
   - Uses adequate experimental task when needed given the topic, conceptual development, sample, and setting.
2. Proper controls.
   - Has appropriate control or comparison groups.
   - Uses truly random assignment procedures and explains them fully, or presents adequate evidence for the comparability of comparison groups.
   - Uses counterbalancing and statistical controls as needed.
3. Valid manipulations.
   - Operationalizes the construct manipulations or intervention correctly given the literature, theory, or conceptualization.

- Avoids obvious artifacts or biases (e.g., demand effects, experimenter expectancy, reactivity, evaluation apprehension, etc.).
- Avoids or minimizes confounding of extraneous variables with the independent variable manipulation.
- Has adequately strong manipulations or interventions, and has equivalence between conditions when needed.
- Includes manipulation checks when needed.
- Has realistic levels of factors in terms of populations and settings to which inferences are to be made, including multiple levels if possible to understand the form of the effect.
- Considers and includes important situational and contextual factors.

4. Threat avoidance.
   - Minimizes and addresses threats to internal validity (e.g., history, instrumentation, testing, maturation, selection, regression, mortality, directionality, confounding, etc.).
   - Minimizes and addresses threats to statistical conclusion validity (e.g., see analyses items, plus reliability of treatment implementation, random irrelevancies in the experimental setting, heterogeneity of respondents, etc.).
   - Minimizes and addresses threats to construct validity (e.g., see other design items, plus construct under-representation or confusion, insufficient definition, mono-operation bias, confounding constructs with levels of constructs, etc.).
   - Minimizes and addresses threats to external validity (e.g., see sampling items).
   - Makes appropriate trade-offs between types of validity (and between rigor and relevance) given the state of the research on the topic and the purpose of the study.
   - Explains, in lab studies, how key dimensions of the phenomenon or process under investigation can be adequately simulated in an artificial environment.

H. Design – Non-experimental and Cross-Sectional

1. Appropriateness.
   - Uses a high quality non-experimental design given the constraints of the topic and setting.
   - Examines questions that are amenable to cross-sectional or other non-experimental research (e.g., tests differential predictions and alternative explanations rather than a generalized null hypothesis, examines generalizability of previous experimental research, examines topics that cannot be examined experimentally, etc.).
   - Has logical implied directions of causation that are theoretically realistic in light of previous findings and theory and are assessed with adequate statistical procedures.
   - Avoids common method variance (i.e., mono-method or percept-percept bias), or explains why it is not a likely counter explanation for results.
2. Threat avoidance.
   - Includes needed control variables.

- Uses logical timing of measurement, especially regarding longitudinal designs.
- Identifies and addresses the influence of restriction of range, unreliability, and other statistical factors on results.
- Emphasizes degree of proof of causation correctly, and avoids hidden causal language.
- Addresses and assesses multicollinearity when needed.
- Addresses model misspecification (e.g., missing variables) when needed.
- Assesses nonlinearity when needed.
- Uses multiple research designs if possible.

I. Design – Meta-Analysis

1. Adequacy of sample of studies.
   - Includes all reasonably available relevant studies (both published and unpublished) in the domain of interest, and addresses the "file drawer" problem.
   - Cumulates a sufficient number of studies to justify a meta-analysis, and avoids second-order sampling as a major limitation.
2. Procedural adequacy.
   - Uses technically correct analytic procedures.
   - Explains and justifies rules for including and excluding studies.
   - Explains and justifies the coding of study variables.
   - Includes a list of the studies examined or makes it available.
   - Explains and justifies the methods of finding studies.
   - Aggregates adequately similar measures or constructs (i.e., variables have similar construct validity).
   - Uses multiple coders, and reports acceptable reliability.
3. Incremental value.
   - Goes beyond simply summarizing the data, but also contributes in some other important manner (e.g., theory, practice, methodology, etc.; see additional criteria for literature reviews and contribution).
   - Explores moderators fully.

J. Design – Qualitative

1. Procedural adequacy.
   - Defines the problem or questions to be addressed by the data.
   - Executes the methods and techniques properly.
   - Examines questions that are amenable to qualitative research (e.g., new topic area, initial stages of research, theory development, alternative methodology, fresh approach to old problem, etc.).

- Uses qualitative methods that are of high quality for the topic, setting, and purpose of the study (e.g., observation, interview, etc.).
- Conducts content analyses correctly, and describes them clearly.
- Describes procedural details fully, such that replication is possible.
- Justifies sampling frame (e.g., persons, observations, time periods, etc.) sufficiently for study purposes.
- Considers advantages and disadvantages of sample and setting.

2. Appropriateness of conclusions.
   - Develops and defines conceptual categories fully.
   - Relates conclusions to the problem or question, and to the methods used.
   - Develops appropriate theory or conceptual model from the data, and data supports the emergence of the theory or model.
   - Specifies and explains linkages among concepts or conceptual categories.
   - Considers important contextual factors and other explanatory conditions.
   - Describes process of change in the phenomenon if needed.
   - (For quantitative studies). Uses some qualitative procedures and data as needed to increase accuracy of measurement, support causal inferences, or otherwise help interpret the data (e.g., uses subject matter experts, qualitative pilot studies, focus groups, or interviews for planning or data interpretation, etc.).

K. Procedures

1. Quality.
   - Uses instructions to participants that are unlikely to improperly influence results.
   - Uses procedures in lab studies that are involving and have enough impact to be realistic.
   - Avoids procedures for data collection in field studies that are so intrusive that there is a risk of changing the phenomenon under examination or creating Hawthorne effects.
   - Follows ethical standards for the use of human subjects (e.g., informed consent, debriefing, etc.).
   - Conducts pilot tests where appropriate.
2. Adequacy of description.
   - Explains procedures clearly and in adequate detail (enough to allow a replication), yet is reasonably succinct.
   - Includes description of selecting or soliciting participants, specific instructions to participants, and efforts to ensure standardization.
   - Describes special conditions clearly which might compromise legitimacy of the results (e.g., relationship between investigator and organization that might reduce objectivity, study originally designed for another purpose or part of another study that might affect interpretation of results, etc.).
   - Compares procedures with those of other studies when needed.

- Describes procedural problems and solutions that might be useful to other investigators.

L. Data Analysis and Results

1. Appropriateness of statistics.
   - Uses analyses that are correct for the research questions or hypotheses, research design, and measures.
   - Reports both descriptive and inferential statistics.
   - Uses both univariate and multivariate statistics as needed.
   - Does not overlook simpler or more sophisticated methods that are more appropriate.
   - Includes basic statistics needed for future reviews and meta-analyses (e.g., means, standard deviations, reliabilities, intercorrelations, etc.).
2. Warranted assumptions and appropriate error rates.
   - Demonstrates awareness of major assumptions (e.g., level of measurement, independence of observations, homoscedasticity, fixed effects, etc.), avoids violating major assumptions or assesses degree of violation, or uses statistical procedures that minimize effect.
   - Uses significance levels that balance Type I and II errors, limits the number of levels used (e.g., to two), and applies them consistently.
   - Controls experiment-wise error rate (e.g., adequate overall test or post hoc procedure).
   - Uses correct data manipulations and transformations.
   - Avoids the apparent selective reporting of data dredging.
   - Avoids or assesses capitalization on chance (e.g., through cross-validation or shrinkage formulas), and has an adequate ratio of sample to variables.
3. Completeness.
   - Reports and discusses effect sizes.
   - Reports confidence intervals and significance levels as needed.
   - Does not report redundant or tangential analyses.
   - Reports analyses and statistics unambiguously and consistently, especially novel or sophisticated techniques. Gives additional explanation and justification as needed, including references.
   - Takes steps to protect the integrity of the data (e.g., quality control over collection and inputting), and examines outliers as needed.
   - Conducts obvious supplemental analyses suggested by the study.
   - Uses tables and figures correctly to help clearly communicate results. Uses tables and figures to complement, but not repeat, text.
   - Describes analyses in a logical sequence (e.g., descriptive statistics and manipulation checks first, followed by tests of primary hypotheses in order, followed by supplemental analyses, etc.).
   - Explores alternative explanations of the findings when possible.

- Shows consistency across analytic details (e.g., correct degrees of freedom, logical interrelationships among statistics, etc.).

M. Discussion and Conclusions

1. Explanation of results.
   - Makes correct inferences from research design and data analyses.
   - Links findings back to original hypotheses and purposes of the study.
   - Does not over-interpret or under-interpret data and results.
   - Does not simply rehash results, but interprets them in a meaningful manner.
   - Separates objective results description from subjective interpretation.
   - Summarizes results briefly.
   - Minimizes discussion of tangential topics or issues.
2. Derivation of implications.
   - Takes proper care in extrapolating from operationalized variables back to constructs.
   - Derives specific theoretical implications.
   - Derives specific practical implications.
   - Relates findings to those of other studies.
   - Places results in larger context of relevant issues where necessary.
   - Provides logical and innovative directions for future research.
3. Description of limitations.
   - Does not overlook or minimize findings contrary to hypotheses.
   - Identifies and addresses alternative explanations for results.
   - Provides a realistic (and adequately self-critical) delineation of limitations and weaknesses.
   - Considers both content and methodological explanations of results.
   - Identifies known or suspected boundary conditions or limits on generalizability.
   - Considers simplest explanations of the results.
   - Explains serendipitous findings as such.

N. Presentation

1. Quality of writing.
   - Presents analyses clearly.
   - Uses ideas in a logical and orderly fashion, and links the parts of the article together.
   - Writes well and readably (e.g., simple sentences, active voice, proper grammar, jargon and acronyms minimized, consistent terminology, parallel style, etc.).
   - Is well organized and correct content in each part of article.
   - Is objective, impartial, and professional.

- Explains importance of topic explicitly and introduces it early.
- Is succinct and parsimonious.
- Writes well-crafted and thorough pieces (e.g., attention to fine details and to broad patterns of integration).
- Tells an integrated story that is complete (i.e., addresses all obvious questions) and flows from beginning to end of the article.
- Frames writing in as interesting a manner as possible.

2. Conformance with publication guidelines.
   - Has length commensurate with the contribution.
   - Makes title and abstract adequate summaries of main content and contributions of the paper.
   - Presents all citations correctly in the reference list.
   - Follows journal style and format requirements.

O. Contribution

1. Overall contribution.
   - Makes a theoretical contribution (e.g., advances or challenges, not just applies, theory).
   - Makes a practical contribution (e.g., derives findings not already commonly accepted by practitioners, evaluates a common practice, etc.). Includes considerations of utility, organizational effectiveness, employee welfare, policy implications, and so forth.
   - Makes a methodological contribution (e.g., evaluates or proposes a new instrument, research strategy, analytical technique, etc.).
   - Provides a constructive replication (e.g., replicates but extends in an important way).
2. Increment to the current literature.
   - Fills gaps in current knowledge.
   - Goes beyond previous literature in the area.
   - Contributes in nontrivial or nonobvious way.
   - Stimulates potential future research.
3. Creativity and scope.
   - Addresses the "so what?" question.
   - Is innovative and creative.
   - Reports large amount of data and ideas not already reported elsewhere (e.g., avoids slicing the data, serial reporting, etc.).
   - Reflects an adequate magnitude or scope of research project.
4. Publication potential.
   - Is likely to improve contribution substantially with revision of article.
   - Has strengths in some parts of the study that offset weaknesses in other parts.