

Educational Researcher

<http://er.aera.net>

Standards for Reporting on Empirical Social Science Research in AERA Publications: American Educational Research Association

EDUCATIONAL RESEARCHER 2006; 35; 33

DOI: 10.3102/0013189X035006033

The online version of this article can be found at:

<http://edr.sagepub.com>

Published on behalf of



<http://www.aera.net>

By



<http://www.sagepublications.com>

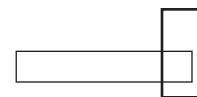
Additional services and information for *Educational Researcher* can be found at:

Email Alerts: <http://er.aera.net/cgi/alerts>

Subscriptions: <http://er.aera.net/subscriptions>

Reprints: <http://www.aera.net/reprints>

Permissions: <http://www.aera.net/permissions>



Standards for Reporting on Empirical Social Science Research in AERA Publications

American Educational Research Association

The American Educational Research Association (AERA) is pleased to provide guidelines for reporting on empirical social science research in AERA publications. These guidelines apply to reports of education research grounded in the empirical traditions of the social sciences. They cover, but are not limited to, what are commonly called qualitative and quantitative methods. Other forms of scholarship equally important to education research include reviews of research; theoretical, conceptual, or methodological essays; critiques of research traditions and practices; and scholarship more grounded in the humanities (e.g., history, philosophy, literary analysis, arts-based inquiry). The latter forms of scholarship are beyond the scope of this document.

The aim of specifying reporting standards for empirical research in education is to assist researchers in the preparation of manuscripts that report such work, editors and reviewers in the consideration of these manuscripts for publication, and readers in learning from and building upon such publications. The primary audience for these standards is researchers who wish to publish reports of empirical research and who review such research for AERA publications.

In adopting these standards, AERA emphasizes that the standards are intended to provide a framework of expectations, or rules of thumb, about what a report of empirical work ordinarily should address. The standards are not intended to define the conduct of empirical research. Although research reporting and research conduct are necessarily related, decisions about how to conduct empirical research are the researcher's responsibility. The standards are also not intended to define or determine the format for writing empirical work. Reports of different kinds of empirical research can take different forms, and authors working in different methodological or intellectual traditions may vary in the modes, order, or form of presentation. Finally, the acceptability

These reporting standards were adopted by the Council of the American Educational Research Association in June 2006. The Council extends its thanks to the AERA Task Force on Reporting of Research Methods in AERA Publications for quality work and breadth of consultation in the preparation of draft standards for Council consideration. The Task Force Members were Richard P. Duran, Margaret A. Eisenhart, Frederick D. Erickson, Carl A. Grant, Judith L. Green, Larry V. Hedges, Felice J. Levine (ex officio), Pamela A. Moss (Chair), James W. Pellegrino, and Barbara L. Schneider.

of a research report does not rest on evidence of literal satisfaction of every standard in this document, and acceptability cannot be determined by using a checklist. In a given case, there may be a sound professional reason why a particular standard is inapplicable. The purpose in specifying these standards is to provide guidance about the kinds of information essential to understanding both the nature of the research and the importance of the results.

While these standards are directed to authors, editors, reviewers, and readers of AERA journals, the substance of the standards and the breadth of methodological coverage are not particular to education research. Thus, in publishing these standards, the Association seeks to offer an educational document that can be useful to other research societies and journal publishers that disseminate empirical work using these same social science methods. Also, as part of AERA's broader educational mission to advance high-quality research in education and to foster excellence in reporting on empirical research, the Association commends use of these standards in the training and preparation of researchers in publishing research.

Preamble to Standards

Two overarching principles underlie the development of these reporting standards: the sufficiency of the warrants and the transparency of the report. First, reports of empirical research should be *warranted*; that is, adequate evidence should be provided to justify the results and conclusions. Second, reports of empirical research should be *transparent*; that is, reporting should make explicit the logic of inquiry and activities that led from the development of the initial interest, topic, problem, or research question; through the definition, collection, and analysis of data or empirical evidence; to the articulated outcomes of the study. Reporting that takes these principles into account permits scholars to understand one another's work, prepares that work for public scrutiny, and enables others to use that work. These standards are therefore intended to promote empirical research reporting that is warranted and transparent.

The reporting standards are divided into eight general areas: problem formulation; design and logic of the study; sources of evidence; measurement and classification; analysis and interpretation; generalization; ethics in reporting; and title, abstract, and headings. Each of these areas is considered in detail in the sections that follow. Each section starts with a general discussion of that area followed by specific numbered standards that pertain to that domain. In some instances, the discussion provides additional detail to clarify the meaning of the standards and how they might apply to different research traditions.

1. Problem Formulation

A research problem is an issue, topic, or question that motivates a study. Such problems may be theoretical, practical, or a combination thereof. The problem formulation answers the question of why the results of the investigation would be of interest to the research community and how an investigation is linked to prior knowledge and research.

Problem formulation can vary in scope and inclusiveness of questions and issues. Authors need to make clear how their formulation defines the limits of what can be addressed and the extent to which it is inclusive of diverse populations or circumstances. A thorough formulation of the problem typically includes a clear statement of the topic, issue, or question; a review of what others have written that bears directly on the problem; a rationale for the conceptual, methodological, and theoretical choices made in addressing the problem; and a consideration of how the study contributes to knowledge or understanding about the problem. These choices can have a significant influence on how a problem is understood, what generalizations can be made, and the extent to which a work can contribute to addressing significant issues. Reporting needs to provide as comprehensive a picture as possible of what the problem is about and how it has been approached.

- 1.1. The *problem formulation* should provide a clear statement of the purpose and scope of the study. It should describe the question, problem, or issue the study addresses, situate it in context, describe the approach taken to addressing it, and explain why it is important to address.
- 1.2. Reporting should make clear how the study is a *contribution to knowledge*.
 - 1.2.a. If the study is a contribution to an established line of theory and empirical research, it should make clear what the contributions are and how the study contributes to testing, elaborating, or enriching that theoretical perspective.
 - 1.2.b. If a study is intended to establish a new line of theory, it should make clear what that new theory is, how it relates to existing theories and evidence, why the new theory is needed, and the intended scope of its application.
 - 1.2.c. If the study is motivated by practical concerns, it should make clear what those concerns are, why they are important, and how this investigation can address those concerns.
 - 1.2.d. If the study is motivated by lack of information about a problem or issue, the problem formulation should make clear what information is lacking, why it is important, and how this investigation will address the need for information.
- 1.3. Reporting should include a *review of the relevant scholarship* that bears directly on the topic of the report. It should include a clear statement of the criteria used to identify and select the relevant scholarship in which the study is grounded. A review should make clear how the study contributes to, challenges, and/or extends theory, practice, methodology, research results, knowledge and/or understandings within an arena of inquiry.
- 1.4. The rationale for the *conceptual, methodological, or theoretical orientation* of the study should be described

and explained with relevant citations to what others have written about it.

- 1.5. A rationale should be provided for the *problem formulation as it relates to the groups studied* (especially with respect to relevant features of the historical, linguistic, social, and cultural backgrounds of the group) where questions about appropriateness of the connections may arise.

2. Design and Logic

The design and logic of a study flows directly from the problem formulation. It is shaped by the intellectual tradition(s) in which the authors are working and the ways in which they view the phenomenon under study. This in turn influences the identification of questions, the choice of methods of data collection, the approach to analysis and interpretation, and the format of reporting. These decisions constitute the logic of inquiry that researchers report.

Many different study designs are used in education research and different designs are appropriate for different problems. Some studies are concerned with drawing causal inferences regarding intended treatments, such as estimating the effects of reducing class size on student achievement. Other studies are concerned with describing particular occurrences and the meanings people give them in a single setting, such as how early literacy is taught and understood in a particular preschool. Some studies describe how a given phenomenon changes over time, such as how the number of women enrolled in doctoral programs in the physical sciences changes across several decades. Other studies examine the subjective experience of a particular group of research participants, such as what it is like, on a daily basis, to be a student in a “last chance” algebra class in a comprehensive high school. Still other studies examine the multiple layers that support or constrain the opportunities afforded to students and teachers, the distribution of resources within a school district, or the symbols and language of reform discourse. These examples illustrate only some of the range of empirical research problems in education research.

Designs for studies with such differing aims require very different approaches to data collection and analysis, from the construction of large-scale experiments or sample surveys to in-depth interviewing of a small number of subjects or the preparation of detailed transcripts of naturally occurring classroom talk. In some situations and in some forms of research, aspects of designs can be well specified in advance; in other circumstances, these may evolve, and the initial research questions become more elaborate or focused, as researchers become more familiar with the contexts in which they work. Moreover, research designs often take account of constraints; some designs are more feasible or practical in certain research situations than others, independent of their suitability in the abstract.

Whatever the study’s central purpose and circumstances might be, description of its design needs to make clear its logic of inquiry, showing how and why the methods and procedures that were used were appropriate for the problem as formulated. It is important as well that significant changes or developments in the design be clearly described, that reasons for changes be provided, and that any substantial implications for interpretation of the results be discussed.

- 2.1. Research reporting should follow a *clear logic of inquiry* that allows readers to trace the path from the initial state-

ment of the problem, issue, or interest; to the review of the relevant scholarship and intellectual orientation of the study; to the research questions initiated and/or developed in the study; to the description of the site, group, and/or participants (demographic information); to the methodology guiding collection and analysis of evidence; to the interpretation and presentation of outcomes and understandings gained from the research process. There should be a coherent presentation of these aspects of the study, and it should be clear how the different parts of the study are related to each other.

- 2.2. There should be a specific and unambiguous *description of the design*—the way the data collection or data identification activities were organized in the investigation. Significant developments or alterations in the research questions or design should be described and a rationale for the changes presented.

3. Sources of Evidence

“Sources of evidence” refers to both the units of study and the data or empirical materials collected or identified to address the research question or problem. Thus, reporting on sources of evidence includes describing relevant characteristics of the site, group, participants, events, or other units studied; the processes and judgments through which they were selected; and a rationale for these choices. It also includes specification of the data or empirical materials that were collected, the processes and judgments through which they were collected, and a rationale for these choices. Data sources typically include participant and nonparticipant observations; unstructured or semistructured interviews; documents and other artifacts; audio- or video-recordings; and standardized instruments like surveys or tests, structured interview protocols, and categorical demographic information that permit aggregation of data across cases or units of analysis. These data can be newly collected for a study or based on secondary sources of evidence. Since the role of the researcher and the relationship between the researcher and the participants can influence the data collected, this relationship is addressed in descriptions of sources of evidence.

- 3.1. The *units of study* (sites, groups, participants, events, or other units) *and the means through which they were selected* should be adequately described.

- 3.1.a. Descriptions should include relevant characteristics of the site, group, participants, events, or other units of study that bear directly on reporting and interpreting outcomes. The social, historical, or cultural context of the phenomena studied should also be described. The number of participants or other units of analysis (e.g., classrooms, schools) should be described unless circumstances make that impossible (e.g., some forms of observation in public places), and where relevant their relation to the more general population from which they were selected should be provided.

- 3.1.b. The means of selection of the sites, groups, participants, events, or other units of study should be described and a rationale provided. This includes the processes and judgments through which the units of study were sampled or selected, the agreements made with participants, and a rationale for these choices. Descriptions should include

how access, selection, and consent of participants were addressed; how rapport was established; what roles were taken by the researcher (e.g., interviewer, observer) and participants (e.g., respondent, informant) in the data collection; and what significant changes, if any, in relationships and roles of researcher and participants occurred over time. If in order to use certain forms of data or information consonant with guarantees of confidentiality, specific techniques have been used to mask or perturb the data or generate synthetic data from the original data, these processes should be noted.

- 3.1.c. Reporting on studies that compare groups as a central feature of their design (e.g., student participants in different community organizations, teacher turnover rates in rural, urban, and suburban schools) should describe those individuals, groups, or entities in sufficient detail to make the salient attributes, choices, and conceptual rationale clear. In laboratory or field studies, if the researcher has recruited participants and made any assignments to groups, the process, rationale, and outcome of assignment should be described.

- 3.1.d. When an intervention or treatment is implemented, the intervention or treatment should be described in sufficient detail so that its key features can be identified and used to account for results, and be compared with related interventions or treatments. Similarly, features of control or comparison groups should be described so that they can be understood and examined in relation to interventions or treatments.

- 3.2. The *collection of data or empirical materials* should be clearly described, including how and when they were gathered, by whom, and for what purposes. Description should also address salient processes and judgments that went into specification of data collection and a rationale for these choices. The description should be precise and sufficiently complete to enable another researcher, where appropriate, to understand what was done and, where appropriate, to replicate or reproduce the methods of data collection under the same or altered research circumstances. The relevance of evidence to the research problem, topic, or question should be clear. The sources and schedules of evidence may be specified in text, tables, or diagrams.

- 3.2.a. Descriptions should include information such as the time and duration of data collection; the schedule, context, and focus of data collection and how it was done (e.g., structured inventory, fieldnotes, audio- or video-recording); the protocol for the administration of any instruments; the documents, records, or artifacts gathered and the ways in which they were identified. Where secondary data sources were used, reference to where the original description can be found may be sufficient.

- 3.2.b. With structured or semistructured interviews, open-ended surveys, or observational inventories, there should be sufficient description of these guides or protocols to convey their properties; with open-ended or informal interviews, there should be sufficient information to place participants' responses in the context of what was asked and of what preceded it. Structured surveys; data

collection protocols; or standardized tests, measures, or instruments should similarly be described in sufficient detail to convey the development process and provide evidence of their technical quality. Information on access to these surveys, instruments, protocols, inventories, and guides should be specified. References should be included for instruments used in a reported study previously developed by the authors or by other investigators.

4. Measurement and Classification

Empirical studies typically entail some process of data selection, reduction, or translation to enable analysis and reporting of outcomes. Measurement is the process by which behavior or observation is converted into quantities, which may, in turn, then be subjected to some kind of quantitative analysis. Classification refers to processes of segmenting data into units of analysis and categorizing or coding them. With qualitative methods in particular, classification is often considered integral to the data analysis. Thus, it is addressed here and referred to again under data analysis standards. The validity of empirical studies depends, in part, on the claim that classifications and measurements preserve important characteristics of the phenomena they represent. The practices of classification and the development of measurement instruments are typically iterative, as researchers seek to provide representations or translations of the data that are most meaningful in light of the phenomena studied and the research questions addressed. Descriptions of the development of classifications and measurements, as well as evidence of their meaningfulness and appropriateness for capturing important characteristics of the groups or participants studied, are important elements of research reporting.

Empirical investigations often involve a large number of data elements, some of which are more important to the logic of inquiry and interpretation of the investigation than others. It is important to distinguish key data elements that are crucial to the logic and interpretation of the outcomes. Such elements will typically include those that are directly involved in the quantitative or qualitative analyses on which interpretations are based. They will also include those that are crucial to any intended extrapolations or generalizations of the results beyond the social phenomena studied.

- 4.1. The *development of measurements and classifications* should be clearly described, showing how the measurement or classification preserves important characteristics of the phenomena under study. When a previously developed measurement instrument or classification scheme is used, reference to a publication where these descriptions are provided may be sufficient.
- 4.2. Any *classification* scheme should be comprehensively described and illustrated with concrete examples that represent the range of phenomena classified.
 - 4.2.a. When the classification involves only parts of the data, the means through which those parts were selected should be described and a rationale provided.
 - 4.2.b. When exhaustive analysis of the relevant data is desirable and appropriate, especially when such analysis is necessary to support the main conclusions that are drawn (e.g., about the “typicality” of an event or the pervasiveness of a pattern), the classification scheme and frequencies of items in each classification should be presented in

a table, chart, or appendix, or the information on their availability should be otherwise provided by the author.

4.2.c. If coding processes are used, the description should include, as relevant, information on the backgrounds and training of the coders; inter-coder reliability or outcomes of reviews by other analysts; and, where relevant, indications of the extent to which those studied (participants) agree with the classifications.

- 4.3. When *measurement* is entailed, reporting should describe data elements and organization in a specific and unambiguous way.

4.3.a. Relevant descriptive statistics (such as means and standard deviations for continuous variables, frequencies for discrete variables with few categories, and correlation matrices) may be provided in tables if the analyses depend on having this information accessible; otherwise, they should be available from the author upon request.

4.3.b. If key data elements are derived from others, as with scales and composites, their derivation should be presented in a specific and unambiguous way. If these derived data elements are conventional (such as a well-known scale or a score on an established test), then a citation to an external reference is sufficient.

4.3.c. Sufficient detail should be provided to make clear that measures are being used appropriately, have suitable dependability (reliability) properties, and are interpreted properly for the groups studied. If the data were reduced or scales, scores, or measures were developed through data reduction techniques or statistical methods, the data reduction procedures should be fully described. Evidence of appropriate use, dependability, or valid interpretation of measures (particularly key measures) should be provided in circumstances where a knowledgeable scholar might reasonably have questions.

- 4.4. When *transcriptions of audio- or video-recordings* are provided, the conventions and symbols used to represent the discourse or characterize the actions or interactions should be clearly described and a rationale provided.
- 4.5. A rationale should be provided for the *relevance of a measurement or classification* as capturing important characteristics of the group studied (especially with respect to relevant features of the historical, linguistic, social, and cultural background of the group) where questions about appropriateness might readily arise.

5. Analysis and Interpretation

An important aspect of reporting is to provide evidence that the outcomes and conclusions are warranted and that disconfirming evidence, counter-examples, or viable alternative interpretations have been appropriately considered. This entails a clear statement of the process and outcomes of data analysis and a discussion of how they address the research questions or problem. Because the processes of analysis tend to follow somewhat different paths in quantitative and qualitative methods, specific standards are discussed for each, after discussion of the general standards. When reporting on multiple methods or research that is not easily classified as quantitative or qualitative, relevant standards from both sets need to be addressed.

In general:

- 5.1. The *procedures used for analysis* should be precisely and transparently described from the beginning of the study through presentation of the outcomes. Reporting should make clear how the analysis procedures address the research question or problem and lead to the outcomes reported. The relevance of the analysis procedures to the problem formulation should be made clear.
- 5.2. *Analytic techniques* should be described in sufficient detail to permit understanding of how the data were analyzed and the processes and assumptions underlying specific techniques (e.g., techniques used to undertake content analysis, discourse or text analysis, deliberation analysis, time use analysis, network analysis, or event history analysis).
- 5.3. The analysis and presentation of the outcomes of the analysis should make clear how they *support claims or conclusions* drawn in the research.
- 5.4. Analysis and interpretation should include information about any *intended or unintended circumstances* that may have significant implications for interpretation of the outcomes, limit their applicability, or compromise their validity. Such circumstances may include, but are not limited to, key actors leaving the site, changes in membership of the group, or withdrawal of access to any part of the study or to people in the study.
- 5.5. The *presentation of conclusions* should (a) provide a statement of how claims and interpretations address the research problem, question, or issue underlying the research; (b) show how the conclusions connect to support, elaborate, or challenge conclusions in earlier scholarship; and (c) emphasize the theoretical, practical, or methodological implications of the study.

With quantitative methods:

With quantitative methods, statistical analyses are typically undertaken and reported and then discussions of the results developed. The results of statistical analysis typically involve both a quantitative index of a relation between variables or a magnitude and an index of its uncertainty. While statistical significance testing has a long history and a useful place in education research, statisticians have long warned against overreliance on significance testing to the exclusion of other methods of interpreting statistical analyses. Statistical significance tests combine both magnitude of relations (or estimates) and their uncertainty into the same quantity. Interpretation of statistical analyses is enhanced by reporting magnitude of relations (e.g., effect sizes) and their uncertainty separately.

While many statistical analyses may be carried out in a study, typically only a subset is critical to the eventual results and interpretations. It is important to report the results of analyses that are critical for interpretation of findings in ways that capture the magnitude as well as the statistical significance of those results. Quantitative indices of effect magnitude (effect size indices) are a useful way to do this.

- 5.6. Reporting should clearly state *what statistical analyses were conducted and the appropriateness of the statistical tests*, linking them to the logic of design and analysis and describing them in enough detail that they

could be replicated by a competent data analyst. When central to the research, tests of rival hypotheses and alternative interpretations should be reported.

- 5.7. *Descriptive and inferential statistics* should be provided for each of the statistical analyses that is essential to the interpretation of the results.
- 5.8. Any *considerations that arose in the data collection and processing* (e.g., attrition, missing data, ceiling or floor effects, deviations from standard administration of instruments, suspected cheating) that might compromise the validity of the statistical analysis or inferences should be reported.
- 5.9. Any *considerations that are identified during the data analysis* (e.g., violations of assumptions of statistical procedures, failure of iterative statistical procedures to converge, changes in data analysis models necessitated by unexpected data patterns) that might compromise the validity of the statistical analyses or inferences should be reported.
- 5.10. For each of the *statistical results* that is critical to the logic of the design and analysis, there should be included:
 - An index of the quantitative relation between variables (an effect size of some kind such as a treatment effect, a regression coefficient, or an odds ratio) or, for studies that principally describe variables, an index of effect that describes the magnitude of the measured variable.
 - An indication of the uncertainty of that index of effect (such as a standard error or a confidence interval).
 - When hypothesis testing is used, the test statistic and its associated significance level.
 - A qualitative interpretation of the index of the effect that describes its meaningfulness in terms of the questions the study was intended to answer. This interpretation should include any qualifications that may be appropriate because of the uncertainty of the findings (e.g., the estimated effect is large enough to be educationally important but these data do not rule out the possibility that the true effect is actually quite small).

With qualitative methods:

With qualitative methods, analyses typically occur during as well as after data collection. Early analyses can help inform subsequent data collection by, for instance, identifying categories of events, actions, or people for further analysis within the ongoing study or for further study. As indicated in the Measurement and Classification section above, during the initial stages of analysis, researchers may develop ways of segmenting the data (e.g., by person; by action, activity, event, or narrative; by time period) and sets of substantive categories or codes into which segments of data can be organized. These classifications help the researcher identify patterns within the data. Patterns refer to configurations of events or other observations that occur repeatedly or consistently in a characteristic arrangement. Sometimes analysis is intended to provide comprehensive in-depth interpretations of a particular text, recording, or other artifact rather than pattern descriptions of extended or multi-component empirical materials. Whatever the approach to initial data analysis, it is important that

researchers fully characterize the processes they used so that others can trace their logic of inquiry.

Once initial classifications, pattern descriptions, or in-depth interpretations are developed, researchers may review the corpus of available data to locate all relevant instances to support the claims, to search for confirming and disconfirming evidence, and to try out alternative interpretations. They may also return to data collection if additional evidence is needed to confirm or disconfirm a pattern. This process results in an initial set of claims or interpretations which represent the preliminary conclusions or learnings from the research. The available sources of evidence may be re-reviewed, and alternative interpretations may be tried out, in the process of developing the final conclusions or learnings that will be reported.

This iterative process of developing claims or interpretations, seeking confirming and disconfirming evidence in the data, sometimes collecting additional evidence, and trying out alternative claims or interpretations, supports the development of warrants for claims or conclusions. Data analysis ceases when researchers are satisfied and can provide evidence that their interpretations meaningfully and comprehensively characterize the data analyzed in light of the problem formulation.

It is the researcher's responsibility to show the reader that the report can be trusted. This begins with the description of evidence and analysis supporting each claim described above. The warrant for the claims can be established through a variety of procedures including triangulation or comparison of evidence from different sources, asking participants to evaluate pattern descriptions and claims, having different analysts examine the same data (independently or collaboratively), searches for disconfirming evidence and counter-interpretations, and representations of differing perspectives among participants and researchers, including attention to their location in the broader social structure. When the evidence does not converge, differences should be noted. Critical examination of the pre-existing perspective, point of view, or standpoint of the researcher(s), of how these might have influenced the collection and analysis of evidence, and of how they were challenged during the course of data collection and analysis, is an important element in enhancing the warrant for each claim.

The following standards are intended to make the process of analysis transparent for reviewers and readers.

5.11. The *process of developing the descriptions, claims, and interpretations* should be clearly described and illustrated. The description should make it possible to follow the course of decisions about the pattern descriptions, claims, and interpretations from the beginning to the end of the analysis process. Sufficient detail should be included to make the process transparent and engender confidence that the results are warranted.

5.12. The *evidence that serves as a warrant for each claim* should be presented. The sources of evidence and the strength and variety of evidence supporting each claim should be described. Qualifications and conditions should be specified; significant counter-examples should be reported. Claims should be illustrated with concrete examples (e.g., fieldnote excerpts, interview quotes, or narrative vignettes) and descriptions of the

social context in which they occurred should be provided. If a warranted claim entails a generalizing statement (e.g., of typicality), it should be supported with evidence of its relative frequency. Speculations that go beyond the available evidence should be clearly represented as such.

5.13. *Practices used to develop and enhance the warrant for the claims* should be described, including the search for disconfirming evidence and alternative interpretations of the same evidence. Significant limitations due, for instance, to insufficient or conflicting evidence, should be described.

5.14. *Interpretive commentary* should provide a deeper understanding of the claims—how and why the patterns described may have occurred; the social, cultural, or historical contexts in which they occurred; how they relate to one another; how they relate to (support or challenge) theory and findings from previous research; and what alternative claims or counter-claims were considered.

6. Generalization

All investigations involve specific participants, take place in specific contexts, and involve specific activities, data collections, or manipulations. However, some investigations are intended to have implications beyond most, if not all, of the specifics occurring in the investigation itself. Where there is an intent to generalize beyond the specifics studied, it is incumbent on the author to indicate the individuals, contexts, activities, data collections, and so forth (the domains) to which the generalization is intended to apply and (at least implicitly) those to which it may not apply. It is also incumbent on the author to provide a justification for the generalization. To justify such generalizations, it is necessary to articulate both the details of the investigation itself and the logic by which the findings of the investigation should apply to the domains intended.

Sometimes the generalization intended is from a sample to a sampling frame (a population or universe). In this case, one logic that could support the generalization is embodied in the sampling theory of generalization. If the investigation is carried out using a probability sample drawn from the population, statistical theory provides guidelines about the uncertainty appropriate for generalizations to the population. Note that this logic involves very strong assumptions about the way in which the sample used in the investigation is chosen, but the validity of this logic depends very little on knowing the specific characteristics of the individuals in the investigation or in the population.

Sometimes the generalization intended is from a sample to a population, but the sample is *not* a probability sample from that population. In this case, an argument that is similar to the probability sampling argument is used, an argument that involves a claim that the sample is “representative,” a nontechnical term that usually means that the sample supports the same kind of generalization as a probability sample. Such claims of representativeness are typically bolstered by evidence that the sample is similar to the population in all important respects. Note that, to be persuasive, this logic requires more evidence about both the population and sample than is necessary to support claims of generalization from probability samples to populations.

Sometimes the generalization intended is from contexts or cases studied to unstudied contexts or cases. Generalizations from studied sites, situations, groups, or social processes require an explicit justification that features of the studied context are sufficiently similar to unstudied contexts to make plausible such inferences. In other instances, the intent is to generalize more broadly based on cases reported in the literature. This logic requires the researcher to draw explicit comparisons to published research that focuses on similar phenomena.

Sometimes an intended generalization is not specified by the author; rather, generalization to a new context is considered to be an empirical matter whose determination is made by the reader. With such studies, it is the obligation of the researcher to describe the phenomena in sufficient detail that readers of the article can draw appropriate comparisons to their own contexts of interest; in other words, it is the specificity of descriptive reporting that allows the reader to decide whether the processes identified in the reported case may also be found in an analogous situation known firsthand by the reader.

Whenever a claim of generalization beyond the reported case is specified or implied, it is the obligation of the researcher to build an explicit argument for that claim.

- 6.1. Whether generalization is intended by the author or not, it is crucial to make clear the *specifics of the participants, contexts, activities, data collections, and manipulations* involved in the study. This includes all of the specifics that are relevant either to the logic by which the study should apply to the generalizations or to permit readers to draw the necessary comparisons to their own contexts of interest.
- 6.2. When generalization is intended, the author should make clear the *intended scope of generalization* of the findings of the study. It may be helpful to delineate the situations (or domains) in which the findings of the investigation do *not* apply to identify the scope of intended generalization. If the primary generalization is to theory, reporting should make clear specifically how the findings falsify, support, extend, or elaborate the relevant line(s) of existing theory. If the primary generalization is to identifiable problems or practical issues, reporting should make clear the situations in which the findings have applications, implications, or practical consequences and why this is plausible.
- 6.3. Generalization that is intended by the author should make clear the *logic by which the findings of the study should apply within the intended scope of generalization*. The logic should provide a clear and persuasive rationale supporting the generalization from the study to the domain to which generalization is intended. The logic should also identify and present evidence that may be necessary to support the validity of the claims of generalizability (such as evidence that the individuals in the study resemble those in the domain of generalization in relevant respects).

7. Ethics in Reporting

AERA has developed and issued a set of ethical standards for the conduct of research to which its members and those who participate in all AERA programs, including publishing, are expected to adhere (see Ethical Standards of AERA at <http://www.aera.net/>

[aboutaera/?id=222](#)). It is assumed that authors seeking publication in AERA journals are familiar with and adhere to these ethical standards. This section describes only those ethical issues that are directly relevant to reporting research. Authors need to address these and any other issues they consider germane to the transparency and ethics of reporting.

In reporting research, authors have an opportunity and responsibility to address ethical decisions that shaped how the inquiry was designed or undertaken or how the empirical evidence or data were organized, maintained, or analyzed. These include, for example, key considerations with respect to consent (or a waiver of consent) or confidentiality agreements (including any agreement with participants to reveal their identities). Discussion of any incentives for participation that were provided and how they were managed would typically also be reported. Research reporting needs to be undertaken consonant with confidentiality guarantees and data protection plans. In cases where researchers may have eliminated or altered descriptions or used other disclosure limitation techniques to mask or perturb the data, these processes need to be noted. If a condition of access or other criteria led to decisions to mask the identity of locations, institutions, or other sites in data files and in data dissemination, these decisions would also be described.

Reporting of research is expected to reflect the highest standards of ethical practice both with respect to human participants and with respect to the execution of professional conduct and judgment in research. Reporting should avoid descriptions that underrepresent or distort differences within and among individuals and groups. Reporting must be accurate and without falsification or fabrication of data or results; reflect the work of the authors with appropriate attribution to others; be free of any plagiarism or misappropriation of the writing or ideas of others; and be sufficiently accessible to be subject to verification, replication, or further analysis. Any prior review of the manuscript by research participants, those providing access to sites, or those funding the research that could have limited the author's autonomy to publish the research or how it was reported would typically also be described. In addition, funding sources or other sources of support that may raise issues of conflict of interest should be noted.

- 7.1. Ethical considerations involved in *data collection, analysis, and reporting* should be explicitly addressed. Not all ethical issues in the conduct of the study or about human research protection need to be addressed in an article, but those relevant to understanding the study, analyses, and results should be set forth. Study approval in accordance with an Institutional Review Board (IRB) should be stated.
- 7.2. Reporting on research and findings should be presented in a way that honors consent *agreements* with human participants and any other agreements with respect to gaining access to research sites or data. Reporting includes all writing (e.g., text of the article, quotes, excerpts of interactions), pictures, maps, or graphical displays that could inadvertently compromise guarantees of anonymity of human participants and the confidentiality of information about them or conflict with other promises made as a condition of access (e.g., masking the identity of school

districts). Even when direct identifying information is not used, inadvertent disclosure of research participants can happen indirectly.

- 7.3. Reporting should include a description of any potential **conflicts of interest or biases** of the researcher that may have influenced or could have the appearance of influencing the research, along with a description of how they were managed in the conduct of the study.
- 7.4. Reporting of research should be **accurately stated** and attention should be given to ensure that there are no omissions or inclusions of information that are false or that fabricate, mislead, or misrepresent how the research or analyses were done.
- 7.5. The **data or empirical materials relevant to the conclusions** should be maintained in a way that a qualified researcher with a copy of the relevant data and description of the analysis procedures could reproduce the analysis or trace the trail of evidence leading to the author's conclusions.
- 7.6. **Funding support** should be acknowledged in a publication note. In special circumstances, where sponsors cannot be acknowledged by name, a description of the nature of the sponsor should be provided.

8. Title, Abstract, and Headings

A well-constructed title and abstract help readers in locating articles relevant to their interest. Since concepts in the title and abstract are typically indexed and searched electronically, it is important that the words be carefully chosen to convey the contents of the article. Using terms likely to be understood both within and beyond one's immediate research community helps to make the work accessible to a broad audience. Well-constructed headings help readers follow the logic of inquiry in an article.

- 8.1. The **title** should clearly convey what the article is about.
- 8.2. The **abstract** should provide a summary of the article that is self-contained, concise, and accurate. Preparation of the abstract should be in accordance with the format and structure required for AERA publications generally or for the specific AERA journal to which the article is submitted. Whenever feasible, it should set forth the research question or problem; descriptions of the research sites, objects, or participants; the conceptual orientation of the study; the methods used for collecting sources of evidence or data; procedures used for analyzing the evidence; and the main conclusions and implications.
- 8.3. **Headings** and subheadings should make clear the logic of inquiry underlying the report.