

## Writing the Methods

**John Coverdale, M.D., M.Ed., F.R.A.N.Z.C.P.**

**Laura Roberts, M.D., M.A., Alan Louie, M.D., Eugene Beresin, M.D.**

In earlier editorials, we have emphasized the role and responsibilities of reviewers for *Academic Psychiatry* (1, 2). The reviewer's task is to assess scientific rigor and scholarly integrity and to provide prompt, constructive, and comprehensive feedback to assist authors in optimizing their written product. Our intention here is to focus on authors' responsibilities in the writing of the Methods section in educational research reports. We focus on the Methods section because of its centrality to determining the validity of research findings.

The central position of this editorial is that authors are responsible for providing sufficient information about the conduct of research that allows accurate replication and an adequate judgment of scientific merit. To this end, authors are obligated to clearly and adequately communicate what was actually done and to delineate potentially biasing factors in the conduct of the research. Though it is not possible to conduct research free of bias, a detailed description of the Methods will facilitate the critical appraisal of the study and decision making about whether to incorporate the findings into educational practice. A detailed description of the Methods will also help those interested in replication of the study and permit modifications as needed to improve validity.

Some problems of relatively low severity, such as withholding certain details of methodology or results or using inadequate or inappropriate research designs, can nevertheless impair the integrity of science (3, 4). One survey of several thousand researchers supported by the National Institutes of Health found that a minority had engaged in such behavior, as determined by anonymous self-report (3). Such findings are not necessarily generalizable to educational research, and we assume that all of our contributors are committed to promoting the integrity of science. Authors should, however, remain vigilant in the conduct and communication of educational research.

---

### Study Design

It is worth appreciating that research originates with an issue or problem, and the nature of the problem and the

associated research question(s) or hypothesis drive the determination of study design. Authors should therefore explain the rationale of the study design in relation to the research question or hypothesis, including how the design was appropriate to the question. Because practical or financial issues can prevent the choice of the most optimal design, especially in an environment where educational research is not formally funded or is substantially underfunded (5), comment on the choice of design and its relative merits could be useful. A comprehensive discussion of the designs that are most suitable for different kinds of questions is beyond the scope of this column, and several excellent general resources are available (6, 7). Papers published on review criteria for components of the Methods section (research design; instrumentation; data collection and quality control; population and sample; data analysis and statistics) are also excellent resources (8–11).

---

### Quantitative Research

There are a number of educational research designs that incorporate randomization; the pre-test-post-test control group design and the related Solomon four group design are the most rigorous (5, 7). Randomization is a sampling method that serves to protect against the possibility of bias or systematic error in that all research subjects have an equal probability of belonging to the intervention or healthy comparison group(s). Thus, all known and unknown relevant variables are likely to be evenly distributed between groups. First, authors should describe the methods of randomization so that readers can judge their adequacy. In settings outside of educational research in psychiatry, providing such detail has not been usual (12). Because randomization does not ensure that groups will be identical on the dependent measures, especially when group sizes are small, the effectiveness of randomization will be judged by a comparison between groups on pre-test scores. In addition, the effectiveness of randomization will be judged by a between-group comparison of all possible

relevant variables. A table is normally provided for this purpose in the Results section.

Cohort designs and volunteer treatment designs that include control groups are common examples of quasi-experimental designs in educational research (7). Because of the lack of formal randomization, authors should describe the demographic and other possible determinants of outcome for each group in addition to the intervention or exposure of interest. These characteristics are also presented in a table along with a determination of which of the variables are significantly different between groups. This will assist readers in evaluating the potential for confounding, which is an important potential weakness of cohort studies in particular (13, 14). Nonexperimental designs, such as correlational research designs, benefit from some justification, based on theoretical considerations and previous research, of the selection of variables that are to be correlated with each other (6).

The following additional information should be provided: the particulars of the setting, possible contextual effects on the procedures, instructions provided to learners, the treatment of the sample(s), similarities and differences in the treatment of groups, and the data collection procedures. The sequence and timeline of procedures, such as whether randomization occurred before or after pre-testing, should be identified. Any discrepancies in the conduct of the study that might influence outcomes should be identified. For example, any variations within or between the treatment of groups should be identified when known and likely relevant. This concept of treatment fidelity has been defined as the extent to which the treatment conditions, as implemented, conform to the researchers' specifications for the treatment (6).

Authors should be attentive to how the Methods are organized and communicated. One way to order the Methods is to describe the research design first, including the specific treatment of groups in an experimental study and data collection procedures. This can be followed by a description of the sample and population from which the sample was drawn. The outcome measures, including their validity and reliability, and the data analysis procedures, including specific statistics used, conclude the Methods. Although flexibility in the ordering of Methods is allowed, we will follow this same general outline for this editorial. The American Psychological Association has published a manual for assisting authors in manuscript preparation that includes a section on how to organize the writing and how to communicate clearly and concisely (15). The limitations of the Methods are normally identified in the Dis-

cussion section so that the results and conclusions can be understood in their light.

---

### Qualitative Research

---

Qualitative research is done by learner (e.g., focus group) interviews, observations in naturalistic settings (e.g., ethnographic research), or historical research (e.g., analysis of documents) (16). Although qualitative research methods are diverse, the standards for judging adequacy are equally demanding (17–19). First, and as is the case for quantitative studies, authors should justify the appropriateness of the Methods in relation to the specific research question (17, 18). Aspects of research design that should be incorporated into the Methods include how the study participants were selected and the reasoning supporting their selection, the specific methods used to generate data, the comprehensiveness of data collection, and procedures for analyzing the data and corroborating the findings (16, 17). In qualitative research, sampling is purposive in that it aims to cover a range of potentially relevant social phenomena and perspectives from an appropriate array of data sources. Sound reasoning for describing and justifying the participant selection strategies should be provided (19). Because of the subjective nature of qualitative research, authors should identify possible biases or the steps that were taken to limit bias in the collection and interpretation of data. This might include investigators keeping personal records of thoughts and experiences in observational research (18) and how results that were contrary to what was anticipated were accommodated (17).

---

### Participants

---

First, the population from which the sample was drawn and the sample procedures used in the selection of the sample should be defined. Second, the numbers of participants recruited and selected and their relevant characteristics should be defined. Third, the completeness of follow up or number of dropouts or nonresponders should be identified. We refer authors to the CONSORT statement (20) which aims to reduce possible bias by increasing the quality of the reporting of trials. This statement serves equally well as a standard for the reporting of experimental studies in education. In providing this detail the representativeness of the sample or the “generalizability” of the observed results can be better appreciated.

## Outcome Measures

---

The outcome measures should be described or alternatively referenced. When these are constructed anew, authors should indicate how they were constructed, whether by focus groups or by modification or development from earlier available instruments. The measures should address the research question. The validity (the extent to which an instrument measures what it purports to measure) and reliability (the consistency with which an instrument measures what it measures) of measures should be identified when these are known (21). It is worth appreciating that rating scales that are unpublished can be a source of bias in that unpublished scales are associated with significant treatment effects (22).

Many of the empirical papers published in *Academic Psychiatry* are surveys. Issues of particular relevance to survey research include how the questions were written, the extent to which questions were closed or open-ended, and how statements were balanced in minimizing the possibility of response bias (23). Authors should indicate what steps, if any, were taken to determine the various types of validity and test-retest reliability. Alternatively, this lack of information should be identified as a limitation.

## Data Analysis Procedures

---

### Power

The power of a study refers to the ability of the methods to detect differences between experimental and comparison groups when these truly exist (6). Small sample sizes are not uncommon in educational research, especially when conducted in a single institution. When sample sizes are small, calculating the power of the study can assist investigators and readers in their appreciation of the probability of showing an effect of a certain size if it is truly there. It should also be appreciated that significant findings are more likely to occur by chance when multiple outcome measures are used. In this case, authors might adjust their level of significance in order to reduce the possibility of a significant finding when such a finding is not true (11).

### Statistics

Statistical tests should be discussed in light of the study question(s), study design, and the nature of sample(s) and measures used. In keeping with others (15), we adopt the view that the reader of *Academic Psychiatry* will have a professional knowledge of statistics so that basic functions in any manuscript need not be explained. If there are likely

questions about the appropriateness of a particular test, however, authors should justify its use. Key issues to be taken into account in the choice of a test concern the nature of the data (whether interval, ratio, nominal or ordinal), normality of distribution, and homogeneity of variance. Another key issue concerns the nature of samples, whether paired, dependent, or unpaired.

For one example, when comparing the significance of differences between two means, a t test is used when the data are intervals or ratios, are normally distributed, and when there is homogeneity of variance. When the samples are paired, a t test for paired samples is used; when unpaired, a t test for unpaired samples is used. Nonparametric equivalents include the Wilcoxon signed rank test and Mann-Whitney U test for paired and unpaired samples, respectively (6).

For another example, when testing the statistical significance of differences between three or more means, the type of analysis of variance used is dependent on both the nature of data and the research design. For example, an analysis of covariance (ANCOVA) is commonly used for pre-test-post-test research designs. The advantage that is conferred by an ANCOVA is that pre-test means can be used as a covariate and post-test means can be adjusted by taking the initial (pre-test) differences into account. Thus, post-test differences between the means of the experimental and comparison groups can be attributed to differences in the intervention, as opposed to initial differences between the means of the groups (6, 7). It should be appreciated, however, that pre-testing can limit the "generalizability" of study results. This is because the results may only apply to when learners are primed (6, 7). The post-test only design, therefore, has an advantage in contrast to the pre-test-post-test comparison group design in this regard, although the post-test only design confers less statistical power.

## Conclusions

---

*Academic Psychiatry* is committed to improving the standards of conduct and communication of educational research. As previously identified (24), we seek to publish articles that inform about educational research design, measures, and common methods of statistical analyses, and we invite your contributions. We are also interested in publishing articles that critically appraise the strengths and weaknesses of the available literature on focused educational questions.

At *Academic Psychiatry*, we also need to consider our

own editorial processes. Editorial policies and practices can make important contributions to increasing the quality of educational intervention research (25). We depend primarily on our own authors and reviewers to be assiduous in their appraisals of the adequacy of methods. We recognize that we frequently ask authors to shorten papers because of space constraints. Recommendations to shorten papers should not usually be understood to compromise the Methods in favor of other sections. We also welcome vigorous invited commentary and letter sections as a check on the adequacy of methods supporting the conclusions published in the *Journal*.

It also follows that researchers should take the requisite time in planning to prevent or limit major study design weaknesses (7, 25). Early stage review with design and measurement experts might prevent some problems. Statisticians often appreciate an opportunity to review survey questionnaires in advance of their distribution and to rectify possible flaws. Taking notes and attention to detail will facilitate disclosure of methodological concerns as these arise. All of these processes in turn should improve the overall quality of the scientific basis of academic psychiatry.

### References

1. Roberts LW: On the centrality of peer review. *Acad Psychiatry* 2002; 26:221–222
2. Roberts LW, Coverdale J, Edenharter K, et al: How to review a manuscript: a “down to earth” approach. *Acad Psychiatry* 2004; 28:81–87
3. Martinson BC, Anderson MS, DeVries R: Scientists behaving badly. *Nature* 2005; 435:737–738
4. DeVries R, Anderson MS, Martinson BC: Normal misbehavior: scientists talk about the ethics of research. *J Empirical Res Human Res Ethics* 2006; 1:43–50
5. Reed DA, Kern DE, Levine RB, et al: Costs and funding for published medical education research. *JAMA* 2005; 294:1052–1057
6. Gall MD, Gall JP, Borg WR: *Educational Research: An Introduction*. White Plains, NY, Longman, 2003
7. Regehr G: The experimental tradition, in *International Handbook of Research in Medical Education*. Edited by Norman GR, Van der Vleuten CPM, Newble DI. Boston, Kluwer Academic, 2002, pp 5–44
8. McGaghie WC, Bordage G, Crandell S, et al: Method, research design. *Acad Med* 2001; 76:929–930
9. Shea JA, McGaghie WC, Pargaro L: Instrumentation, data collection, and quality control. *Acad Med* 2001; 76:931–933
10. McGaghie WC, Crandall S: Population and sample. *Acad Med* 2001; 76:934–935
11. McGaghie WC, Crandall S: Data analysis and statistics. *Acad Med* 2001; 76:936–938
12. Ogundipe LO, Boardmen AP, Masterson A: Randomization in clinical trials. *Br J Psychiatry* 1999; 175:581–584
13. Gurwitz JH, Sykora K, Mamdani M, et al: Reader’s guide to critical appraisal of cohort studies, 1: role and design. *BMJ* 2005; 330:895–897
14. Mamdani M, Sykora K, Li P, et al: Reader’s guide to critical appraisal of cohort studies, 2: assessing potential for confounding. *BMJ* 2005; 330:960–962
15. American Psychological Association. Washington DC, *Publication Manual of the American Psychological Association*, 5th ed, 2001, pp 31–76
16. Harris IB. Qualitative methods, in *International Handbook of Research in Medical Education*. Edited by Norman GR, van der Vleuten CPM, Newble DI. Boston, Kluwer Academic Publishers, 2002, pp 45–96
17. Inui TS, Frankel RM: Evaluating the quality of qualitative research: a proposed pro tem. *J Gen Internal Med* 1991; 6:485–486
18. Giacomini MK, Cook DJ: Users’ guide to the medical literature XXIII: Qualitative research in health care: A: are the results of the study valid? *JAMA* 2000; 284:357–362
19. Giacomini MK, Cook DJ: Users’ guide to the medical literature XXIII: Qualitative research in health care: B: what are the results and how do they help me care for my patients? *JAMA* 2000; 284:478–482
20. Begg C, Cho M, Eastwood S, et al: Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 1996; 276:637–639
21. Shea JA, Fostna GS: Psychometric methods, in *International Handbook of Research in Medical Education*. Edited by Norman GR, van der Vleuten CPM, Newble DI. Boston, Kluwer Academic, 2002, pp 97–126
22. Marshall M, Lockwood A, Bradley C, et al: Unpublished rating scales: a major source of bias in randomised controlled trials of treatments for schizophrenia. *Br J Psychiatry* 2000; 176:249–252
23. Sierles FS: How to do research with self-administered surveys. *Acad Psychiatry* 2003; 27:104–113
24. Roberts L, Coverdale J, Louie A: Evidence, methods, and psychiatric education. *Acad Psychiatry* 2003; 27:227–228
25. Pressley M, Harris KR: Increasing the quality of educational intervention research. *Educ Psychol Rev* 1994; 6:191–208