Quantitative Articles:
Developing Studies for Publication
in Counseling Journals
Jerry Trusty

This article is presented as a guide for developing quantitative studies and preparing quantitative manuscripts for publication in counseling journals. It is intended as an aid for aspiring authors in conceptualizing studies and formulating valid research designs. Material is presented on choosing variables and measures and on selecting statistical analyses models. The second section of the article addresses preparation of manuscripts for publication, including article organization and presentation of data.

The purposes of my article are twofold: (a) to present material on developing quantitative studies, an area that falls outside the scope of the APA Publication Manual, and (b) to present material on organizing and writing manuscripts describing quantitative studies, an area that falls within the scope of the Publication Manual. My intent is to highlight important—and sometimes overlooked—material in the Publication Manual, expand and clarify some of the guidelines presented in the Manual, and provide context particular to the field of counseling. In the initial section on developing quantitative studies, I provide suggestions for (a) defining research parameters, (b) designing quantitative studies, (c) operationalizing variables, and (d) selecting appropriate statistical analysis procedures.

Developing Quantitative Studies

The decisions researchers make before data are collected are the most important decisions. There is most often no remedy for poor data other than starting over and correcting mistakes. Thus, there are two major suggestions for getting started with a quantitative study: (a) know the literature in your area thoroughly and (b) conduct a pilot study.

Defining Research Parameters

The researcher's first task is to define the area to be researched. There are two major parameters that require definition: (a) what will be researched and (b) who will be the target population. The what involves the variables to be studied and the models to be tested in the quantitative research paradigm. The who involves determining the parameters of the participants and population to be studied (e.g., age range, cultural experience, presenting concern, environment). To connect the study to the larger knowledge base, researchers must first become intensely knowledgeable about the literature in the research area. The literature review will define the what and the who of the study.

Many researchers, including me, have collected data and later found literature that either specified some variable not included in the study or suggested some measure or practice other than the one used. Thus, a full review of the literature is needed before a study is designed, and this literature review should expand beyond those resources that are eventually cited and referenced in the article. It follows that literature reviews should encompass fields outside counseling and psychology, being more broadly inclusive of the social sciences (e.g., sociology, human resources, communication, education). The fields to which a literature review expands depend entirely...
on the area or population of focus. For example, a study using veterans returning from Iraq or Afghanistan as a target population might warrant review of literature in various areas of psychology, medicine, sociology, and family studies. It is likely that many of the works reviewed during the research design process would not appear in a cogent literature review in the final product, a quantitative manuscript submitted to a counseling journal.

Whereas no study can be absolutely flawless, no aspiring author sets out to include a major flaw in the design. Minor imperfections and compromises are usually known by authors. Major flaws, in contrast, are unknowns. It is, therefore, highly advisable to conduct a pilot study so that unknowns may become known and flaws can be corrected before resources are spent conducting the study. Pilot studies are needed most when researchers are starting their research careers or initiating a new area of research. All researchers should at least pilot their procedures. A pilot of the research procedures often helps researchers refine protocols, instructions, and practices; and the pilot may sometimes lead to major changes in the conduct and content of the research. Postpilot study interviews with participants often reveal vague wording, response difficulties, or other practices or situations that might undermine the validity of the study.

An initial task in developing a study is defining the research problem. Research problems are knowledge gaps that exist in the literature, and these gaps are identified through the literature review. By extension, research knowledge gaps are practice knowledge gaps; filling these gaps helps us to better serve our clients, students, constituencies, and profession. Often, the sections of articles that contain suggestions for future research are useful in formulating research problems. These suggestions are found near the ends of articles, and they often are explicit research problems.

Research problems might also be an overlooked aspect or assumption in the literature. For example, it was long assumed in the literature that African American young adults preferred social occupations and college majors over other types of occupations and majors. When choice of major, however, was examined by socioeconomic status (SES) level, African Americans differed little from their White counterparts (see Trusty, Ng, & Plata, 2000). Researchers who failed to specify SES in their studies, thus, had made a specification error by not including SES as a variable. Research problems or knowledge gaps often exist around population subgroups. For example, a particular research finding might not hold true for both women and men, for various age groups, cultural groups, differently abled groups, and so forth.

The next decision to be made is whether a quantitative or qualitative (or perhaps a mixed-method) study is needed, and research problems are the driving force behind research methods (Silverman & Marvasti, 2008). Generally, the more subjective the research problem, the more appropriate a qualitative method would be; inversely, the more objective the research problem, the more appropriate a quantitative method would be. If little is known about a research area or target population, it is likely that a qualitative study would be needed first. Qualitative studies often identify critical phenomena and illuminate important research problems. Subsequently, quantitative studies might be used to support or disconfirm theoretical formulations resulting from qualitative studies. Creswell (2007) noted that qualitative studies are more inductive in approach (researcher decisions arise from the data), whereas quantitative studies are more deductive (researcher decisions are based on theory or previous empirical findings), and the research design is set before data are collected.

Research problems produce research questions or research hypotheses. There are various traditions around research questions and research hypotheses, but many researchers hold to the logic that researchers should use hypotheses in experimental or quasi-experimental studies and research questions in nonexperimental studies. There is no need to use both research questions and hypotheses. Research questions and hypotheses—when stated clearly and explicitly—set the framework for the research design and, subsequently, for data analysis procedures.

**Designing Quantitative Studies**

Numerous experimental, quasi-experimental, and nonexperimental designs exist. Several resources (e.g., Heppner, Kivlinghan, & Wampold, 2008; McMillan & Schumacher, 2001) are dedicated to describing various research designs, and descriptions are beyond the scope of this article. Rather, I focus on specification of variables and specification of models to be tested. Specification (see Asher, 1983; Pedhazur, 1997) involves naming the variables that will be used and how these variables are conceptualized, arranged, and treated (e.g., independent variables, dependent variables, control variables, sequence of variables, causal linkages).

**Specification of variables.** Variable specification is important across all quantitative research designs, even purely descriptive studies. For example, if the research problem is to describe some population subgroup, it is salient to specify the variables that would adequately describe it. Variable specification evolves from theory and the empirical literature. Any quantitative study benefits from a strong theoretical foundation. Theories not only specify which variables are important but also tie research to the broader knowledge base, thereby making it more likely that the study will contribute to the knowledge base. A major reason for rejection of submitted articles in the anonymous review process is there is little or no contribution to the knowledge base. Theories also supply conceptual definitions of variables, and researchers should be careful to specify variables that are conceptually consistent with theoretical bases.

**Specification of models.** Theories specify how the variables should be classified and arranged in the design of the study.
Causal linkages and causal processes are specified by theories, and researchers' results—including results from correlational studies—either support or disconfirm the causal linkages or predictions specified in theories (see Asher, 1983, and Pedhazur, 1997, for an informed discussion of causality with correlational research).

Know that the term theory is conceptualized broadly. That is, a theory as a basis for a study may not be called a theory or even a model or approach. Rather, it might be a widely held axiom, an accepted assumption, or a conclusion reached by other researchers or practitioners. It becomes, however, the theory that one's research will test. Some informative studies are based on multiple theories because often, no one theory specifies all the variables important to the outcome of interest. For example, extant studies in adolescent development (e.g., Eccles, 1994; Farmer, 1985) use sociological and psychological theories as their comprehensive theoretical bases.

Previous empirical studies should also drive model and variable specification. It is hoped that theories have been formed based on empirical studies and that other researchers have tested one or more tenets of the theory. Sometimes, it is clear when a particular variable should be included in the study being developed, but often results will be mixed, with some supporting inclusion of the variable and others disconfirming its usefulness to the model. In this case, it is best to include the variable, and perhaps the planned study will shed new light on the variable and effects relationships.

Operationalizing Variables

Another major reason for the rejection of studies submitted to counseling journals is weakness in measurement. After variables are specified, the researcher must decide how variables will be quantified or operationalized. One common and useful method for deciding how to operationalize variables is precedent. That is, if previous researchers have been successful in quantifying variables in particular ways, there may be no reason to deviate from established practice. If, however, previous researchers have suggested other ways of operationalizing variables, these suggestions become part of the rationale for selecting a new or different measure. In considering the use of new or different measurement instruments, researchers should be aware of the following: (a) Some instruments are constructed and used for research, whereas others are constructed and used for clinical or evaluation purposes; (b) often, the more established, well-researched instrument is better than one about which little is known; (c) measurement error can undermine the validity of any study; (d) new or unknown measures demand a more thorough description of the instrument in the manuscript; and (e) any measure selected should be conceptually consistent with the theory used to specify the variables and model.

Researchers often develop their own measures for quantifying demographic or background variables. All studies should provide a thorough description of the research participants (see APA, 2010, pp. 29–30). Also, demographic variables are often used as independent variables, control variables, or moderator variables. Thus, it is important that demographic variables be quantified in ways that minimize measurement error. For example, it is common to see collapsed response categories in surveys. Participant ages, for example, might be listed as 20 to 29 years, 30 to 39 years, and so forth. Age could be quantified, however, with much less error if respondents simply indicated their ages in exact years. Researchers, therefore, should be careful to avoid collapsing measurement categories.

In the social sciences, in general, and in the counseling field, in particular, method variance is a problem that frequently undermines the usefulness of studies (see Heppner et al., 2008), particularly correlational studies. When the same methods (e.g., self-report methods, rating methods) are used to measure both the dependent and independent variables, method variance is a common and expected source of response bias. Method variance is a psychometric validity concept; when method variance is present, the relationship between the independent and dependent variable is artificially and spuriously strong. There are several sources of method variance, including socially desirable responding (SDR), participant trait negativity, and participant or rater mood states. Method variance is particularly a problem when the nature of both the independent and dependent variables is the same. For example, if both variables are self-reports of perceptions, method variance should be expected to inflate the relationship or effect.

There are several ways to eliminate or minimize method variance. The best and most obvious way is to design the study so that the independent and dependent variables are measured by different methods (e.g., participant self-report for the independent variable, observer ratings for the dependent variable; perceptions for the independent variables, behaviors for the dependent variable). For established and well-researched measurement instruments, multitrait–multimethod studies may exist. When they do, researchers can examine the discriminant validity of the instrument (how well the instrument discriminates from unlike constructs measured by the same method). If the discriminant validity is sound, method variance is likely minimized. Researchers can statistically control for SDR by including SDR instruments (see Paulhus, 1991) in their analyses. Method variance can also be controlled statistically in a structural equation modeling (SEM) analysis paradigm (see Heppner et al., 2008).

Selecting Statistical Analysis Methods

It is important for researchers to be aware that the design of the study and the nature of the variables determine the statistics to be used and that the statistic used does not determine the design of the study. For example, some researchers erroneously assume that if chi-square statistics are used, the design of the study is necessarily descriptive. This is simply
Researches developed chi-square-related statistics for cases in which either variable in a cross-tabulation is the dependent variable. Similarly, regression does not necessarily mean the study is a correlational design because regression is often used in experimental studies. Similarly, using an analysis of variance (ANOVA) does not mean a study has a causal-comparative or an ex post facto design.

The nature of the specified variables and the nature of the specified model are the major criteria that determine the statistical analysis model or method. The reader should note that in many cases, different analysis models might be appropriately used. In many cases, however, there is one, best analysis model; and in some cases, there is only one possible analysis model.

Particular statistical methods were designed to accommodate particular forms of variable scaling (nominal, purely categorical scales; ordinal scales; interval-ratio scales). All forms of ANOVA and most forms of regression, for example, were designed for dependent variables of interval or ratio scaling. When dependent variables of inappropriate scaling are used, the assumptions of the analysis method are violated. Thus, researchers should pay close attention to variable scaling and the assumptions of the analysis method when selecting statistical procedures. The scaling of independent variables also matters when selecting analysis methods, and particular statistical methods (e.g., least squares regression, logistic regression) can accommodate various scalings of independent variables. In every case, researchers should pay close attention to the assumptions of the chosen analysis model.

As stated previously in the section Operationalizing Variables, researchers should not collapse variables (thereby creating measurement error in variables) to fit them into a particular analysis paradigm. This is most commonly seen with ANOVA or multivariate analysis of variance (MANOVA) when researchers collapse an interval-ratio variable into ordinal categories to use as an independent variable. Some instruments themselves collapse interval scales into typologies, and this practice also creates error. Thus, it is best to use interval-scaled variables in their original form. For example, Fraley and Waller (1998) found that the two continuous adult attachment variables of avoidance and anxiety offer more validity, reliability, and statistical power than the collapsed four-category adult attachment typology.

There are two schools of thought around the relative complexity of statistical models. Some researchers assert that the simplest possible analysis method should be used; others argue for the most complex and robust statistical model. If a more complex model offers no advantages, then the simpler model is best. If a simpler model produces more straightforward results than a more complex model, the advantage lies with the simpler model. More complex models, however, often offer particular advantages, including more statistical power, more flexibility and control over the analysis, and more appropriate analysis methods.

Multivariate analyses methods, although complex, should be used when there are multiple dependent variables in a study. There are two main advantages of multivariate methods (e.g., MANOVA) over univariate methods (e.g., ANOVA): (a) Multivariate methods set the overall alpha level at the selected level (e.g., .05), and (b) multivariate methods account for the relationships among the dependent variables. When a series of separate ANOVAs, for example, are generated, the alphas of the statistical significance tests are additive, thereby increasing the Type I error rate. In contrast, a MANOVA using all the dependent variables in one analysis would keep the overall alpha at the designated, lower level. When a study has multiple dependent variables, the dependent variables are assumed to be related to one another (if not, they should likely not be in the same study). The relationships among the dependent variables should be accounted for in the analysis because this often gives the study more statistical power and the dynamics of relationships may be revealed. Additionally, multivariate methods are often consistent with the theoretical bases of studies.

One complex, yet flexible and powerful method is structural equation modeling (SEM). SEM has increased in frequency in social sciences journals and in counseling journals. SEM computer programs have become easier to use over the last several years, with programs reading data files from standard statistical programs such as SPSS (e.g., SPSS data files are read directly by AMOS, Analysis of Moment Structures; http://www.spss.com/amos/) and offering a graphics interface. From a graphics interface, the researcher draws (specifies) the model to be analyzed on the computer screen. Historically, statistical programs have used a language interface, whether that interface is through syntax or accessed through pull-down menus. SEM is highly flexible, including applications for confirmatory factor analysis, time-series data, path analysis, and analyses involving latent (unobserved) variables. SEM offers a high degree of control over the analysis. With path analysis, for example, SEM programs allow the researcher to specify some paths as unanalyzed when the theory underlying the study specifies no relationship between a pair of variables. In contrast, the traditional path analysis method (a series of standard regressions) does not allow a particular relationship to be unausnalyzed. In addition, SEM programs provide valuable information on the fit of the model to the data.

Preparing Quantitative Manuscripts for Publication

In this section, I focus on article organization and presentation of data. Organization among subheadings is important, as is organization within subheadings. Effective data presentation—in text, tables, and figures—is salient to effectively communicating findings. I highlight commonly encountered inconsistencies found in manuscripts submitted for publica-
tion, and I clarify and expand on important points regarding the review and publication process.

**Article Organization**

First and foremost, researchers should follow the set structure of article organization and headings and subheadings commonly used in counseling journals. The temptation for counselors to be creative is entirely understandable. Adherence to heading and subheading conventions, however, makes the article much more manageable for the reader. Readers expect to find information in particular places. The *Publication Manual* (APA, 2010) provides detailed information on the structure of articles, and this material and the sample papers are useful guides to organization.

There are several commonsense practices that are less explicit in the *Publication Manual* and more a matter of tradition. In the Participants section, the reader will expect to see the *N*, or number of participants, early in this section, preferably in the first sentence. Sampling methods and a description of the participants follow. The Instruments section is typically organized by instrument, and the dependent variable or variables are most often presented first. Within the description of each measure, it is common to first see a general description of the instrument, including purpose, number of items, response format, nature of the items, and scoring practices. Information on instrument validity is presented next; it is desirable to tie the instrument conceptually to the model specified, referring the reader to the theoretical basis of the variable and study. Each instrument description ends with information on reliability, including the reliability calculated with the sample in the researcher’s study.

Other headings in the Method section are somewhat inconsistent in the order they are presented. For example, sometimes the Research Design section fits best as the first section under the Method heading; sometimes it fits best after the Instruments section; and sometimes a separate heading is not needed to describe the research design. The Procedure section usually follows the Participants section or the Instruments section. It is typical to see the Data Analysis section as the last subheading under Method.

There are multiple ways to organize the Results section, and common sense and readability should guide organization. One useful way is to organize the results according to the research questions or hypotheses. For particular types of analyses, organization is dictated by the analysis steps. For example, when using multiple regression, univariate, bivariate, and multivariate analyses are performed. In the univariate phase, the distributions of variables are examined and variables may undergo nonlinear transformations. In the bivariate phase, correlations are generated and relationships are inspected for linearity. In the multivariate phase, the regression is generated. Thus, the Results section subheadings in this example would be Univariate Analysis, Bivariate Analysis, and Multivariate Analysis.

There are several other important topics that, in most cases, should be covered in the Results section, including (a) how missing data were treated, (b) characteristics of participants with unit or item nonresponse, (c) whether the assumptions for the particular analysis were met or approximated, and (d) effect sizes and practical significance. Authors should be careful to include only the findings in the Results section. The subsequent section (Discussion and Implications) provides the reasons for and meaning of the findings.

Discussions and implications are organized in various ways. At times, the best fit is a Discussion and Implications section; in other cases, it is best to have a Discussion section followed by an Implications section. Most often, this works best as one Discussion and Implications section. In any case, it is important in counseling journals to provide the implications of the findings for the practice of counseling, including training and policy implications where relevant.

The Discussion and Implications section is often where submissions to journals fall short. There are several important elements in this section. It is advisable to start this section with a brief summary of the study (two or three sentences) and the most important findings from the study. The Discussion and Implications section should include the *why* of the findings (why you found what you found) and the meaning of the findings. With regard to the why, for example, an unreliable measurement instrument may have been responsible for a non-significant relationship. The meaning of the findings involves putting your data into words that speak to the phenomena studied and the target population. The why and the meaning should be discussed within the framework of the theoretical bases of the study, and works cited in the introduction (introduction and literature review) are cited again and discussed in light of the findings. Effect size (practical significance) is an important component that ties your findings to counseling practice (see Trusty, Thompson, & Petrocelli, 2004, for a discussion of effect size and practical significance).

The challenge in the Discussion and Implications section is finding a balance of extending the meaning and implications of your findings with avoiding extrapolation beyond your findings. It is advisable to have a colleague examine this balance or imbalance before the manuscript is submitted. It is important to use tentative language (e.g., may be, supports, suggests, seems) when writing about the meaning and implications of the findings. If findings are consistent with the findings of others, language could be strengthened (e.g., this body of evidence shows . . . ; these studies, taken together, reveal . . . ).

The Discussion and Implications section should include at least a paragraph on the limitations of the study. It is useful for researchers to think in terms of internal validity (the validity of the study itself, including the measures) and external validity (the generalizability of the findings). Researchers are encouraged to be critical of their own work. It is common,
however, for new researchers to be overly critical of their own work and to not highlight the strengths of the study. It is best, therefore, if both the limitations and strengths are presented, and usually in that order.

The Discussion and Implications section typically ends with suggestions for future research. Studies are designed and embarked upon to answer questions. Studies often, however, generate questions or next steps. A study may have bridged a part of the knowledge gap, but persistent gaps might be illuminated by the study.

Presentation of Data

The Publication Manual (APA, 2010) offers extensive guidance on the presentation and display of data, and numerous examples of data presentation (in text, tables, and figures) are included. Authors should follow these examples closely, working to be explicit and clear in presenting data. For example, there are particular ways to present data in text, and a particular order to what is presented. It is important to avoid redundancy, and if numbers are presented in tables or figures, reiteration in text should be minimized. Note that, in general, two decimal places are sufficient; however, exact probabilities should be reported. Numbers and symbols are expected in the Results section, but the use of numbers should be minimal in the Discussion and Implications section. Effect sizes and a few important numbers are appropriate in the Discussion and Implications section.

Authors should be aware that journals have limits on the numbers of tables and figures. Authors, therefore, should be judicious in what data are displayed and how they are presented. A useful guideline is when there are three or fewer numbers to present, they should be in text. Authors should also examine ways to combine data in tables. There are several good examples in the literature (e.g., Pistole, Roberts, & Mosko, 2010) showing how to combine descriptive statistics (means and standard deviations) with a correlation matrix and how to combine regression results. I suggest that authors not use the Tables function in their word processing programs and, instead, follow the examples in the Publication Manual (APA, 2010). Authors should ensure that tables are visually pleasing and that acronyms for variable names are avoided when possible, thereby making the table easier to read.

Figures are often the best way to present relationships and differences. It is not advisable to present simple descriptive data in figures. That is, bar charts and pie charts typically present only a limited number of discrete data points, and these data would be more straightforward in a table or in text. Bar charts and pie charts also create erroneous illusions because of the areas the bars and slices encompass. Figures should be reserved for describing trends or changes across time (e.g., group means at pretest, posttest, follow up), complex relationships (e.g., causal models with variables and arrows), or other data that are visual in nature (e.g., decision trees, flow charts).

SEM models are best described in figures. Several common practices and traditions have arisen around how to display SEM results in figures. Indicator variables (observed variables) are encased in rectangles, latent (unobserved variables) are encased in ovals, and error terms are encased in circles. Straight single-headed arrows signify causality and causal direction; double-headed curved arrows signify correlation. Circular causality (two-way causality) is depicted by two straight, single-headed arrows signifying reciprocal (nonrecursive model) causality. Parameter estimates are included alongside arrows or adjacent to variables, and chi-square model fit statistics are included in the figure.

It is most useful to generate SEM figures within the SEM computer program that generated the analysis. Most of these programs allow the researcher to name or rename variables, move variables and arrows for clearer display, insert path coefficients along arrows, insert explained variance adjacent to variables, remove unanalyzed effects, show model modifications, insert chi-square and other model fit statistics, and insert notes or keys to the figure. When model fit modifications are included in SEM analyses (finding a best-fitting model), it is usually best when only the final model is displayed. Previous models can be described in text.

Summary and Conclusion

The most important decisions regarding any quantitative research effort are those made before the data are collected. Researchers should engage in a thorough review of the literature. Literature reviews will inform the researcher of important knowledge gaps, and thorough reviews will help researchers define the phenomena to be studied and the target population. It is advisable to conduct a pilot study before the larger study is embarked upon because pilot studies often serve to strengthen the larger study.

Research problems are knowledge gaps identified through thorough literature reviews, and in the counseling context, knowledge gaps are also counseling practice gaps. Research questions and research hypotheses evolve from research problems. Researchers specify the variables to be studied and the model to be tested, using theoretical bases and empirical findings from the literature to guide variable and model specification. Theoretical bases also help researchers connect their studies to the larger knowledge base.

When quantitative manuscripts are rejected in the review process, the problem is often related to variable operationalization and measurement. Researchers should take caution to select measures that are appropriate for research and the theoretical model on which the research is based. Additionally, researchers should be careful to operationalize all variables in ways that minimize measurement error. Method variance is a problem that often arises in counseling research, and researchers can minimize or eliminate method variance by
using different methods for measuring the independent and dependent variables. Also, researchers can statistically control for method variance.

The choice of statistical methods is dictated by the design of the study and the nature of the variables used. The scaling of variables is important in meeting the assumptions of particular statistical analysis models. If multiple dependent variables are used, multivariate procedures work best because multivariate methods take into account the relationships among dependent variables. More complex analysis models (e.g., SEM) offer the researcher a high degree of control over the analysis, and they provide a high degree of flexibility for various research designs.

The Publication Manual (APA, 2010) offers general and specific directions on the organization and content of quantitative articles, including sections as well as headings and subheadings. Authors should closely adhere to the Manual when preparing manuscripts for publication. When authors organize articles in the traditional way, it is much easier for the reader to find the information needed; this is particularly important in the Method section. In the field of counseling, it is important that authors pay close attention to the implications of their findings for the practice of counseling. The challenge for authors is to illuminate the practical meaning of their findings without extrapolating beyond their findings.

Authors should use tables and figures judiciously, combining tables when appropriate and reserving figures for showing complex relationships, data across time, and visual phenomena. When using SEM figures, it is easiest and best if these are generated through the SEM statistical analysis software used.

It is my hope that this article will be helpful for aspiring researchers. From researchers’ perspectives, the natural, overriding desire is to have works published. This perspective is understandable because continuing employment in academia often depends on being published. From editors’ and reviewers’ perspectives, however, the overriding desire is to add to our knowledge bases. When authors approach research projects from the knowledge-base perspective, I think authors’ potential for publication is enhanced.

References


