
Arthur W. Bangert and Julie P. Baumberger

This article provides a methodological review of research designs and statistical analyses published over an 11-year period by the Journal of Counseling & Development (JCD). Results suggest that the majority of published research is more advanced than the research training experienced by the majority of JCD’s readership. Suggestions are made for authors publishing research in JCD to assist the 55,000 plus readers in their efforts to interpret and evaluate the research outcomes reported.

The Journal of Counseling & Development (JCD) is the principal journal of the American Counseling Association (ACA) and is considered by many to be the premier counseling journal in the world. JCD is the only association-wide journal for ACA and is read by more than 55,000 counselors (ACA, 2002) who rely on it as a scholarly resource for improving their work with clients (Borders, 1996). JCD publishes articles focused on counseling theory, practice, research, opinion, and professional association-related matters. The Research section of JCD is devoted primarily to publishing empirical investigations related to counseling theory and methods, counselor training and education, and the efficacy of specific counseling interventions.

The publication of research in peer-reviewed journals is one of the major vehicles for serving and promoting the investigation and dissemination of counseling theory and practice (Lindsey, 1978). However, an overriding issue among practitioners is the degree to which results of empirical research are presented in an understandable and usable format (Sabatino, 1981). This concern suggests that the generalization of research outcomes to practice is dependent on the reader’s ability to understand the research methods and analyses used in studies published in journals like JCD. Familiarity with common research practices found in the studies published by JCD is essential if counselors and other mental health professionals are to make informed decisions regarding the application of research outcomes to their own individual practice.

The historical analysis of research articles like those published in JCD can provide empirical insights into the most widely used research methods and statistical analyses that readers would be expected to understand in order to critically evaluate outcomes with respect to their own practice. According to Thompson and Snyder (1998), because both the counseling field and the methodological fields do evolve over time, and change the consensus about what constitutes accepted practice, it is important to self-evaluate contemporary practice on a regular basis to ensure that ongoing practice reflects current thinking.

Thompson and Snyder’s observation suggests that as research methods evolve, readers’ knowledge and understanding of commonly used methodologies must also advance if they are to stay current in the counseling field. Awareness of current research practices is vitally important for counselor training programs that are charged with ensuring that graduates are equipped with the necessary skills to conduct their own investigations as well as to fully understand the investigations of others.

JCD has been the subject of numerous content analyses regarding general research topics and methodologies. For example, Sexton (1996) investigated trends in counseling outcome research focus and methods by reviewing research published in mental health journals from 1988 to 1994. Results from this analysis indicated that most research conducted during that time period used descriptive field research to explore the efficacy of individual counseling interventions. Whiston and Sexton (1998) found similar results when summarizing school counseling outcome research published between 1988 and 1995. Their study found that most published research during that time frame focused on counseling interventions and incorporated quantitative methods using statistical techniques to answer the primary research questions investigated by the authors. Thompson and Snyder (1998) studied the use and interpretation of statistical significance testing for quantitative research articles published by JCD in 1996. Their research specifically focused on the frequency with which authors who used tests of statistical significance also reported and interpreted effect.
sizes. Their analysis found at least one effect size such as Cohen’s $d$, $r^2$, or $\eta^2$ was reported for 15 of the 25 research articles examined. However, only 2 of the articles reviewed interpreted the relative magnitude of the effect sizes reported.

*JCD* has experienced substantial growth in published research since 1990, when the first Research section was included as a regular feature of the journal (Williams & Buboltz, 1999). However, despite the importance and wealth of periodic literature available to the readership of *JCD*, little attention has been given to examining the extent to which research found in *JCD* can be realistically consumed by readers and translated to professional practice. To date, no comprehensive reviews have been conducted that summarize the prevalent patterns of research methods and statistical analyses most often found in research articles published by *JCD*. Thus, the purpose of this study was to identify the research designs and statistical techniques used most frequently by authors who have published research in Volumes 69–79 of *JCD*.

### Method

The articles reviewed for this study included the total number of research articles published in *JCD* from January 1990 through December 2001. Those research techniques used to answer the main research questions of each article were tabulated. Nonresearch publications included in the Practice, Theoretical, Assessment and Diagnosis, Profiles, and Trends sections of *JCD* were not reviewed for this study. The data-gathering procedure involved a two-step process. First, the research designs were identified based on a classification system devised by Gay and Airasian (1999). Their coding system served as an a priori framework from which *JCD* research articles were surveyed and then categorized into design codes used to answer our primary research questions. As a result of this preliminary review, we decided that research designs would be identified according to specific methodology and classified as either intervention or nonintervention studies. Intervention studies were considered research designs in which the independent variable was under the control of the researcher. This category of designs included true experiments (using random assignment), quasi-experiments, one-group pretest–posttest designs, and single-case studies.

Nonintervention studies were defined as those designs in which a treatment or intervention was not introduced by the researcher. This group of designs included correlational, comparative, qualitative, and test development studies. The test development category consisted of research related to documenting reliability and validity for recently developed test instruments. Survey, content-analyses, cross-sectional, and longitudinal-developmental studies were considered descriptive research and classified as such. The qualitative research category contained investigations in which researchers collected interview data that were synthesized to identify major themes intended to describe the specific phenomena of interest.

Next, the categories for statistical methods used in each research article were developed. Categories for statistical methods were based on the categorization scheme developed byGoodwin and Goodwin (1985a, 1985b) who tabulated and summarized statistical techniques used in research articles published by the *Journal of Educational Psychology* and the *American Educational Research Journal* (AERJ). In addition, commonly used social science statistics texts (e.g., Glass & Hopkins, 1996; Hopkins, Hopkins, & Glass, 1996; Howell, 1999) were reviewed to assist in developing the final categories for coding the statistical techniques. Best practice for reporting results in published research recommends that authors consistently report and interpret effect size and power to enhance the meaning of results from traditional tests of statistical significance (Azar, 1997; Thompson & Snyder, 1998). Thus, for all of the 256 articles reviewed, frequency of reported effect size in terms of variance-accounted-for (e.g., $r^2$, $\eta^2$, multiple $R^2$ or canonical $R^2$) and Cohen’s $d$ (Cohen, 1988) as well as power analyses were tabulated.

Each of the 256 research articles was numbered consecutively beginning with Volume 69 (January, 1990) and ending with Volume 79 (December, 2001), as they appeared in the table of contents of each issue. The first author coded research designs and statistical techniques for all 256 articles included in this study. Interrater reliability was estimated by the second author’s independent coding of a 10% random sample of the articles ($n = 26$) by year (Goodwin & Goodwin, 1985a, 1985b). Reliability was estimated using the following formula: number of coding agreements, divided by the number of coding agreements, plus the number of coding disagreements. The coauthors rated both the classification of the research designs and the statistical techniques. Interrater agreement was calculated separately both for research designs and for statistical techniques. Interrater reliability estimates calculated for the research designs and the statistical techniques were found to be .96 and .92, respectively.

### Results

#### Research Designs

Table 1 summarizes the frequencies of research designs tabulated in the 256 studies reviewed. Where appropriate, $z$ tests (Glass & Hopkins, 1996) were conducted to determine if proportions of research designs differed significantly from the category of interest. In addition, the effect size $h$ was reported for all $z$ tests conducted. According to Cohen (1988), $h$ values would be interpreted as follows: $.20 = $ small, $.50 = $ medium, and $.80 = $ large. Overall, the percentage of nonintervention studies (85%) were coded at a significantly greater frequency than were intervention designs (15%), $z = 11.2, p = 0.000$, $h = .78$. The two most frequently coded designs, which could be grouped under the nonintervention category, were those using correlational (29%) and comparative (29%) methods.
These designs collectively were found to make up a significant proportion of all nonintervention designs coded, \( z = 5.58, p = .000, \ h = .40 \). Of the 16% \((n = 41)\) of articles using descriptive research, the majority (73%) used survey methods to collect data to answer the primary research questions.

Table 1, one-group pretest–posttest designs were used in a very small percentage (3%) of all articles reviewed. However, pretest and posttest data were collected in almost half (43%) of the articles using either true or quasi-experimental designs, suggesting that this data collection method may have been somewhat underrepresented by the one-group pretest–posttest research design category. Test development (3%) and single-case (1%) research were the least identified of all research methods classified.

### Statistical Techniques

Table 2 shows the frequency and percentages of occurrence of statistical methods used to answer the major research questions or test primary hypotheses for each study coded. A total of 745 statistical techniques were coded for all articles reviewed for this study. It should be noted that the number of statistical procedures identified was far greater than the number of articles reviewed because most studies incorporated more than one technique. The most frequently identified statistical procedures for all articles reviewed were those classified as basic (62%), followed by intermediate (22%) and advanced (16%). Basic statistics were found in articles at a significantly greater frequency when compared to intermediate, \( z = 12, p = .000, \ h = .50 \), or advanced statistical procedures, \( z = 13.80, p = .000, \ h = .62 \). The most frequently tabulated statistics across all research reviewed were descriptive procedures (31%), Pearson correlation (12%), multiple regression (8%), one-way analysis of variance (ANOVA; 7%), and one-way multivariate analysis of variance/covariance (MANOVA/MANCOVA; 7%).

### Table 2

<table>
<thead>
<tr>
<th>Major Statistical Techniques</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic</td>
<td>230</td>
<td>31</td>
</tr>
<tr>
<td>Descriptive statistics</td>
<td>88</td>
<td>12</td>
</tr>
<tr>
<td>Pearson correlation</td>
<td>54</td>
<td>7</td>
</tr>
<tr>
<td>One-way ANOVA</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>Chi-square</td>
<td>31</td>
<td>4</td>
</tr>
<tr>
<td>Independent t test</td>
<td>19</td>
<td>3</td>
</tr>
<tr>
<td>Dependent t test</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>Other correlation</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>Other nonparametric</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>Intermediate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple regression</td>
<td>62</td>
<td>8</td>
</tr>
<tr>
<td>Post hoc multiple comparisons</td>
<td>35</td>
<td>5</td>
</tr>
<tr>
<td>Factorial ANOVA</td>
<td>26</td>
<td>4</td>
</tr>
<tr>
<td>One-way ANCOVA</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>Repeated measures ANOVA/ANCOVA</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>Planned orthogonal contrasts</td>
<td>5</td>
<td>0.6</td>
</tr>
<tr>
<td>Factorial ANCOVA</td>
<td>3</td>
<td>0.4</td>
</tr>
<tr>
<td>MANOVA/MANCOVA</td>
<td>53</td>
<td>7</td>
</tr>
<tr>
<td>Factorial MANOVA/MANCOVA</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>Canonical correlation</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>Factor analysis</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>Repeated measures MANCOVA</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>Discriminant analysis</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>Cluster analysis</td>
<td>6</td>
<td>0.8</td>
</tr>
<tr>
<td>Structural equation modeling</td>
<td>5</td>
<td>0.7</td>
</tr>
<tr>
<td>Logistic regression</td>
<td>2</td>
<td>0.2</td>
</tr>
<tr>
<td>Hierarchical linear modeling</td>
<td>2</td>
<td>0.2</td>
</tr>
<tr>
<td>Multidimensional scaling</td>
<td>1</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Note. ANOVA = analysis of variance; ANCOVA = analysis of covariance; MANOVA = multivariate analysis of variance; MANCOVA = multivariate analysis of covariance.

Results from this analysis indicate that descriptive statistics were responsible for one third of the statistical procedures coded. This result is not unexpected when considering that descriptive techniques were always reported, even when other basic, intermediate, or advanced statistical analyses were conducted. The least used basic and intermediate techniques included nonparametric procedures such as the Mann-Whitney \( U \) test, factorial ANOVA, and planned orthogonal contrasts. Cluster analysis, structural equation modeling, logistic regression, hierarchical linear modeling, and multidimensional scaling were the least used advanced techniques across all the research articles examined in this study. These results not only provide insight regarding prevalent patterns of research practice but also show that the \( JCD \) publications reviewed for this study present a broad representation of the types of research methods covered in most basic, intermediate, and advanced social science research and statistics texts.

### Effect Size and Power Analysis

Of the 204 articles reporting results from significance tests, 43% \((n = 74)\) reported at least one effect size (e.g., multiple \( R^2 \), \( \eta^2 \), or Cohen’s \( d \)). Eighty-one percent \((n = 60)\) of the articles...
using tests of statistical significance reported multiple $R^2$. We expected this outcome because 29% of the studies coded used correlational research designs involving multiple regression. Cohen’s $d$ and $\eta^2$ were reported in the remaining 19% of articles that used tests of significance to compare group means. Two of the articles using tests of significance conducted power analyses. Of these 2 articles, 1 calculated the sample size needed to obtain the power required to detect significant differences based on effect size, while the other article reported posterior power based on obtained results.

**Discussion**

**Designs Characterizing Published Research in JCD**

Results indicate that nonintervention designs were the most frequently identified research methods used in studies published by JCD over the 11-year period in which articles were reviewed. Correlational and comparative research methods were found to be the most often used nonintervention designs. The high occurrence of comparative designs is not unexpected when considering that much counseling research is focused on how certain subgroups (e.g., male vs. female or depressed vs. nondepressed) respond to various types of counseling interventions. The correlational research identified in this study almost always involved the use of multiple regression analysis to explain how groups of variables (e.g., self-efficacy and anxiety) contribute to the variability of or predict a dependent variable (e.g., procrastination). Both research techniques are well suited for counseling studies because they provide mental health professionals with insights regarding variables that are important to consider when treating individuals or groups of individuals.

True experiments ($n = 16$) and quasi-experiments ($n = 14$) were found to be the most frequently recorded intervention designs. However, these designs accounted for only 12% of all research coded. One explanation for the small percentage of true experimental studies is that researchers, in many cases, are unable to find the necessary freedom within schools and agencies to randomly assign participants to control and treatment group conditions. Considering the obstacles to forming randomized groups, it was thought that more quasi-experimental research would have been conducted. However, this was not the case. Our results found that quasi-experiments were used even less frequently than true experimental designs in JCD research. This finding suggests that even when randomization is not required, researchers may still have difficulty obtaining enough participants representing clinical populations of interest to form groups of adequate sample size for making valid comparisons.

**Randomized Experimental Designs in Social Science Research**

The high occurrence of nonintervention designs most often used by authors in published JCD research is not surprising when considering historical trends that show that randomized experimental designs have not been prevalent in literature related to educational, social, and psychological research (Boruch, De Moya, & Snyder, 2002). However, there are indications that this trend is changing. For example, over the past three decades, the U.S. Congress has increasingly required the use of randomized experimental designs as a key method for evaluating the effectiveness of federally funded programs (Cook & Payne, 2002). The Early Head Start Program and the Comprehensive Child Care Program (Boruch et al., 2002) are two examples of federally supported preschool programs that used randomized research designs to document intervention effectiveness. The most recent reauthorization of the Elementary and Secondary Education Act, popularly known as the No Child Left Behind Act of 2001, has specifically designated funding for broad-based programs such as enhancing teacher quality, promoting safe and drug-free schools, and supporting Title I goals. However, in order to receive funding, states and localities must demonstrate that the effectiveness of programs they plan to implement are supported by scientifically based research (Feuer, Towne, & Shavelson, 2002). The Coalition for Evidenced-Based Policy sponsored by The Council for Excellence in Government recommends that, as in the field of medicine, randomized field trials be considered the gold standard for evaluating the efficacy of intervention-based programs eligible for funding under No Child Left Behind (Feuer et al., 2002). According to Feuer et al., “when well-specified causal hypotheses can be formulated and randomization to treatment and control conditions is ethical and feasible, a randomized experiment is the best method for estimating effects” (p. 8).

The practices of withholding treatments from certain groups or disrupting ongoing routines have been cited as causing more harm than good in relation to the benefits that the experimental results may provide regarding intervention or program effectiveness. However, despite these objections, the use of rigorous experimental designs has been shown to be feasible in educational and in other settings where it is necessary to explore and evaluate the effects of specific psychological variables and social interventions. We expect that the frequency of published research in JCD using true and quasi-experimental designs will increase dramatically over the next few years as funding for social science research using only the most rigorous of research methods continues to increase.

**Single-Case Research**

Only two of the studies reviewed used single-case research designs. This finding was unexpected because counselors spend a considerable amount of time providing individual therapies to clients who require highly specific interventions. The small representation of single-case investigations is problematic because it demonstrates a lack of research related to the efficacy of individual interventions for clientele with specialized counseling needs. Single-case research, like well-designed group research, can be rigorously de-
signed to control for extraneous sources of error and provide quantitative results that can be used to make objective decisions regarding treatment effectiveness of interventions for clinical populations (Barlow & Herson, 1984). Single-case research methods are also advantageous because they eliminate the need to obtain large numbers of participants representing rare clinical populations necessary for conducting group experimental studies. Finally, single-case research methods are also beneficial for identifying treatment effects that group designs often fail to identify for individuals with unique characteristics beyond the clinical populations they represent (Richards, Taylor, Ramasamy, & Richards, 1999).

One possible reason for the paucity of single-case studies published over the past 10 years may be due to JCD’s emphasis on publishing research conducted by academics who are more interested in investigating the esoteric aspects of counseling theory and methodology rather than the practical aspects of intervention efficacy (Lundevold & Belwood, 2000). The use of applied research such as that represented by single-case investigations is aligned with the Council for Accreditation of Counseling and Related Educational Programs (CACREP; 2001) standards and considered an exemplary practice by training programs that emphasize a scientist-practitioner model of counselor training. Counselors and other mental health professionals would benefit tremendously from greater exposure to the various applications of single-case research methodology for evaluating and improving practice.

Statistical Analyses Reported in JCD Research

Results from the tabulation of statistical techniques indicate that basic statistical procedures were the most frequently used techniques, followed by intermediate and advanced techniques respectively. Descriptive statistics were by far the most frequently used type of statistical procedure and were responsible for 50% of all basic techniques tabulated. Of all the intermediate techniques coded, multiple regression was the most frequently identified statistical analysis. This result was not unexpected given the high frequency of correlational studies using regression analysis to investigate major research questions. One-way MANOVA/MANCOVA was by far the most frequently coded advanced statistical technique.

Approximately 84% of the statistical techniques used in research published by JCD are taught in most introductory and intermediate statistics courses. This suggests that at least a basic to moderate understanding of statistics is necessary to evaluate and understand articles published. However, this finding may be somewhat misleading given the fact that descriptive statistical procedures (i.e., M, SD, percentages) were almost always (97% of the time) included in research studies that used more advanced types of data analysis. When descriptive statistics are removed from the total number of statistical procedures coded, the percentage of advanced statistical techniques rises to 24%, reducing dramatically the difference in proportions between basic and advanced as well as basic and intermediate statistical techniques. With the removal of descriptive statistics, the differences in percentages between basic and advanced statistics and basic and intermediate statistical procedures become small and nonsignificant in practical terms. For example, the recalculated effect sizes with descriptive statistics removed for the comparisons between basic and intermediate statistical techniques were found to be .16 and .30, respectively. These results suggest that basic, intermediate, and advanced statistical techniques were used in similar proportions by articles published in JCD. Considering this reanalysis, professionals relying on JCD for current research findings would need to have, at the very least, a conceptual understanding of the more advanced multivariate statistical techniques to cogently read and understand almost one fourth of all the research reviewed for this study. These findings and implications are very similar to those found as a result of previous content analyses of research articles published in the Journal of Learning Disabilities (Baumberger & Bangert, 1996) and in the AERJ (Goodwin & Goodwin, 1985b).

Effect Size Reporting

Effect sizes are essential when reporting results in terms of statistical significance because they provide a metric for consumers to evaluate the practical importance of research outcomes (Kirk, 1996). Of the 204 articles reporting tests of significance, less than half reported some type of effect size for the statistical analysis conducted. Our findings are consistent with research conducted by Kieffer, Reese, and Thompson (2001), which summarized effect sizes reported in 757 research articles published by the AERJ and the Journal of Counseling Psychology. Most of the studies reporting effect size used multiple regression analysis and reported variance-accounted-for in terms of multiple $R^2$. This result is not unexpected because authors routinely report bivariate and multiple correlation coefficients without interpreting them as effect size indices. It is interesting to note that the last year (i.e., 2001) that research designs were examined for this study, only 7 of 12 articles reporting results from significance tests reported any type of effect size. This is surprising when considering the author guidelines outlined by JCD and the best practices for reporting results of significance testing recommended by the American Psychological Association (2001).

Consistent reporting of effect size in the literature is important because this practice allows the counseling researcher to consider prior effect sizes from related previous studies when planning research powerful enough to detect important effects of counseling treatments and therapies (Kieffer et al., 2001). Once research is completed, effect sizes from a specific study can then be directly compared with effect sizes from other related research. This comparison provides
counseling researchers as well as consumers of counseling research with insights regarding the consistency or generalizability of intervention effectiveness across various settings or situations.

The APA Task Force on Statistical Inference (Wilkinson & APA Task Force on Statistical Inference, 1999), requirements for reporting results of quantitative research outlined in the American Psychological Association (2001) fifth edition of the Publication Manual of the American Psychological Association, and ongoing discussions by research methodologists (e.g., DeVaney, 2001; Thompson, 2002) have been the stimulus for the increased frequency of effect size reporting in published research. According to Thompson (2003), 23 journals now have editorial policies in place that require that effect sizes be reported. For example, the editorial policy for The Journal of Early Intervention now requires that authors not only report but also interpret effect size measures (e.g., Cohen’s \( d \), \( R^2 \), \( \eta^2 \)) for all reported tests of statistical significance (Snyder, 2000).

### Using Effect Size to Enhance Understanding

An important advantage of including effect sizes in addition to tests of significance for group comparisons is that they can be easily converted to results that are practical and lend themselves to easier interpretation of important outcomes for readers with a minimal understanding of statistics (Lipsey, 1990; Rosenthal & Rubin, 1982). For example, suppose that two groups of attention-deficit/hyperactivity disorder (ADHD) students are compared to determine the effects of self-monitoring skills on positive peer interactions. The researcher conducting this investigation finds that the ADHD group taught to use self-monitoring skills demonstrates a .5 standard deviation unit increase over the control group (i.e., ADHD students not taught to use self-monitoring skills) when compared on the mean number of positive peer interactions. This result would be reported as an effect size of .5 in favor of the self-monitoring group and interpreted as moderate in magnitude. However, this effect size could also be interpreted as the following: Seventy percent of ADHD students who were taught to use self-monitoring skills exceeded the mean number of positive peer interactions observed for ADHD students who were not taught to use self-monitoring skills (see Lipsey, 1990, p. 58, for a detailed discussion on converting effect sizes to proportions). Intervention results reported in this manner provide the practitioner who relies on JCD for guidance regarding best practice with a more user-friendly way to evaluate the importance of research outcomes. Another benefit of reporting effect size is to assist researchers interested in conducting meta-analytic studies designed to synthesize major outcomes of counseling research. The failure of authors to report effect sizes consistently in conjunction with tests of significance may have been the primary reason that only one meta-analysis was published during the entire 11-year period that JCD articles were studied.

### Clinical Significance and Individual Treatment Effects

There is no question that reporting and interpreting effect size for results of group research provide practitioners with results that are easier to interpret when evaluating the importance of specific counseling interventions. Although results from group research are essential to consider, practitioners are most often interested in the clinical significance or real differences that potential interventions can create in the everyday life of clients or others with whom clients interact (Kadzin, 1999). According to Jacobson, Roberts, Berns, and McGlinchey (1999), clinical significance is reached when clients (a) show a statistically reliable improvement on assessment measures and, by the end of therapy, (b) posttest scores on clinical assessments fall within normal ranges of functioning. However, Kadzin argued that although large practical effects are the primary index of clinical significance, interventions that yield very small or no effects on outcomes measures (e.g., MMPI-2, Child Behavior Checklist, BASC) still produce important results considered to be clinically relevant. For example, qualitative observations may suggest that a client is better able to cope with symptoms and shows improved social interactions with family members and close friends, even though a positive and significant change is not documented on formal, standardized assessments. Regardless of results from standardized clinical assessments in this case, one could argue that the observed positive changes related to social interactions has resulted in an improved quality of life for the client.

Direct observation has been suggested as a more accurate method for assessing the clinical significance of interventions on behaviors targeted by therapeutic goals. Data from direct observations analyzed using single-case research methodology have been shown to be well-suited for assessing the clinical treatment effects for individuals that otherwise may go undetected on standardized assessments (Barlow & Herson, 1984; Richards et al., 1999). Single-case research methodology also provides a method for therapists to evaluate the efficacy of interventions across multiple goals and multiple perspectives.

The primary difficulty with research designs that use group comparisons is that the results cannot be directly generalized to any individual within a group (Gay & Airasian, 1999). Therefore, it is suggested that, whenever possible, the clinical significance of research outcomes as they pertain to individuals be reported and discussed along with results based on group comparisons. This recommendation suggests that researchers should make every effort to collect qualitative data, in the form of observations and interviews, that can be further analyzed to examine the effects of interventions as they pertain to individuals within a group of individuals receiving an experimental treatment. The methodology surrounding the notion of clinically significant
results is unique to the field of mental health and should be considered by researchers as another potential research method for examining the efficacy of various interventions.

**Conclusion**

Research designs using intermediate to advanced statistical procedures, such as MANOVA, factor analysis, and cluster analysis, have become commonplace in the counseling literature. Access to powerful desktop computers and statistical software packages has greatly increased the researchers’ utilization of these and other more advanced techniques (Baumberger & Bangert, 1996). The prevalence of complex research designs and statistical analyses found in studies published by JCD and other professional journals related to the field of counseling suggests that consumers of this research would profit from at least a conceptual understanding of intermediate and advanced research methods and statistical techniques.

The majority of ACA membership who receive JCD have reported a master’s-level degree as their highest level of professional training (ACA, 2002). An informal review of graduate programs endorsed by CACREP indicates that most master’s-level counselor preparation programs require only one introductory, graduate-level research and/or statistics course. Some programs reviewed require no formal research or statistics course. This finding suggests that many ACA members will most likely find it difficult to comprehend and evaluate the usefulness of much of the research published by JCD.

One solution to this problem would be to require that authors provide explanations of research methods and results from statistical analyses that are detailed, presented in less technical terms, and emphasize the practical importance of findings. This guideline would increase the likelihood that a greater number of professionals reading JCD would have a clearer understanding of the methods and analyses used to determine conclusions based on research outcomes. Another recommendation would be to conduct a comprehensive review of the research and statistics courses required by counselor preparation programs to further investigate the extent to which graduates are prepared to read and comprehend current research studies. Results from a study of this type would be important for providing direction to institutions involved in revising existing counselor training programs.

The intent of this study was not to evaluate the quality of published research but rather to provide the readership with an overview of the most frequently used research designs and statistical procedures used in research published by JCD. The degree of familiarity that readers have with these techniques will determine their capabilities for critically evaluating and profiting from the results of published studies. Interpreting results from published research can be challenging to professionals who depend on the literature to inform their practice. Optimal use of published research will only occur when readers possess enough understanding of research design and statistical procedures to adequately evaluate the contribution of research as it applies to their own individual settings and when authors of published research interpret results in a manner that is comprehensible to the less sophisticated reader.

**References**


