

causality. True experimental designs (or randomized clinical trials) are commonly accepted as the gold standard in offering evidence about the efficacy of an intervention because they are organized to meet the criteria of association, time order, and internal validity (American Psychological Association, 2006; Gambrill, 2006; Johnston et al., 2006). Ideally, there are a number of randomized experimental trials of the intervention's effectiveness relative to a particular outcome. Quasi-experimental and nonexperimental designs provide less conclusive evidence about the effectiveness of the intervention.

But we do not mean to suggest that you need not look critically at the evidence learned from a true experiment, let alone quasi-experimental and nonexperimental designs. Throughout this chapter, we have suggested that there are specific issues you should consider as you read the results of research studies, including the following:

- *Randomization process.* Many authors report using random assignment of participants to the experimental and comparison groups without clarifying how the actual assignment was made. This is important information as you assess the findings' internal validity.
- *Sample size.* In Chapter 4, we briefly mentioned the concept of statistical power; the study needs to have a sample size that is sufficient to detect a statistically significant difference. With small samples, the chances of finding no treatment effect are greater than with a larger sample size; in other words, there may indeed be a treatment effect but the sample size may be too small to detect the impact of the treatment.
- *Attrition.* It is likely that some participants will drop out of the study, and there may be differential rates of attrition for the experimental and comparison groups. You should consider how attrition is handled in the analysis.

Even after you are convinced that the results are meaningful and not the outcome of a poor process of random assignment, small sample size, or attrition, you will have to address the external validity of the findings. Remember, you are taking research-derived knowledge and applying that knowledge to your individual clients. McNeil (2006) notes that “clinical expertise is indispensable for deciding whether external evidence applies to an individual client and, if so, how it should be integrated into treatment” (p. 151). Will the treatment's effects hold true for your clients who, for example, may differ by race, gender, social class, or sexual orientation from the people in the intervention studies? Does the study's setting and location impact the findings? These are all considerations in determining an appropriate intervention.

## DIVERSITY, GROUP DESIGN, AND EVIDENCE-BASED PRACTICE

---

In Chapter 4, we described how historically racial minorities and women had not been adequately represented in research studies. Under the provisions of the *NIH Revitalization Act of 1993* (Pub. L. No. 103-43), women and minorities must be included in clinical research supported by the National Institute of Health. We described recruitment strategies

in Chapter 4. In this section, we highlight the link between adequate representation in research and evidence-based practice.

The most important consideration is the external validity of the findings from group research designs. For what population groups has the intervention been determined to be effective? Although social work research extends to many different areas, the evidence about the inclusion of people of color and women is probably best developed in the evaluations of mental health studies. Miranda, Nakamura, and Bernal (2003) found that between 1986 and 1997 there were few minority participants in studies using true experimental designs to test the effectiveness of treatment for depression, bipolar disorder, schizophrenia, and ADHD; of nearly 10,000 participants, they could only identify 561 African Americans, 99 Latinos, 11 Asian Americans and Pacific Islanders, and no American Indians or Alaska Natives (U.S. Department of Health and Human Services, 2001; Miranda et al., 2003). A more recent analysis of 379 National Institute of Mental Health–funded studies reported that women were adequately included whereas only Whites and African Americans were adequately represented (Mak, Law, Alvidrez, & Perez-Stable, 2007).

Therefore, as you review the available research and try to answer the “for whom” question, it is necessary to identify the characteristics of those who participated in the research. This is likely to be challenging because many of the studies described by Miranda and Mak included no information about ethnicity or lumped all those who were not White into a single category.

Representation alone is insufficient, because there needs to be a sufficient number of participants so that subgroups can be analyzed. Researchers often fail to do an analysis just for women in their sample or just for African Americans. Rather, the results are often only analyzed for the entire sample; different treatment effects for women or people of color may not be reported or observed.

Finally, the broad categories we use to depict racial or ethnic groups tend to imply that all African Americans, all Latinos, or all Asian Americans share the same cultural, social, and historical legacies. Yet, within these groups, there are differences in cultural definitions, language, history, and immigration experience. For example, Vega et al. (1998) found that Mexican immigrants have lower rates of depression than do Mexican Americans born in the United States. You can see that even within what seems like a narrowly defined ethnic group—Mexican Americans—there can be significant differences, in this case based on birthplace. Given that there can be so many variations, evaluating the evidence becomes even more difficult. Therefore, any intervention should at least have a theoretical base, and there is some evidence to link that theoretical base to culture (Miranda et al., 2003).

## ETHICAL ISSUES UNIQUE TO EXPERIMENTAL RESEARCH

---

Social science experiments often involve subject deception. Primarily because of this feature, some experiments have prompted contentious debates about research ethics.