

Criteria for Choosing Knowledge and Assessing Evidence-Based Intervention

Two of the steps in conducting evidence-based practice, as discussed in the previous chapter, are (1) finding the best available evidence relevant for practice decisions and (2) applying critical thinking in analyzing the evidence for its validity, impact on client outcomes, and applicability in practice settings (Gambrill, 2006; Sackett, Straus, Richardson, Rosenberg, & Haynes, 2000). In order to complete these steps successfully, practitioners require the ability to evaluate research studies and translate research into practice generalizations (Briggs & Rzepnicki, 2004; Roberts & Yeager, 2004). This chapter reviews the criteria that may aid practitioners in the assessment of behavioral science research, the use of scientific and practice criteria that can be used to determine the relevance of a study for social work practice, and the characteristics of effective treatment programs for use in comparative evaluations.

SCIENTIFIC CRITERIA FOR THE EVALUATION OF RESEARCH STUDIES

An essential foundation for evidence-based practice is an understanding of the research literature. Evidence-based practitioners employ critical thinking to assess the quality of the research and its relevance for social work practice. Following initial acquaintance with the research process, which can be accomplished through a basic social work research course, practitioners should begin to review research studies in their chosen areas of specialization and evaluate the quality of studies according to the following criteria:

1. *Testability.* Are the basic premises of the study stated in a manner that allows them to be tested adequately? Are the concepts linked to observable events? If not, the basic premises cannot be tested. Moreover, are the key concepts logically interrelated in a consistent, clear, and explicit manner? What are the basic questions and hypotheses of the

study? Are the independent variables specified? Is the question linked in a logical and consistent manner to relevant literature and concepts contained within a particular conceptual framework?

2. *Internal consistency.* How well integrated are the various procedures composing the study? All procedures must be logically consistent. Are concepts operationalized well, meaning that concepts are stated in such a way that they can be measured and observed. Are experimental treatments validated, samples adequately derived, and data inferences checked? Is the study constructed in such a manner that enables answering of the question?

3. *Subsumptive power.* To what extent does the study draw on the available knowledge in the literature? This is assessed by examining the literature review for timeliness, relevancy, and sufficiency in number of references cited. Are research questions formulated in a manner that will add to the available knowledge of social work practice?

4. *Parsimony.* Are the basic relationships between theoretical concepts stated and tested simply and clearly? Practitioners should ask themselves, "Do I really know what is happening in this study?"

5. *Communicability.* To what extent can the findings of the study be communicated to other professionals without distortion and vagueness? Can another practitioner read the study and derive the same conclusions and practice applications? Are the data clearly and concisely communicated? Science is a cumulative social process; hence, it is essential for communications to be clear.

6. *Stimulation value.* To what extent does the study generate other research? How often is it cited in the literature? This criterion reflects the usefulness of a study in producing incentives for other investigators to develop new insights, generate discoveries, and restructure their research endeavors in more profitable ways. How much stimulus value does this study have for me?

7. *Rival factors.* How easily can the basic findings of the study be accounted for by events other than the posited ones, such as history, maturation, testing, instrumentation, regression, selection, mortality, interaction of selection and maturation, interaction of testing and the experimental variable, reactive arrangements, and multiple experimental variable interference (Bloom, Fischer, & Orme, 2009; Campbell, 1967)? These factors are discussed in Chapter 6.

8. *Procedural clarity.* How explicit is the study regarding agreement among various assumptions, relationships, hypotheses, measurement devices, data collection procedures, data analyses, and conclusions? Generally, how well does the study hold together? How well does the study read?

CRITERIA FOR EVALUATING THE RELEVANCE OF RESEARCH FOR SOCIAL WORK PRACTICE

Social work practice typically involves helping clients increase positive behaviors and cognitions related to their treatment goals and decrease those that impede achieving their

goals. Therefore, in addition to assessing the quality of a research study, practitioners assess its relevance for informing practice decisions. The following criteria are relevant in determining whether a study adds to practice knowledge:

1. Does the study develop knowledge that will help explain and predict worker and client behaviors in interactional situations in which services are to be provided? Are client and worker variables that influence the interactional situation, such as gender, age, ethnic background, socioeconomic status, and other relevant social attributes, specified?
2. Does the study lead to knowledge that explains what is involved in forming the relationship? What are the basic features of relationship formulation in terms of body language or verbal expression, for example? Maintenance procedures also must be explained. Are these procedures specified for other practitioners?
3. What are the behaviors involved in the attempt to motivate clients to change, teach new skills, or provide important knowledge? How and when should these behaviors be exhibited by the worker? Are criteria clear as to when the different intervention techniques are to occur: how to proceed, at what pace, how long, and when to terminate?
4. Are treatment techniques related to outcome variables? How valid are the assumptions of the study about explaining and predicting behavior? How accurately measured is the amount of change that took place? Is treatment related to behavioral change?
5. If a study provides relevant practice principles, how useful to workers is the knowledge in terms of the accessibility of the variables involved? Can the variables be identified and manipulated? Is the cost-benefit ratio too great? Does the knowledge violate the values and ethics of the profession?
6. Are procedures for relapse prevention addressed? What procedures are specified to ensure the maintenance and generalization of changed behaviors? Have the change agents created the right conditions within the environment to maintain positive change by, for example, substituting “naturally occurring” reinforcers, training relatives or other individuals in the client’s environment, gradually removing or fading the contingencies, varying the conditions of training, using different schedules of reinforcement, and using delayed reinforcement and self-control procedures (Kazdin, 2001)? Such procedures are important components of effective social work interventions (Wodarski, 2009).

CHARACTERISTICS OF EVIDENCE-BASED INTERVENTION STRATEGIES

In addition to the criteria above that help practitioners assess whether a study is methodologically strong and useful for informing practice, a growing body of literature describes the common characteristics that effective practices tend to share. Interventions designated as evidence-based practices often share several core characteristics. By familiarizing themselves with these, practitioners can critically assess interventions described in the literature in terms of their consistency with research on best practices. Schinke, Brounstein, and Gardner (2002) report that common characteristics of research-based interventions include

1. program content on general life skills;
2. opportunities to practice newly learned skills through modeling and practicing behaviors during sessions and between sessions;
3. emphasis on consistent support for behavior change from the family, school, and community;
4. use of materials that are clear and easy to follow, such as written manuals that provide step-by-step guidelines for each session;
5. emphasis on relationship building;
6. emphasis on strengths rather than deficits;
7. materials tailored to the target group, implemented by bicultural facilitators when offering the curriculum to minority youth;
8. implementation consistent with curriculum instructions; and
9. booster sessions to maintain positive changes.

These are merely guidelines for assessing the quality of interventions. Some promising and research-based interventions do not meet all the above criteria. For example, the absence of a treatment manual or formal curriculum does not mean that the intervention is ineffective. Social workers serve many populations for whom there are no well-researched interventions. In this case, practitioners can use the tools discussed later in this volume to evaluate their interventions. The following sections provide guidance in evaluating the information provided in the research literature and its consistency with evidence-based practice standards.

CONCEPTUALIZATION AND OPERATIONALIZATION OF TREATMENT

Appropriate conceptualization and operationalization of treatment interventions are imperative for the development of effective programs. Workers must be able to specify which behaviors to implement for a given treatment strategy. This represents a difficult requirement for many theoretical frameworks. Usually, therapeutic services are described on a global level and are assigned a broad label, such as transactional analysis, behavior modification, or family therapy. However, such labels are valuable only so long as they specify the operations involved in implementing the services. For instance, the global label of behavior modification can be separated into the following distinct behavioral acts: directions, positive contact, praise, positive attention, holding, criticism, threats, punishment, negative attention, time-out, and application of a token economy (Kazdin, 2001; Wodarski & Pedi, 1977). Moreover, essential attributes of the change agent that facilitate the implementation of treatment should be delineated. The use of treatment manuals can increase the integrity of the intervention (Wodarski, Wodarski, & Dulmus, 2003).

Rationale for Service Provided

The rationale for offering a program should be based primarily on empirical grounds. This decision-making process includes considering the fit between the agency and the service provided. Relevant considerations include potential barriers to implementing the program, communication structures necessary for effective implementation, types of measures that can be used to evaluate the service, accountability mechanisms that need to be prepared, and procedures for monitoring execution of the program (Bellamy, Bledsoe, Mullen, Fang, & Manuel, 2008; Carpenter-Aeby, Salloum, & Aeby, 2001; Dagenais, Brière, Gratton, & Dupont, 2009; Reid, 1978).

Duration

What criteria will indicate that clients have improved to the extent that services are no longer necessary? When these criteria are defined at the beginning of treatment, they enable workers to determine whether or not a service is meeting the needs of the client and to identify the particular factors involved in deciding whether a service should be terminated. Defining these criteria in specific, behavioral terms will help clinicians and clients make an informed decision about when to begin termination.

For example, a clinician and client may agree that when the client has completed vocational training and obtained employment, they will begin to terminate treatment. A school social worker working with a student who is repeatedly asked to leave the classroom due to disruptive behavior may decide, with the student, family, and teacher, that treatment will be terminated when the student is able to remain in the classroom every day for 3 weeks.

Adequate Specification of Behaviors and Baselines

An adequate treatment program must take into account the need for reliable specification of target behaviors, or those behaviors that are to be changed. For example, a treatment program to alleviate antisocial behavior might employ behavioral rating scales in which the behaviors are concretely specified. These could include such observable behaviors as hitting others, damaging physical property, running away, climbing and jumping out windows, throwing objects, and making loud noises and aggressive or threatening verbal statements.

A prerequisite for the adequate evaluation of any therapeutic service is securing a baseline before treatment. A baseline measurement provides a record of the state of the presenting problem before an intervention is provided. For example, a practitioner may observe a child in a classroom and record the number of behavioral incidents 2 days per week for 2 weeks prior to beginning an intervention. Subsequent measurement data can be compared with the baseline data to determine whether clients have improved in targeted areas. Referring back to the example, the practitioner could continue to conduct behavioral observations twice per week while providing a behavior intervention with the child. This enables the practitioner to assess how treatment interventions compare with no treatment interventions.

Measures of Therapist and Client Behaviors

Various measures such as checklists filled out by children and significant others (e.g., group leaders, parents, referral agencies, or grandparents) and behavioral time-sampling schedules can be used to assess change in clients (Bloom et al., 2009; Wodarski & Rittner, 1995). Likewise, behavioral rating scales can be used to assess the behaviors exhibited by a change agent. These and other measures are reviewed in Chapter 5.

The literature of the past decade has called for multicriteria measurement processes for the evaluation of therapeutic services. However, investigators who have used multicriteria measurement indicate that many changes observed with some inventories do not necessarily correspond with results of other inventories used in the same evaluation. For example, studies by Wodarski and colleagues (Wodarski & Buckholdt, 1975; Wodarski, Feldman, & Pedi, 1976; Wodarski & Pedi, 1977, 1978; Wodarski, 2009) found little correlation between self-inventory and behavioral rating scales. In many instances, a change can occur on one of the measurements and not on another measurement. When using multiple measures, it is helpful to include a direct measure of behavior, such as behavioral observation, in addition to indirect measures of behavior, such as self-report inventories (Bloom et al., 2009). It is important to consider the strengths and limitations of any measure used in a study. Because every measurement strategy is likely to have some limitations, using multiple types of measures strengthens a study. If the measures are weak, it is difficult to know whether your data reflect change in client outcomes, or inconsistencies or bias in the measure. Measurement issues will be discussed in greater detail in Chapter 5.

Designs

It frequently has been assumed that the only way therapeutic services can be evaluated is by employing classical experimental designs—those in which participants are assigned randomly to one or more experimental or control groups. However, such designs may not be the most appropriate for the evaluation of services. Implementing these designs may be costly in money, energy, and administration. Moreover, the criterion of random assignment of participants is usually hard to meet in the evaluation of services provided to clients. New time-series designs, however, are emerging from behavior modification literature (see Chapter 7). The designs are easily implemented in social work practice contexts; they cost less money, energy, and administrative execution. These designs provide pilot data that enable a worker to determine if interventions have had an effect on client behaviors.

The emphasis on the use of traditional experimental designs, which involve grouping clients into experimental and control groups, in the evaluation of services in social work is diametrically opposed to a basic practice assumption—namely, that every individual is unique and needs to be considered in his or her own gestalt. The single-case study, which has been championed in recent behavior modification research, may alleviate many of the measurement problems discussed. In this approach, clients serve as their own control, and a client's change is evaluated against data provided during a baseline period that precedes the application of treatment. Chapters 6 and 7 provide criteria by which the worker can determine the appropriateness of the design chosen for evaluation.

Statistics

Evaluation involves several means of assessing whether significant change has taken place. Evaluation of therapeutic services entails the construction of tables and graphs of client and therapist behaviors. Usually, graphs are constructed from measures of central tendencies, such as the mean, mode, or median. A common error in social work practice is to focus solely on what is to be changed in the client and to proceed only to measure that change. Sophisticated evaluation programs measure the behaviors of the client and the change agent simultaneously to enable assessment of how the change agent's behavior has affected the client.

Guidelines on acceptable levels of change are being developed through meta-analysis of numerous studies. Meta-analyses synthesize findings from multiple studies of the same intervention and provide the typical size of the effect of that intervention (Johnson, Scott-Sheldon, & Carey, 2010). These studies typically report an average effect size, a statistic indicating the magnitude of a change in behavior. For example, a meta-analysis of behavioral intervention research would examine the effects of behavioral interventions across a range of studies and provide the average effect size for the intervention. These results can be helpful in determining the typical size of the intervention's effect so that the research can set reasonable expectations for behavior change.

Treatment Monitoring

Having met all prerequisites, it then becomes necessary to monitor the implementation of treatment throughout so that necessary adjustments can be made over time if the quality of treatment varies. If behavioral change is obtained and the investigator can provide data to indicate that treatments were differentially implemented, the change agents can claim with confidence that their treatment is responsible for the observed modifications in behavior. However, if such data cannot be provided when client change has occurred, many rival hypotheses can be postulated to account for the results.

Follow-up

The proper assessment of any therapeutic program with clients involves follow-up. Crucial questions answered by follow-up include whether a therapeutic program has changed behaviors in a desired direction, how long these behaviors were maintained, and to what other contexts they generalized. Has the practitioner worked to maintain positive changes by substituting "naturally occurring" reinforcers, training relatives or other individuals in the client's environment, gradually removing or fading the contingencies, varying the conditions of training, using different schedules of reinforcement, and using delayed reinforcement and self-control procedures (Kazdin, 2001)? Such procedures will be employed in future sophisticated and effective social service delivery systems. Pertinent questions remain concerning when and where a follow-up should occur, how long it should last, and who should secure the measurement. Empirical guidelines for these questions are yet to be developed. Usual procedures include follow-ups 1 and 2 years after service has been provided (Wodarski, 2009).

Overview

This chapter equips practitioners with an overview of the tools used to evaluate studies and assess whether or not the knowledge is relevant to social work practice. A set of criteria are provided that enable the evaluation of treatment programs reported in the literature. As the knowledge produced by the behavioral sciences increases, such criteria will become part of a skill repertoire to aid empirical practitioners in choosing the complex knowledge needed in practice.

Questions for Discussion

1. An article reports that a behavioral intervention resulted in significant improvement in the behavior of a group of fourth-graders with a history of disruptive behavior. What criteria would you use to determine whether you should use these findings to inform your practice as a school social worker?
2. Write a script that you could use to discuss with a client the research on a particular intervention. Practice reading the script aloud. Consider how the information would be received by a client.
3. How would you apply the steps of evidence-based practice if you were working with a client whose presenting problem has not been well researched?

References

- Bellamy, J. L., Bledsoe, S. E., Mullen, E. J., Fang, L., & Manuel, J. I. (2008). Agency-university partnership for evidence-based practice in social work. *Journal of Social Work Education, 44*(3), 55–75.
- Bloom, M., Fischer, J., & Orme, J. (2009). *Evaluating practice: Guidelines for the accountable professional* (6th ed.). Boston: Allyn & Bacon.
- Briggs, H. E., & Rzepnicki, T. L. (Eds.). (2004). *Using evidence in social work practice: Behavioral perspectives*. Chicago: Lyceum.
- Campbell, D. T. (1967). From description to experimentation: Interpreting trends as quasi-experiments. In C. W. Harris (Ed.), *Problems in measuring change* (pp. 112–142). Madison: University of Wisconsin Press.
- Carpenter-Aeby, T., Salloum, M., & Aeby, V. G. (2001). A process evaluation of school social work services in a disciplinary alternative educational program. *Children and Schools, 23*(3), 171–181.
- Dagenais, C., Brière, F. N., Gratton, G., & Dupont, D. (2009). Brief and intensive family support program to prevent emergency placements: Lessons learned from a process evaluation. *Children and Youth Services Review, 31*(5), 594–600.
- Gambrill, E. (2006). Evidence-based practice and policy: Choices ahead. *Research on Social Work Practice, 16*, 338–357.

- Johnson, B. T., Scott-Sheldon, L. A. J., & Carey, M. P. (2010). Meta-synthesis of health behavior change meta-analyses. *American Journal of Public Health, 100*(11), 2193–2198.
- Kazdin, A. E. (2001). *Behavior modification in applied settings* (6th ed.). Homewood, IL: Dorsey.
- Reid, W. J. (1978). The social agency as a research machine. *Journal of Social Service Research, 2*(1), 11–23.
- Roberts, A. R., & Yeager, K. R. (2004). *Evidence-based practice manual: Research and outcome measures in health and human services*. New York: Oxford University Press.
- Sackett, D. L., Straus, S. E., Richardson, W. C, Rosenberg, W., & Haynes, R. M. (2000). *Evidence-based medicine: How to practice and teach EBM* (2nd ed.). New York: Churchill Livingstone.
- Schinke, S., Brounstein, P., & Gardner, S. (2002). *Science-based prevention programs and principles, 2002*. Rockville, MD: Substance Abuse and Mental Health Services Administration, Center for Substance Abuse Prevention.
- Wodarski, J. S. (2009). *Behavioral medicine: A social worker's guide*. New York: Routledge.
- Wodarski, J. S., & Buckholdt, D. (1975). Behavioral instruction in college classrooms: A review of methodological procedures. In J. M. Johnston (Ed.), *Behavior research and technology in higher education*. Springfield, IL: Charles C Thomas.
- Wodarski, J. S., Feldman, R. A., & Pedi, S. J. (1976). The comparison of prosocial and antisocial children on multicriterion measures at summer camp: A three-year study. *Social Service Review, 3*, 255–273.
- Wodarski, J. S., & Pedi, S. J. (1977). The comparison of antisocial and prosocial children on multicriterion measures at a community center: A three-year study. *Social Work, 22*, 290–296.
- Wodarski, J. S., & Pedi, S. J. (1978). The empirical evaluation of the effects of different group-treatment strategies against a controlled treatment strategy on behavior exhibited by antisocial children, behavior of the therapist, and two self-ratings measuring antisocial behavior. *Journal of Clinical Psychology 34*, 471–481.
- Wodarski, J. S., & Rittner, B. (1995). Clinical instruments: Assessing and treating children and families. *Early Child Development and Care, 106*, 43–58.
- Wodarski, J. S., Wodarski, L. A., & Dulmus, C. N. (2003). *Adolescent depression and suicide: A comprehensive empirical intervention for prevention and treatment*. Springfield, IL: Charles C Thomas.