Strengthening Evidence-Based Practice

Zvi Gellis, PhD
William J. Reid, DSW

Social work practitioners have historically been viewed as those who use not only scientific orientation for case work but also the results of research studies for clinical decision making. Dominant conceptions of evidence-based practice emphasize the social worker as a user of scientific findings, with only a minor role given to the social worker as a scientifically oriented practitioner. Problems with this emphasis include a limited base of validated interventions for social work practice, overestimations and shortfalls in the effectiveness of such interventions, and unresolved issues in transporting research findings to actual work with clients. Evidence-based practice could be strengthened by broadening the notion of evidence to include data obtained by practitioners in the application of research-supported interventions—data that might indicate, for example, whether the intervention (or one of its modifications) worked or needed to be changed.

[Brief Treatment and Crisis Intervention 4:155–165 (2004)]

KEY WORDS: evidence-based, single-subject methods, empirical practice, empirically supported treatment.

In this paper we view the current trend toward evidence-based practice (EBP) in social work as the latest development in the profession’s age-old struggle to place its practice on a scientific footing. We use this historical context as a lens through which to examine the potentials and limitations of EBP (as currently defined) and to suggest a view of EBP that may further the cause of a scientifically based social work practice. Although EBP can relate to any level of practice, we focus on applications to direct, or clinical, practice; and, within that focus, on methods of intervention.

Historical Context

EBP is not a new notion in social work. Indeed, the ideal of practice grounded in science was present at the dawn of the profession and was given expression not only in the writings of the early pioneers (Lowell, 1884) but also in the movement of scientific philanthropy. The profession was to be scientific in two senses:
practitioners were to act as scientists in their work with cases or in the administration of charity. In respect to cases, they were to make hypotheses about clients, and they were to test these hypotheses with case data. They were also to make use of the results of scientific studies to inform their practice at both micro- and macrolevels (Kirk & Reid, 2002; Reid, 2002). The ideal remained in force as the profession evolved during the early decades of the 20th century. However, as is often the case in social work, there was a rather large gap between the professed ideal and the actual reality. To paraphrase one wry observation, “Social work was a profession based on a science yet to be invented.”

The invention of a scientific foundation has been a slow, incremental process. The idea that social workers should act as scientists in their practice took shape in the work of Mary Richmond (1917), whose *Social Diagnosis*, as its title suggests, was strongly influenced by a scientific model of medical practice. It is quite likely that her influential book had some effect on the practice of its day. The psychoanalytic movement that dominated professional practice in the middle decades of the past century also reinforce a scientific approach to case study. The fullest expression to date of the social-worker-as-scientist appeared in the 1970s, with the emergence of the empirical practice movement (Reid, 1994). Borrowing from developments in behavioral psychology, the movement introduced methods of scientific case study into social work practice, notably, the single-subject design (SSD). Even the simplest form of this design called for an unprecedented infusion of research methods into practice. The client’s problems were to be defined in specific, observable terms, usually expressed as the occurrence of some behavioral difficulty. Data on the frequency and severity of the problem were to be gathered to provide a baseline before beginning intervention. Data collection was to proceed as rigorously as possible, using such measurement techniques as direct observation, client self-monitoring, and standardized instruments. The purposes of the baseline were to determine if predicted changes occurred after intervention and to guide subsequent treatment decisions. In more-elaborate forms of the design, the intervention could be manipulated to rule out extraneous factors that might be contributing to a client’s change.

The use of research findings as a guide to practice emerged more slowly. Although the first experiments to test the efficacy of different forms of social work practice began to appear in the middle of the last century, results were not particularly promising. A number of major experiments failed to demonstrate that professional social work services could accomplish more than lesser forms of service or more than what clients could accomplish on their own (Fischer, 1976; Mullen & Dumpson, 1972). However, beginning in the 1970s, experimental tests of social work interventions began to show more positive results, a trend that has continued since. Thus, in their review of 129 experimental evaluations of social work programs conducted during the 1990s, Reid and Fortune (2003) found that in 88% of the studies, programs tested outperformed control groups or alternative interventions. Moreover, in the majority of these studies, positive findings were supported by similar experiments conducted in related fields (Reid, Fortune, & Kenaley, 2002; Reid, Kenaley, & Burton, 2003).

In the 1990s, the attention of those concerned with promoting a more scientifically based social work appears to have shifted from the practitioner-as-scientist using single-subject methods to the utilization of research-tested interventions. A decline in interest has been attributed to feasibility problems in agency application of SSD and criticisms of its usefulness (Wakefield & Kirk, 1996). The application of intervention research findings to actual practice had begun to
emerge as the exciting new hope for advancing the cause of science in social work.

**Defining EBP**

As the foregoing has suggested, a critical mass of experimentally tested interventions began to form in social work and related fields during the last quarter of the past century—enough to provide a basis for forms of practice that have been variously called either “research-based,” “empirically supported,” or “evidence-based.” The current trend, at least in social work, is to refer to such practice as “evidence-based”—that is, practice based on tested interventions.

In social work, two major conceptions of EBP have emerged, the first of which has been developed by Gambrill (1999, 2003) and Gibbs (2003). Their view was derived from Sackett, Rosenberg, et al. (1996); Sackett, Straus, et al. (2000); and Gray (2001), who have been among the prime movers of the EBP movement in medicine. Accordingly, EBP consists of a number of steps:

1. Needs for knowledge are translated into answerable questions.
2. Questions are answered through scientifically based evidence, acquired primarily through search of computerized databases.
3. The evidence is critically appraised and then applied to the decision at hand.
4. The outcome of the application is recorded.

The emphasis rests on the client’s concerns and expectations, the sharing of evidence with the clients, and the collaborative decision making about the course of action suggested by the evidence (Gambrill, 2003).

The other major conception of EBP generally refers to practice that uses knowledge and interventions with research validation but leaves open the question regarding how to expedite this practice. For example, Hoagwood, Burns, Kiser, Ringeisen, and Schoenwald (2001) define EBP in child and adolescent mental health services research as “a body of scientific knowledge about service practices—for example, referral, assessment, and case management— or about the impact of clinical treatments on the mental health problems of children and adolescents” (p. 1179). In her *Evidence-Based Social Work Practice With Children and Families*, Corcoran (2000) refers to EBP as “approaches with demonstrated effectiveness,” but she has little more to say in the way of definition or procedures (p. xi). Drake et al. (2001) and Torrey et al. (2001) discuss implementation of EBP on the assumption that such practice consists of service practitioners’ using research-supported treatment guidelines. These notions of EBP are equivalent to research-based, or empirically supported, practice; in fact, the terms are often used interchangeably. This second conception of EBP has become dominant in the human services literature.

Current EBP paradigms stress the role of the practitioner as a user of scientific findings. In these paradigms, interventions are tested through randomized controlled trials (RCTs). Those that prove to be efficacious are then “transported” to service programs, for implementation by line practitioners. The evidence in this form of EBP is generated by group experiments. Little attention is given to the evidence that practitioners might obtain in the application of these presumably efficacious interventions. Largely set aside is the conception of the scientifically oriented practitioner using single-subject methods to test and guide implementation of these interventions.

An alternative definition of EBP would result from broadening the notion of evidence to include not only evidence obtained from research studies but also evidence garnered...
by practitioners in their work with clients. A key idea in this conception would be an integration of the two kinds of evidence in ways that would provide a stronger scientific base for practice than would be the case if only one kind were used. In the following sections, we examine some of the limitations of the current conception of EBP as utilization of existing research, to provide justification for our broader view. We then flesh out this view and try to show how it may help offset limitations of the narrower conception.

**Limited Research Foundations**

Research conducted in recent decades has provided social workers with the beginnings of a base of empirically tested interventions—its adequacy, however, varies considerably by field of practice. Empirical backing for practice is probably most robust in the mental health field, with well-tested interventions available for anxiety disorders, for adult and geriatric depression, and for serious mental illness such as dementia (Baucom, Shoham, Mueser, & Daiuto, 1998; Cohen-Mansfield, 2001; Hoagwood, 2001; Hoffman & Tompson, 2002; Kasl-Godley & Gatz, 2000; Lebowitz et al., 1997; Nathan & Gorman, 2002; Penn & Meuser, 1996). In social work, the field of mental health comprised over a quarter of the more than 100 controlled experiments that had positive results and that were conducted during the 1990s (Reid & Fortune, 2003)—a solid contribution to the much larger body of experimental work that has emanated from psychology and psychiatry (Chambless & Ollendick, 2001; McHugo, Drake, Teague, & Xie, 1999). As one leaves the field of mental health, the grounding for EBP is much more of a patchwork. For example, in the field of child welfare (a major practice arena for social work), tested programs have been developed for the purpose of training parents to improve their child-care skills and to avoid practicing abusive behavior (Meezan & O’Keefe, 1998; Videka-Sherman, 1989), but the effectiveness of family preservation programs has still not been consistently demonstrated, despite numerous experimental evaluations over the past several decades (Littel & Schuerman, 1995; U.S. Department of Health and Human Services, 2001). In the field of health, EBP has a solid foundation with respect to work with clients who have certain disabilities or illnesses, such as Type 1 diabetes (Delamater, Alvarez-Salvat, & McCullough, 2003), back pain (Nicholas, Wilson, & Goyen, 1992), hypochondriasis (Bouman & Visser, 1998), irritable bowel syndrome (Van Dulmen, Fennis, & Bleijenberg, 1996), chronic fatigue (Sharpe et al., 1996), and noncardiac chest pain (Klimes, Mayou, Pearce, Coles, & Fagg, 1990); but EBP lacks many of the core functions of medical social work, such as discharge planning. The unevenness of the research base across fields of practice looms as a major limiting factor in the advance of EBP in social work.

**Inflation of Evidence**

The prevailing paradigm for EBP calls for the rigorous testing of interventions, preferably through RCTs. Such experimental evaluations fall prey to at least two common types of biases that each tends to give an inflated picture of the effectiveness tested. One of these is the experimental demand, or “investigator allegiance,” a tendency for investigators or their assistants to give “a leg up” to interventions they favor. Such investments or convictions can lead to a serious and pervasive form of bias in effectiveness research—what Rubin (1999) has referred to as “expectation for improvement” (p. 637). This kind of bias may operate quite unwittingly, through such means as case exclusions or interpretations of ambiguous
data. A good deal of evidence suggests that such effects occur (Gorey, 1996; Luborsky et al., 1999; Robinson, Berman, & Neimeyer, 1990; Smith, Glass, & Miller, 1980) and are large enough to be a matter of concern.

In social work experiments, investigator allegiance emerged as a problem in the 1970s. Before then, investigators conducting experimental evaluations of social work programs tended to be formed of research teams who were not responsible for developing the intervention tested and, hence, lacked a strong emotional investment in its success. By the 1970s, experimental testing was more likely to be carried out by researchers with practice backgrounds who had developed the programs they were evaluating. By the 1990s, the majority of experimental evaluations of social work programs were directed by program developers or by adherents (Reid & Fortune, 2003). (One cannot help but wonder what role investigator allegiance might have played in the shift from negative to positive findings that also began to occur in the 1970s.) Even though the evidence suggests that investigator allegiance does not account for all intervention effects (Gorey, 1996), such a factor may inflate the results of most experimental tests of interventions, not only in social work, but in related fields as well.

Another common source of bias originates in the dependence on clients’ self-reporting to measure intervention outcomes. Such self-reports take the form of interviews, questionnaires, and standardized instruments. Clients’ self-reports of symptoms or behavior, even on well-validated instruments, may reflect social desirability, expectancy, and cognitive dissonance effects (Karoly & Wheeler-Anderson, 2000). The use of control groups may not be of much help with this form of bias. Clients who have invested time and effort (and often money) in a bona fide intervention program (with the expectancy that it will make things better) may be more likely to exaggerate the program’s benefits or report transient gains than “control” clients who have received either no treatment or minimal treatment. As with investigator allegiance, the consequence of a client’s tendency to exaggerate benefits may be seen as another inflation factor in outcome findings. These forms of bias do not necessarily negate the evidence that suggests that many forms of psychotherapeutic and social work interventions are effective (Reid & Fortune, 2003; Wampold, 2001), but they do raise questions about the extent of their effectiveness.

Limitations on Effectiveness

Even if findings from psychosocial interventions are accepted at face value, the question still remains: What do such interventions accomplish? The benefits seem to occur often as part of the initial response to treatment, but even these may be modest. For example, in their meta-analysis of 34 experiments, Westen and Morrison (2001) found that 40 percent of the clients who began treatment for panic, generalized anxiety disorder (GAD), or depression were improved at the end of treatment, with the percentage higher for panic than for depression or GAD. Improvement, however, was more likely to mean relief rather than absence of symptoms, as the durability of the gains was questionable. For example, just over a quarter of the clients treated for depression had maintained their improvement for 2 years following treatment. Moreover, the clients in these studies were screened to exclude those with different comorbid conditions (e.g., alcoholism), thereby giving the interventions an advantage they may not have in ordinary agency practice. As Westen and Morrison (2001) point out, evidence that interventions are effective is much stronger for mental states, such as depressive episodes, than for disorders, such as a major depressive...
disorder. Following their review of a number of meta-analyses covering a larger variety of problems, Karoly and Anderson (2000) come to essentially the same conclusion about durability of effects: “Treatment gains for complex and chronic problems cannot be expected to persist” (p. 172).

In a similar vein, Foa and Kozak (1997) refer to the “efficacy ceiling” in cognitive-behavioral therapy (CBT), which has become the staple of EBP. Initial optimism about the efficacy of CBT has been tempered by research that suggests that an appreciable proportion of clients fail to improve or they relapse. Moreover, there is no convincing evidence that suggests that the addition of cognitive methods has resulted in any dramatic increase in effectiveness over the more traditional forms of behavior therapy or that CBT or other forms of behavior therapy are increasing in effectiveness. As Hayes (1997) has commented “there seems to be little evidence that effect sizes are noticeably increasing over time” (p. 518). A major reason given for this stall in progress has been the growing disconnect between clinical practice and basic research in the behavioral and cognitive sciences (Eifert & Plaud, 1998; Foa & Kozak, 1997; Hayes, 1997).

**Issues in Transportability**

As noted, interventions found to be efficacious in RCTs are to be transported to line practitioners for implementation with clients. The development of the best method of transportation is still at a beginning stage. It is unlikely that many practicing social workers, left to their own devices, will seek out and adopt empirically supported interventions either from the periodical literature or from Internet databases. The most promising vehicle for dissemination appears to be some form of agency staff training in applying tested intervention protocols. Although the evidence suggests that training of this kind can lead to effective application of EBP (Addis, 2002; Drake et al., 2001; Torrey, 2001), this approach may be more feasible with larger agencies, and, in any case, it may require funding that may not be available. Generalization from RCTs raises additional issues. Client samples used in validation studies may differ substantially from the clientele in a typical agency program. The latter may be less well motivated, may have a different demographic profile, and may be more likely to have multiple diagnoses. Will an intervention for depression that is based on clients who do not have substance abuse problems be effective with an agency client who suffers from both depression and alcoholism? Finally, experimental tests of an intervention are conducted in a supportive (or at least tolerant) organizational environment. Their applications in agency programs may need to be carried out in a less-welcoming environment, especially if such applications are forced “top-down” on staff who do not particularly want them.

Transportability barriers may be especially troublesome in social work, where clinical programs are multidimensional, incorporate a broad definition of service, and form a multimodal ecological approach to intervention. Such programs may generate a diversity of practice goals that may not be consistent with a goal of standardized intervention. Such considerations suggest that interventions tested in an RCT may not necessarily be applied in the same way or with the same degree of efficacy when eventually applied at the level of agency practice.

**What Is Needed**

The foregoing review suggests that existing research on intervention efficacy can, at best, provide only part of the evidence needed for
EBP in social work. If major gaps in research-tested interventions exist, if the interventions that have been tested are likely to have shortfalls in effectiveness, and if their effectiveness is compromised further in the transportation process, then clearly something else is needed. What is most needed, in our opinion, is evidence collected in the application of research-supported interventions. The notion of evidence in EBP needs to be broadened to include local, case-specific evidence, which would help determine whether the tested interventions are effective in actual application. The social worker as a user of research needs to become a scientifically oriented practitioner, to translate research to the case at hand.

Our intent is not to suggest that social workers use data from the cases to augment data from intervention research in EBP; rather, as Gibbs (2003) suggests, single-subject studies should be viewed as a means of testing and supplementing available research findings. Westen and Morrison (2001) propose that clinicians evaluate interventions in their own practices as a means of identifying those with sufficient promise to warrant testing through RCTs.

Perhaps the most comprehensive of these efforts is the Planning and Assessment in Clinical Care (PACC) scheme developed by Woody, Detweiler, Teachman, and O’Hearn (2003). Their rationale for PACC nicely expresses one of our main points: “Because only a portion of clients resemble participants in clinical trials, a fully relevant model of evidence-based practice must use evidence obtained locally from each client” (p. viii). PACC essentially presents an array of empirically based methods regarding assessment, treatment, and the tracking of client progress.

Although such initiatives are available, they have not become part of the mainstream relevant literature, nor are they included in the current thinking about EBP. In our view they need to be, to provide the base for the optimal utility and effectiveness of that form of practice. We now consider some specific ways in which single case methods can be used in the implementation of interventions that have been (presumably) tested through controlled trials.

**Application Evidence**

The term *application evidence* refers to data collected during a clinically based intervention involving a single client, a family, or a small group. Whenever possible, the intervention applied should have an effectiveness supported by controlled research (for the purpose of this discussion, we shall assume such is the case). Various kinds and amounts of application evidence can be collected; however, if the approach is to succeed, it must be made sufficiently flexible to accommodate the varied needs of case situations as well as the pressures faced by overworked practitioners. Thus, what may be preferable to a fixed set of procedures may be a menu of methods that permits users discretion in what to select.

**Benchmarking.** A simple way to connect a tested intervention to the case at hand is to use one of the same measures used in the controlled trial (Addis, 2002; Woody, 2000). So doing permits practitioners to determine how well the intervention is working with their own clients. For example, if the practitioner is currently using a form of interpersonal therapy with a depressed client and if the validating study had used the Beck Depression Inventory (BDI) as its primary outcome measure, the practitioner would then use the same inventory to track the client’s progress. Thus, the failure of the BDI to show the kind of progress obtained in the controlled study’s successful cases would be an indicator that the method chosen was not effective for the client and that a different form of treatment might need to be considered.
Multiple Tracking. As noted earlier, comorbidity is a vexing problem in applying results of a controlled trial to actual cases. How do we know if the treatment for depression works with a client who is both depressed and alcoholic, when the controlled study excluded clients with substance abuse problems? Suppose we want to use another tested intervention for the alcohol problem. How can we be confident that the prior research results—either for the treatment for depression or for the alcohol problem—will hold when both interventions are used together? In working with a client who is depressed and struggling with a drinking problem, a practitioner may want to work on both issues concurrently since they may be mutually reinforcing. The application evidence in this situation might consist of using tested interventions (one for depression and the other for the drinking) and of tracking progress on both problems, again with the same instruments that were used in the controlled trials. The case evidence could then be used to evaluate the effectiveness of both interventions as well as for deciding if a change in approach was indicated.

Modifying and Testing an Intervention. Practitioners often have good reason to modify a tested intervention before it is used in a particular setting. For instance, presume that an intervention was validated with middle-class, Caucasian suburban women and that an agency’s clientele consists of poor African-American single mothers living in the inner city. In this case, a practitioner may need to consider adaptation and may decide that, for to be effective with these clients in particular, the intervention needs to address environmental issues. However, the practitioner would not know whether the effectiveness of the intervention would be compromised by adding the environmental focus or whether such a focus is even needed. Tracking in this case might include both an instrument used in the validation study and one aimed at detecting the impact of the environmental interventions.

The examples here are just some of the ways in which single case methods might be used in EBP. All are compatible with clinical practice; in fact, they all should provide data directly useful for such practice. Moreover, all are relatively simple to use, and none are inordinately time-consuming. (For literature on relevant methodology, see Bloom, Fischer, & Orme, 1999; Herson & Barlow, 1984; and Streiner 1998).

Using the Evidence. Application evidence may be gathered for just a single case or for a series of cases. The latter option may be indicated if an agency plans a substantial implementation of an intervention. If the implementation covered a large number of cases, a random sample could be selected for study. If evidence was gathered on multiple cases, data could be aggregated, perhaps as a part of the agency’s information system (Benbenishty, 1997). In addition to shedding light on the effectiveness of the interventions actually used, the data could not only provide feedback about the controlled trials’ generalizability but also could suggest promising interventions for more rigorous testing.

Conclusion

The potential benefits of EBP for social work are indeed enormous. Social workers will have access to a broad base of reliable information about interventions, thereby facilitating decisions for clinical services. By selecting interventions that are demonstrably effective, the likelihood of improving social work outcomes over the long term will be greatly increased. Moreover, as fewer resources are wasted on efforts that are not effective, more cost-effective practices will be supported. In our view, the realization of these benefits depends on the development of an adequate base of tested
interventions, the control of bias in conducting these tests, the resolution of some knotty transportability issues, and the systematic use of evidence from the case situations to which presumably effective interventions are applied—the last of which has been the focus of this paper. The joining of the results of local applications with the results of experimental studies would create a broader and more useful notion of evidence, one that could facilitate the development of EBP on all fronts.

References


