Meeting the Challenges of Evidence-Based Practice: Can Mental Health Therapists Evaluate their Practice?

Susan M. Love, PhD
Jeffrey J. Koob, PhD
Larry E. Hill, MSW

This cross-sectional study looked at the relationship between mental health therapists’ perception of treatments’ impact and actual change. Twenty-three licensed mental health therapists, providing “treatment-as-usual” to foster children, rated their impact on children’s complaints of depression, anxiety, behavior problems, and self-esteem after 6 months of treatment. Actual change was measured by taking the difference from pre- and posttreatment on standardized instruments: Children’s Depression Inventory, Beck Anxiety Inventory, Child Behavior Checklist, and Rosenberg Self-Esteem Scale. Of the 23 therapists surveyed, only a single therapist on a single outcome variable negatively evaluated the efficacy of their practice. No correlation was found between the therapists’ perceptions of client’s progress and actual child’s outcomes. Mental health therapists are unable to subjectively evaluate their own practice accurately. [Brief Treatment and Crisis Intervention 7:184–193 (2007)]

KEY WORDS: child welfare, foster care, evidence-based practice, mental health.
intervention is achieving its intended goals. To be effective, practitioners must adopt a critical perspective and look for refuting evidence. Interventions that are scientifically supported in one population or across populations may have different outcomes if they are administered under different circumstances (Rubin & Babbie, 2007). Furthermore, there are no published studies that have established complete success across all clients and all circumstances, leaving the possibility that even a well-selected treatment may fail or harm a specific client. The ability to accurately evaluate a client’s response to a given intervention is a critical aspect of protecting his/her well-being. Unfortunately, just as “practice wisdom” often guides treatment, a therapist’s subjective feelings may come into play when determining the effectiveness of his/her practice decisions. Without a critical examination of the client’s goals and progress toward those goals, subjective appraisals are fraught with bias. This can be further complicated if the practitioners’ operating theory is not constructed to invite criticism or refutation. Traditional practitioners either assess their client’s progress favorably or blame the client for being “unmotivated” or “resistive.” None of these assessments encourage the therapist to challenge the efficacy of the intervention.

Background

Social work professionals are reaffirming their commitment to using science as a cornerstone of practice decisions (Gibbs & Gambrill, 2002; Howard, McMillen, & Pollio, 2003; Rosen, 2003). Social work educators have argued for using science to inform practice since Mary Richardson published *Social Diagnosis* in 1917. Unfortunately, however, the application of research in clinical practice is “fuzzy”. Joel Fischer’s 1978 scathing article, *Does anything work?*, found little evidence for the effectiveness of social case work and that professionals were practicing interventions that had minimal or no empirical support (Fischer, 1978). Social workers are not unique, in this aspect. It is estimated that in 1997 only 20% of medical decisions were based on “solid scientific evidence” (Sackett et al. 2000). Rosen (2003) argues, however, that social workers are a group particularly vulnerable to relying on problem heuristics because they share life experiences (e.g., problems of living, relationships) similar to those of their clients. He believes that social workers bring to the profession already well-formed belief systems about what causes these problems and the efficacy of certain interventions in treating them. These interventions are perceived as “intuitive” and not open to refutation—the hallmark of a scientific approach.

The promotion of EBP in social care began in 1994 with the Barnardo’s “What Works?” initiative (Newman & McNeish, 2002), subsequently published as the “best evidence” reviews by Macdonald and Roberts in 1995. This scientific movement aims to minimize the problems engendered by poor interventions (as noted by Rosen, 2003) including expensive lawsuits for incompetence and failure. Child welfare specialists have realized that poor treatment outcomes are not only costly to society (allowing the cycles of poverty, homelessness, and intergenerational maltreatment to continue), but they have significant negative ramifications on a child’s individual development as well, including an increased propensity toward violence, substance abuse, and mental illness (Casey National Alumni Study, 2005; Clarke et al. 1999; Courtney et al. 1990; Wulczyn, Kogan, and Harden, 2003). Fundamentally, the client has a right to effective treatment (Myers & Thyer, 1997) and social workers have an ethical obligation to provide that treatment as effectively as possible (National Association of Social Workers, Code of Ethics 4.01c, 1999).

EBP constitutes a “paradigm shift” (Howard et al. 2003) in how social work practitioners...
learn, practice, and evaluate their trade. Essentially, EBP is the integration of the client’s values and expectations, the healthcare practitioner’s scope of practice, and the best research evidence (Shlonsky & Gibbs, 2004). EBP is a systematic approach to practice that begins by formulating an answerable question. The practitioner then seeks the best evidence, critically appraises it for validity using recognized scientific principles, and then integrates those answers with both the client’s values/wishes and the practitioner’s expertise. Furthermore, it allows the clinician to evaluate the efficacy and efficiency of his/her practice (Sackett et al. 2000).

Currently, there are a menu of options and tools to objectively evaluate clinical practice. Thyer, Artelt, and Shek (2003) report that over the last 40 years, there has been an investment and confidence in single-system research designs, SSRD, to help practitioners evaluate their practice. SSRD are relatively easy to implement and have only two prerequisites: that the client’s issue, problem, or solution lends itself to objective measures and that those measures can be repeated over time (Thyer et al. 2003). An alternative assessment is the use of rapid assessment instruments (Thyer et al. 2003), short (typically take about 10 minutes or less to administer) standardized instrument such as the Beck Anxiety Inventory, or the Children Depression Scale. Ideally the use of a rapid assessment instruments pre–post in concert with SSRD, triangulation, will provide valuable information to the practitioner to the efficacy of the intervention. Relying, though, exclusively on impressions is at risk for desirability bias (Rubin & Babbie, 2007). Objective tools and scientific methods to evaluate practice have been available, but getting practitioners in the field to adopt EBP is proving to be a “formidable challenge” (Proctor, 2004).

The standard arguments for not applying EBP are paucity of available research, limited resources including training and time to apply research in strapped community agencies, and organizational culture (Kessler et al. 2005). The number of well-designed studies in social work has grown from 11 in modern social work history (Fischer, 1978) to over a hundred fully randomized or quasi-experimental studies (Howard et al. 2003). Contractors of community agencies are demanding more measurable outcomes and will soon tip the balance from perceived “unnecessary” costs of training to benefits of EBP to help secure funding. Furthermore, as new social workers enter the field and older workers retire, organizational culture will reflect the values of the new workers.

Essentially, the goal of EBP is to enhance the client’s well-being by asking the question “Did the clinical encounter serve the client?” Rubin and Babbie (2007) suggest that any meaningful evaluation is a collaborative effort. The client and practitioner together identify specific measurable indicators, determine a system by which the indicators can reliably be tracked, and then periodically assess the data. The information gathered will help to inform the therapist if he/she should continue, change, intensify or stop the intervention. Given the importance of accurately assessing a client’s progress, in this study we asked “Can mental health therapists evaluate the efficacy of their practice?” Or, do therapists, as strongly recommended by EBP advocates, need to engage in more rigorous, objective evaluation endeavors?

Hypotheses 1: If mental health therapists were accurate in evaluating their clients’ change, then there would be a correlation between therapists’ impression of client progress and actual client change.

Null hypothesis: If mental health therapists are inaccurate in evaluating their clients’ change, then there would be no correlation between therapists’ impression of client progress and actual client change.
Methods

Participants

Twenty-three licensed mental health professionals were independent contractors for a specific agency that specialized in the mental health treatment of foster children. Their clients were foster children at first entry into public child welfare and ranged in ages from 6 to 17. The treatment-as-usual therapy was conducted in the foster home, weekly, over 6 months during the first half of 2004. California State University, Long Beach’s Institutional Review Board reviewed and approved the study’s protocol. The study was conducted in compliance with the ethical standards of the National Association of Social Workers.

Instruments and Procedures

The foster children were assessed at Time 1 (pretreatment) and 6 months later at Time 2 (posttreatment) with standardized instruments to measure depression, Children’s Depression Inventory (CDI); anxiety, Beck Anxiety Inventory-Youth (BAI-Y); negative behaviors at home, Child Behavior Checklist, Externalizing Scale (CBC-ES); and, self-esteem, Rosenberg Self-Esteem Scale (RSE). The difference between pre and post scores were calculated, divided by the total possible score of each instrument, and multiplied by 100 for a percent change with a positive difference indicating improvement.

Children’s Depression Inventory. Depression was measured using the CDI. In a study, Kovacs (1985) found the CDI coefficient alpha to be .86, indicating good internal consistency. The CDI also has been found to have good construct validity, with scores correlating significantly with the Child Depression Scale ($r = .70$, $p < .01$) (Bartell & Reynolds, 1986).

Beck Anxiety Inventory-Youth. Anxiety was measured using the BAI-Y (Beck, Beck, & Jolly, 2001). The BAI-Y shows strong reliability (Cronbach’s alpha .90) and validity (Beck et al. 2001).

Child Behavior Checklist, Externalizing Scale. Negative behaviors at home were measured using the CBC-ES for ages 6–18 (Achenbach & Rescorla, 2001). The foster parent was asked to complete the form. A Spanish version was made available for Spanish-speaking caregivers. The CBCL has eight syndrome subscales and two composite scales, internalizing and externalizing. The externalizing composite score was a summary of rule-breaking behavior (e.g., “truant,” “set fires”) and aggressive behavior (e.g., “mean,” “teases a lot”) syndrome subscales. The CBCL is a widely used instrument with well-established validity, an inter-interviewer item reliability of .96, and over “four decades of research, consultation, feedback and refinement” (Achenbach & Rescorla, 2001, p. 109).

Rosenberg Self-Esteem Scale. The RSE measured self-esteem. According to Rosenberg (1989), the measure has a reproducibility of 93%, 73% scalability of items, and 72% scalability of individuals. These findings were consistent in the studies by Lorenzo-Hernandez and Oullettee (1998) that report good reliability in terms of internal consistency, Cronbach’s alpha = .78, and Fleming and Courtner (1984) report of a test–retest correlation of .82 at a 1-week interval. The RSE correlated .72 with the Lerner Self-Esteem Scale, showing high convergent validity (Savin-Williams & Jaquish, 1981).

The mental health therapists were given an anonymous survey for each client they treated; both the survey and the children’s posttest were administered 6 months (Time 2) from initiating mental health services. The therapists were asked “Did your intervention seek to
effect this client’s: Depression, Anxiety, Self-esteem, or Negative behaviors at home?’” Therapists were given a “yes” or “no” option to each of the four indicators of client mental health. If the therapist answered “yes” to any of the indicators, then they were asked to respond to “If yes, to what extent do you agree that a significant improvement resulted from the intervention?” The response categories were presented in a Likert scale from strongly disagree (SD = −2), disagree (D = −1), neither agree nor disagree (N = 0), agree (A = 1), and strongly agree (SA = 1). Furthermore, therapists were asked to identify their “professional title” from a menu and “How many years of professional experience do you have in counseling?”

The therapists’ responses were compared to the children’s percent change from pre to post on each of the four indicators of child mental health: depression, anxiety, self-esteem, or negative behaviors at home.

Results

A total of 23 licensed therapists provided individual therapy to a specific foster child in the foster home, weekly, for 6 months. This particular investigation involved clinical treatment-as-usual by community providers. Of the 23 therapists, all of them (100%) volunteered to complete the anonymous survey. Two of the 23 clients did not have pre and post information (both children were too immature to complete the pretesting) and the two corresponding therapists’ surveys were not included in the correlation data analysis. The responses on the excluded surveys were consistent with the responses given by the majority of their colleagues.

Descriptive statistics were run on therapists’ professional title and years of experience. The therapists were found to have an average of 11.5 years experience ranging from 6 to 25 years. Seventy-six percent of the therapists were licensed MFT’s, 14% were licensed MSW’s, and the remaining 10% were licensed PhD-level Psychologists.

Of the 23 therapists, 14 reported that they sought to effect change in all four indicators of child mental health: depression, anxiety, self-esteem, or negative behaviors at home. One therapist did not seek to improve depression, three therapists did not seek to improve anxiety, and 5 of the 23 therapists did not seek to improve negative behaviors at home. All of the therapists sought to improve self-esteem.

Eighteen of the 23 therapists “agreed” or “strongly agreed” that the provided intervention made a significant improvement on all four indicators of child mental health. When asked whether the intervention provided made a significant improvement, only 4 of the 23 therapists reported that he/she “neither agreed nor disagreed.” Uniquely, only a single therapist, and only on a single outcome of focus, “disagreed” that the intervention provided made a significant improvement. This single therapist was the only one out of all 23 licensed mental health professionals studied to negatively evaluate his/her practice effectiveness. See Table 1.

Nonparametric correlations were run on the relationship between the outcome scores and the therapists’ perceptions of each of the four indicators of the child’s mental health: anxiety, depression, self-esteem, and problem behaviors at home. Utilizing the most appropriate and commonly used nonparametric formulas for calculating correlation coefficients with variables that are at the ordinal level of measurement and have a number of ties, Kendall’s tau-b was chosen (Rubin, 2007, p. 171). Correlation is a measure of the strength of relatedness from −1, indicating an inverse relationship, to +1, indicating a positive relationship, with zero indicating no relationship. See Figures 1–4.
Using SPSS to calculate the correlation coefficient for therapists’ impression of client progress and anxiety, we found no correlation, Kendall’s tau-b = 0.099, and no significance, 0.625. Because significance is strongly influenced by the size of the sample, the coefficient of determination was also calculated to measure the shared variance of 0.98% (Pallant, 2001, p. 121). The correlation coefficient for therapists’ impression of client progress and depression, Kendall’s tau-b = 0.212, showed a weak but not significant, 0.269, correlation with a shared variance of 4.5%. The correlation coefficient for therapists’ impression of client progress and self-esteem, Kendall’s tau-b = −0.066, showed no correlation and not significant, 0.729, with a shared variance of 0.44%. The correlation coefficient for therapists’ impression of client progress and negative behaviors at home, Kendall’s tau-b = 0.233, was weak and not significant, 0.281, with a shared variance of 5.4%. See Table 2.

Discussion

The results failed to reject the null hypothesis: if mental health therapists are inaccurate in evaluating their clients’ change, then there would be no correlation between therapists’ impression of client progress and actual client change. When examining the measure of association between therapists’ perception of effectiveness and anxiety with the Kendall’s tau-b, there was no association found. Also, there was no association between therapists’ perception of effectiveness in improving the child’s self-esteem and the child’s measures of self-esteem. Self-esteem was of particular interest given both the therapist’s focus on self-esteem and

| TABLE 1. Frequency of Therapist Perceptions (N = 23) |
|------------------------------------------|--------|--------|--------|--------|--------|
| Depression                              | 1      | 1      | 14     | 7      |
| Anxiety                                 | 3      | 2      | 13     | 5      |
| Behavior                                | 5      | 1      | 9      | 8      |
| Self-esteem                             | 1      | 12     | 10     |

![FIGURE 1](Therapists' perception of their treatment's effectiveness with anxiety.)
the therapist’s operating theory of its importance in mental health. There was a weak association between therapist’s perception of effectiveness with both depression and behavior problems in the foster home. The association, though, was not significant and most likely due to sampling error. The associations, also, only shared less than 5% of the variance. Knowing the therapists’ perception of effectiveness only helped to guess the child’s actual improvement by 5%.

Also, of interest is what the therapist did not focus on. Five of the 23 therapists did not focus on the 6 months of individual therapy on the child’s behavior in the foster home. Another 3 of the 23 therapists did not focus on the child’s anxiety. It would be difficult to believe that a child who has been recently, and for the first time, removed from his/her birth family and placed with strangers in a foster home would not be anxious. If anxiety is the apprehension of future negative events that the mind is preparing to cope with (Brown, O’Leary, & Barlow, 2001), then the circumstances of initial foster care placement has a high risk of overwhelming a child’s perception of coping with...
an unknown future. It is also inconsistent with the research, that children removed from a home environment of severe neglect or abuse would not have concerning behaviors. In the only published prospective study on maltreated children entering the foster care system, Taussig (2002) found that school-age children brought to foster care risks: sexual, substance abuse, self-destructive/suicidal, and delinquent/violent behaviors. Children’s problem behavior is also a major risk factor for placement disruption (National Survey of Child and Adolescent Well-Being, 2005). In the largest foster care system in the United States, Los Angeles County recently reported that 24% of its foster care children experience multiple placement disruptions (C-CFSR, 2006). The more disruptions, the more likely the child will drift in foster care until he/she “ages out” at maturity.

“Empirical evidence suggests that multiple placements lead to psychopathology and other problematic outcomes in children” (Wulczyn et al., 2003, p. 213), including premature parenting.

Of the 23 therapists’ evaluations of their treatment’s effectiveness on depression, anxiety, behavior, and self-esteem, only a single therapist on a single variable, self-esteem, disagreed that the intervention was helpful. Of the therapists who focused on the depression, anxiety, behavior, and self-esteem, 93% agreed or strongly agreed that their interventions made a significant improvement. There appeared to be no reality check. The therapists so strongly believed that their interventions worked, that they did not challenge them. When checked against the reality of standardized instruments, therapists were unable to accurately assess their efficacy regardless of the therapists’ extensive experience, professional license, and 6 months of individual therapy with the client.

### Limitations of the Study

The results of this study should be interpreted with caution. First, the sample size was small...
and did not allow us to compare therapists outside of a single agency. Second, all the therapists were supervised by the same supervisor and shared many of the same values, treatment approaches, and views of mental health. Third, although the surveys were anonymous, social desirability bias may also have influenced the results.

And, finally the study was limited to 6 months of intervention and does not tell us of any potential improvement in the children subsequent to the completion of the study. The results, though, do demonstrate the therapists’ inability to take an accurate reading of their client’s progress at a given point in time.

Conclusions

This study supports the idea that theory-driven interventions, if not objectively assessed, can blind the therapists’ perceptions. Universities preparing mental health professionals have an obligation to teach students to critically think about theories and to put theories “up to the light” by measuring both the intervention (what is delivered) and the outcome (what is its impact on the client)—fundamental principles of EBP.

Future studies need to not only address the relationship between therapists’ perception of clients’ progress and actual change but also how therapists derived their perceptions. A qualitative study in concert with a quantitative study may begin to illuminate this important inquiry. As future research explores disconnect between perceptions and outcomes and universities expose their students to the information, it could help move therapists from the old paradigm to current EBP.

Acknowledgments

Conflict of Interest: None declared.

References


