Examining the Effectiveness of Solution-Focused Brief Therapy: A Meta-Analysis Using Random Effects Modeling

Johnny S. Kim, Ph.D.
University of Texas at Austin
Austin, Texas

Statement of the Research Problem

Over the past two decades, solution-focused brief therapy (SFBT) has become a popular therapeutic model for social workers and other counseling professionals. Practitioners from all disciplines, especially social work, have embraced solution-focused brief therapy because of its flexibility, its collaborative nature, and its emphasis on client strengths. Practitioners do not need to have the answers to a client’s problem because they collaborate with the client to identify the problems, to define goals, and to look for solutions to meet those goals. While the popularity of solution-focused brief therapy has grown amongst practitioners in the U.S. and around the world, research on its effectiveness is still limited (Gingerich & Eisengart, 2000; Triantafillou, 1997; Zimmerman, Prest, & Wetzel, 1997). The purpose of this article is to further examine the effectiveness of SFBT through a systematic review using meta-analytic procedures.

Research Background and Hypotheses

Two of the earliest studies on the effectiveness of SFBT were conducted by the team at the Brief Family Therapy Center (BFTC). De Jong and Hopwood (1996) provide an overview of the first study conducted by Kiser (1988), which consisted of follow-up surveys (6, 12, and 18 months after termination of therapy) of clients to see if they had met their goals or felt they had made significant progress. Results showed 80 percent success rate with 65.6 percent meeting their goals and 14.7 percent making significant improvements. At the 18 month follow-up, 86 percent of the clients contacted reported success. This study, however, used subjective measures and did not use a control group or comparison group which threatens internal and external validity. Furthermore, the researchers in this study counted a client’s report of “some progress” as success (Stalker, Levene, & Coady, 1999).

The second study, conducted by De Jong and Hopwood (1996), involved 275 clients seen at the BFTC from November 1992 to August 1993. Similar to Kiser’s (1988) study, participants were contacted 7-9 months after termination and asked if they had met
their goals. Additionally, therapists asked each participant scaling questions (1 being problem is worse than before and 10 being problem solved) at each session to gauge progress. The final scaling question score was then subtracted from the first scaling question score to come up with an intermediate outcome measurement. These scores were then collapsed into three groups: score -3 to 0 equals “no progress;” score 1 to 3 equals “moderate progress;” and score 4 to 8 equals “significant progress.”

Results from this study found that out of 136 participant responses, 45 percent reported meeting their goals, 32 percent reported some progress towards their goal, and 23 percent reporting no progress after termination from therapy. On the intermediate score measure, 141 responses were calculated based on therapists’ session notes. Results from this measure showed 25 percent reported significant progress, 49 percent reported moderate progress, and 26 percent reported no progress. Limitations of this study are similar to the first study because it lacked standardized measures and lacked multiple measures. Despite the lack of rigorous designs in these two early studies, the initial success and positive results were impressive enough to warrant further research on this promising model.

Most recently, Gingerich and Eisengart (2000) conducted the first systematic qualitative review of all the controlled outcome studies on solution-focused brief therapy up to 1999 (N=15). All of these studies either used a comparison group or single-case repeated measures design to measure various client behaviors or functioning. The studies were divided into three groups according to the degree of experimental control employed.

Five studies met the well-controlled standard, 4 studies met the moderately-controlled standard, and 6 studies met the poorly-controlled standard. The five well-controlled studies reported significant benefits from solution-focused therapy with four out of these five reporting statically significant better results than no treatment or institutional services. The remaining ten studies that did not meet the well-control standard also had similar general conclusions about the effectiveness of solution-focused therapy. However, due to the methodological limitations of these ten studies, it is difficult to draw firm conclusions for the efficacy of solution-focused brief therapy.

Past research studies on solution-focused brief therapy have shown promise as an effective intervention (Gingerich & Eisengart, 2000) and research on this model is still growing with recent studies finding mixed results (Adams, Piercy, & Jurich, 1991; Ingersoll-Dayton, Schroepfer, & Pryce, 1999; Springer, Lynch, & Rubin, 2000; Newsome, 2004). However, solution-focused brief therapy lacks the empirical support to be deemed evidence-based (Triantafillou, 1997; Zimmerman, Prest, & Wetzel, 1997). Because several studies with diverse designs, populations, and findings are emerging on SFBT, a meta-analysis and systematic review appears to be a good approach to examine the state of the empirical evidence for this model. Research syntheses are being used increasingly to inform decision makers about the effects of a particular policy (Matt, 1997).

This study will systematically examine the overall effectiveness of solution-focused brief therapy through the statistical method of meta-analysis using random effects modeling. The specific research questions this dissertation will address are:
1. How effective is solution-focused brief therapy for different outcomes such as externalizing behavior problems, internalizing behavior problems, and family or relationship problem?

2. Do the effect sizes vary across studies?

Methodology

In defining the problem statement, the unifying construct for this present meta-analysis is the effectiveness of the therapeutic model of solution-focused brief therapy. Studies were identified through various electronic databases (PsycINFO, Academic Search Premier, UMI Dissertation Abstract, and the Behavioral and Social Science Index) using the keywords, “brief solution-focused therapy,” “solution-focused brief therapy,” “solution focused therapy,” and “solution-building.” Other unpublished studies on SFBT were also obtained by contacting researchers who attended the annual solution-focused brief therapy conference.

Only primary outcome studies (from 1988 to 2005) that examined the effectiveness of solution-focused brief therapy were included in the meta-analysis. Solution-focused brief therapy will be operationalized based on the criteria set by de Shazer and Berg’s (1997) article as well as Gingerich and Eisengart’s (2000) systematic qualitative review:

1. The therapist uses the “miracle question;”
2. Use of scaling questions;
3. A consulting break and giving the client a set of compliments;
4. Assigning homework tasks;
5. Looking for strengths or solutions;
6. Goal-setting;
7. Looking for exceptions to the problem.

Currently, these core components remain important techniques for change in the SFBT and are an integral part of doing SFBT as identified by the main developers of the model. Therefore, it was determined a priori that at least one of these core components must be utilized in order for a study to be considered solution-focused brief therapy and the authors of the primary studies must identify the intervention as solution-focused. This decision was based on the Gingerich and Eisengart’s (2000) published article which used this similar selection criterion. Additionally, if a study does not contain at least one of these core components, or if a study combined these components with other elements from different therapeutic interventions, then it was excluded from this meta-analysis.

For each study, effect sizes were calculated using independent-groups pretest-posttest design sample estimator when pretest and posttest scores for both groups were available (see Morris & DeShon, 2002 for formulas) and Hedges’ g with the unbiased estimate when only mean posttest scores were available. Hedges’ g with the unbiased estimate correction provides a better estimate for smaller samples and is the recommended effect size estimator for single studies (Hedges & Olkin, 1985). Accompanying variances were also calculated for each study and used to calculate weights by taking the inverse of the variance score. For those studies that report non-
significant results without providing any detailed statistical information, an effect size of zero was substituted for non-significant outcomes. This provides a more conservative pooled point estimate of the effect size (Perry, 1997).

A common issue that arises when calculating effect sizes for a primary study is what to do when there are multiple measures for a single construct. The approach taken for this study is based on Lipsey’s (1994) suggestion to calculate individual effect sizes for each of the different measures in a single study and then average them to come up with one effect size for that measure. Similarly, a study may provide an effect size for all the dependent variables, which measure different constructs, in that primary study. It is recommended that only one effect size value should represent a study in any analysis in order to ensure statistical independence of the data (Bangert-Drowns, 1997; Devine, 1997).

Hierarchical linear modeling (HLM) software was used to synthesize the primary studies to calculate an overall effect size estimate as well as test for between-study variability. The application of HLM is appropriate for meta-analysis because meta-analysis can be viewed as a hierarchical data set with sample subjects within each primary study at the first level and primary studies at second level (Hox, 2002). Hierarchical linear modeling takes into account variations at the subject level as well as at the study level. The first model in HLM typically investigates the unconditional model that has no predictor variables at either level as well as a test of the statistical significance of the variability between studies. A statistically significant result indicates that study outcomes are heterogeneous and study descriptors can be added to the unconditional model (making it a conditional model) to see whether they help explain some of this variability. If the variance component in the unconditional model is not statistically significant, then we can assume homogeneity and that sampling error and random error accounts for the differences in effect size estimates across studies (Raudenbush & Bryk, 2002).

The conditional model analysis investigates variation among the solution-focused brief therapy study effect sizes as a function of study characteristics, in addition to estimating the variance component of the unexplained basis of heterogeneity among studies (Kalaian, Mullan, & Kasim, 1999). The level 1 or within-studies model in the conditional model is the same as in the unconditional model and the level 2 or between-study model is modified to include study characteristics to account for the variation among effect size parameters (Raudenbush & Bryk, 2002). Thus, this meta-analytic study on the effectiveness of solution-focused brief therapy will use a random effects model (estimated using HLM software) to calculate overall synthesized effect size estimates.

Results

The results from the literature search produced 22 studies, both published and unpublished dissertations, which met the criteria to be included in the meta-analysis. The studies were divided and grouped into three categories based on the outcome problem each study targeted (externalizing behavior problems, internalizing behavior problems, and family and relationship problems). The three categories had between 8 to 12 studies each with 5 studies (Huang, 2001; Marinaccio, 2001; Moore et al., in press; Seagram,
Results from the study found that solution-focused brief therapy demonstrated small, but positive treatment effects favoring SFBT group on the outcome measures. The unconditional random effects model shows an overall weighted mean effect size estimate of .11 for externalizing behavior problems with a 95% confidence interval range of -.14 to .36. An overall weighted mean effect size estimate for internalizing problem behaviors was .26 with a 95% confidence interval range of .05 to .47. And finally, an overall weighted mean effect size estimate for family and relationship problems was of .26 with a 95% confidence interval ranged from -.03 to .55. Only the magnitude of the effect for internalizing behavior problems (such as depression, anxiety, self-concept and self-esteem) was statistically significant at the p<.05 level thereby indicating that the treatment effect for SFBT group is different than the treatment effect for control group. It doesn’t appear SFBT is as effective with externalizing behavior problem such as hyperactivity, conduct problems, aggression or with family and relationship problems.

In addition, externalizing behavior problems and family and relationship problems both had statistically significant between-study variance estimates in the unconditional model. Therefore, we rejected the null hypothesis of homogeneity and concluded that variability between studies was not explained by sampling and random error alone and that considerable variability remains to be explained between study effect sizes. This significant variability warranted further exploration through a conditional model using predictor variables. Unfortunately, the limited number of studies for both outcome measures does not allow for enough statistical power to test predictor variables through a conditional model and therefore was not performed.

Utility for Social Work Practice

While results from the meta-analysis reports small treatment effects for all three outcome measures, the results were comparable to other meta-analyses that examined the effectiveness of social work practice models (Gorey, 1996) and psychotherapies (Loesel & Koeferl, 1987; Weisz, Weiss, Han, Granger, & Morton, 1995; Weisz & Jensen, 1999; Weiz, McCarty, & Valeri, 2006), especially under real world settings. Furthermore, there is some evidence of clinical significance given the brief nature of the model, the wide array of treatment problems the model has been used, and the applied nature of the studies’ setting. Results from this meta-analysis will help practitioners in their pursuit to apply evidence-based practice. To date, this is the first meta-analysis on SFBT and allows those interested in the therapy model to examine the empirical evidence quickly and with more quantitative information than traditional primary research studies.
References

References marked with an asterisk indicate studies included in the meta-analysis.


Appendix

Table 1: Externalizing Behavior Outcome Results

<table>
<thead>
<tr>
<th>Author</th>
<th>Population</th>
<th>Sample Size</th>
<th>Outcome Measure</th>
<th>Effect Size (d)</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Franklin et al. (forthcoming)</td>
<td>Students</td>
<td>85</td>
<td>Credits Earned</td>
<td>.47</td>
<td>(.03, .91)</td>
</tr>
<tr>
<td>Gallardo-Cooper (1997)</td>
<td>Mothers &amp; Teachers</td>
<td>66</td>
<td>Eyberg Child Behavior Scale, Sutter-Eyberg Student Behavior Scale</td>
<td>-.14</td>
<td>(-.56, .28)</td>
</tr>
<tr>
<td>Huang (2001)</td>
<td>Couples</td>
<td>39</td>
<td>Conflict Tactics Scale, Scaling Questions</td>
<td>-.43</td>
<td>(-1.24, .38)</td>
</tr>
<tr>
<td>Ingersoll-Dayton et al. (1999)</td>
<td>Elderly</td>
<td>21</td>
<td>Modified Caretaker Obstreperous-Behavior Rating Assessment</td>
<td>.32</td>
<td>(-.30, .94)</td>
</tr>
<tr>
<td>Marinaccio (2001)</td>
<td>Students, Mothers, &amp; Teachers</td>
<td>120</td>
<td>Behavioral Assessment System for Children (BASC)- conduct &amp; aggression subscale</td>
<td>-.25</td>
<td>(-.56, .06)</td>
</tr>
<tr>
<td>Moore et al. (under review)</td>
<td>Students</td>
<td>59</td>
<td>Achenbach Behavioral Checklist- Teacher &amp; Student externalizing behavior subscale</td>
<td>.74</td>
<td>(.20, 1.28)</td>
</tr>
<tr>
<td>Newsome (2004)</td>
<td>Students</td>
<td>52</td>
<td>Grades &amp; Attendance</td>
<td>0</td>
<td>(-.55, .55)</td>
</tr>
<tr>
<td>Triantafillou (2002)</td>
<td>Children</td>
<td>30</td>
<td>Devereux Scales of Mental Disorder-externalizing score &amp; critical pathology score, Social Skills Rating System, Total Number of Problem Behaviors, Total Number of Physical Restraints</td>
<td>.17</td>
<td>(-.59, .93)</td>
</tr>
</tbody>
</table>
### Table 2: Internalizing Behavior Outcome Results

<table>
<thead>
<tr>
<th>Author</th>
<th>Population</th>
<th>Sample Size</th>
<th>Outcome Measure</th>
<th>Effect Size ($d$)</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bozeman (1999)</td>
<td>Psychiatric Patients</td>
<td>52</td>
<td>Beck Depression Inventory, Nowotny Hope Scale</td>
<td>.56</td>
<td>(-.01, 1.13)</td>
</tr>
<tr>
<td>Cook (1998)</td>
<td>Students</td>
<td>68</td>
<td>Piers-Harris Children’s Self-Concept Scale</td>
<td>.28</td>
<td>(-.21, .77)</td>
</tr>
<tr>
<td>Huang (2001)</td>
<td>Couples</td>
<td>39</td>
<td>Beck Depression Inventory</td>
<td>.23</td>
<td>(-.58, 1.04)</td>
</tr>
<tr>
<td>Leggett (2004)</td>
<td>Students</td>
<td>67</td>
<td>Coopersmith Self-Esteem Inventory, Children’s Hope Scale</td>
<td>.04</td>
<td>(-.45, .53)</td>
</tr>
<tr>
<td>Marinaccio (2001)</td>
<td>Students</td>
<td>48</td>
<td>Student Report of Personality, BASC-adaptability, anxiety, social skills subscales</td>
<td>.06</td>
<td>(-.24, .37)</td>
</tr>
<tr>
<td>Moore et al.</td>
<td>Students</td>
<td>59</td>
<td>Achenbach Behavioral Checklist- Teacher &amp; Student internalizing behavior subscale</td>
<td>.74</td>
<td>(.20, 1.28)</td>
</tr>
<tr>
<td>Seagram (1997)</td>
<td>Youth Offenders</td>
<td>40</td>
<td>Coopersmith Self-Esteem Inventory</td>
<td>-.06</td>
<td>(-.70, .58)</td>
</tr>
<tr>
<td>Springer et al.</td>
<td>Students</td>
<td>10</td>
<td>Hare Self-Esteem Scale</td>
<td>.57</td>
<td>(-.91, 2.05)</td>
</tr>
<tr>
<td>Sundstrom (1993)</td>
<td>College Students</td>
<td>40</td>
<td>Beck Depression Inventory, Depression Adjective Checklist</td>
<td>1.18</td>
<td>(.48, 1.88)</td>
</tr>
<tr>
<td>Triantafillou (2002)</td>
<td>Children</td>
<td>30</td>
<td>Devereux Scales of Mental Disorder- internalizing score</td>
<td>-.46</td>
<td>(-1.23, .31)</td>
</tr>
<tr>
<td>Villalba (2002)</td>
<td>Students</td>
<td>59</td>
<td>Piers-Harris Children’s Self-Concept Scale</td>
<td>.11</td>
<td>(-.41, .63)</td>
</tr>
<tr>
<td>Wettersten (2002)</td>
<td>Adults</td>
<td>65</td>
<td>Brief Symptom Inventory</td>
<td>.26</td>
<td>(-.24, .76)</td>
</tr>
</tbody>
</table>
Table 3: Family and Relationship Outcome Results

<table>
<thead>
<tr>
<th>Author</th>
<th>Population</th>
<th>Sample Size</th>
<th>Outcome Measure</th>
<th>Effect Size (d)</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adams et al. (1991)</td>
<td>Families</td>
<td>40</td>
<td>Immediate Outcome Rating Scale-Goal Clarity, Optimism, &amp; Compliance</td>
<td>.70</td>
<td>(.04, 1.36)</td>
</tr>
<tr>
<td>Cockburn et al. (1997)</td>
<td>Orthopedic Patients</td>
<td>48</td>
<td>Family Crisis Oriented Personal Evaluation Scales, Psychological Adjustment to Illness Scale-Self Report</td>
<td>1.23</td>
<td>(.30, 2.16)</td>
</tr>
<tr>
<td>Eakes et al. (1997)</td>
<td>Families</td>
<td>10</td>
<td>Family Environment Scale</td>
<td>.52</td>
<td>(-.38, 1.42)</td>
</tr>
<tr>
<td>Huang (2001)</td>
<td>Couples</td>
<td>39</td>
<td>Marital Status Inventory, Dyadic Adjustment Scale</td>
<td>.25</td>
<td>(-.56, 1.06)</td>
</tr>
<tr>
<td>Sundman (1997)</td>
<td>Adults</td>
<td>200</td>
<td>Therapist&amp; Client Completed Questionnaire</td>
<td>0</td>
<td>(-.28, .28)</td>
</tr>
<tr>
<td>Triantafillou (2002)</td>
<td>Children</td>
<td>30</td>
<td>Parent-Adolescent Communication Scale, Family Adaptability &amp; Cohesion Scales II</td>
<td>-.56</td>
<td>(-1.33, .21)</td>
</tr>
<tr>
<td>Zimmerman et al. (1996)</td>
<td>Parents</td>
<td>42</td>
<td>Parent Skills Inventory</td>
<td>.17</td>
<td>(-.52, .86)</td>
</tr>
<tr>
<td>Zimmerman et al. (1997)</td>
<td>Couples</td>
<td>36</td>
<td>Dyadic Adjustment Scale</td>
<td>.29</td>
<td>(-.20, .78)</td>
</tr>
</tbody>
</table>