Objective: To follow distressed married couples for 5 years after their participation in a randomized clinical trial. Method: A total of 134 chronically and seriously distressed married couples were randomly assigned to approximately 8 months of either traditional behavioral couple therapy (TBCT; Jacobson & Margolin, 1979) or integrative behavioral couple therapy (IBCT; Jacobson & Christensen, 1998). Marital status and satisfaction were assessed approximately every 3 months during treatment and every 6 months for 5 years after treatment. Results: Pre- to posttreatment effect sizes on marital satisfaction were \( d = 0.90 \) for IBCT and \( d = 0.71 \) for TBCT, which were not significantly different. However, data through 2-year follow-ups revealed statistically significant superiority of IBCT over TBCT in relationship satisfaction, but subsequent data showed increasing similarity and nonsignificant differences in outcome. At 5-year follow-up for marital satisfaction relative to pretreatment, effect sizes were \( d = 1.03 \) for IBCT and \( d = 0.92 \) for TBCT; 50.0% of IBCT couples and 45.9% of TBCT couples showed clinically significant improvement. Relationship status, obtained on all 134 couples, revealed that 25.7% of IBCT couples and 27.9% of TBCT couples were separated or divorced. These follow-up data compared favorably to other, long-term results of couple therapy. Conclusion: TBCT and IBCT both produced substantial effect sizes in even seriously and chronically distressed couples. IBCT produced significantly but not dramatically superior outcomes through the first 2 years after treatment termination but without further intervention; outcomes for the 2 treatments converged over longer follow-up periods.

Keywords: behavioral couple therapy, posttherapy adjustment

A large number of randomized clinical trials have convincingly demonstrated that couple therapy, in general, and behavioral approaches to couple therapy, in particular, lead to substantial improvements in relationship quality (Shadish & Baldwin, 2005; Snyder, Castellani, & Whisman, 2006). These improvements tend to persist over the short term—six months to a year after treatment termination (Christensen & Heavey, 1999). However, very few studies have examined the impact of couple therapy two years or more after the end of treatment. This lacunae in the literature is unfortunate, allowing limited conclusions about the durability of treatment effects.

The few studies that do exist cast doubt on the long-term maintenance of treatment gains in couple therapy. In a study comparing several versions of traditional behavioral couple therapy, Jacobson, Schmaling, and Holtzworth-Munroe (1987) found that 56% of couples were unchanged or had deteriorated from their pretreatment status by two years following therapy. Snyder, Wills, and Grady-Fletcher (1991) assessed couples four years after the end of behavioral couple therapy and found that 58% of couples were unchanged or had deteriorated from their pretreatment status; 38% had divorced. In the longest follow-up of couple therapy to date, Cookerly (1980) examined the status of 163 couples five years posttherapy, treated in either conjoint forms of marital therapy or nonconjoint forms of marital therapy. Only 56.4% of those treated in the conjoint format were still married at the end of treatment; only 29.8% of those treated in nonconjoint formats remained married. Thus, the two clinical trials described above on behavioral couple therapy suggested that over half of couples are
unchanged or deteriorated two or more years after treatment termination, with almost 40% divorced. The Cookerly study, which was not specifically focused on behavioral approaches but rather on general marital therapy, found that over half of couples were divorced at five years posttreatment. Apart from these disappointing results, these studies examined only endpoint measures of long-term outcome, so it is not possible to see the trajectory of change, namely, how couples fare over multiple time points after treatment termination.

In the largest clinical trial of couple therapy to date, Christensen et al. (2004) compared the effectiveness of traditional behavioral couple therapy (TBCT; Jacobson & Margolin, 1979) and integrative behavioral couple therapy (IBCT; Jacobson & Christensen, 1998). TBCT uses direct encouragement to increase positive behavior as well as specific training in communication and problem-solving skills to create positive change in couples. In contrast, IBCT focuses on increasing emotional acceptance, as well as direct change, in partners. IBCT assumes that relationship problems result not just from the egregious actions and inactions of partners but also in their emotional reactivity to those behaviors. Therefore, IBCT focuses on the emotional context between partners and strives to achieve greater acceptance and intimacy between partners as well as make deliberate changes in target problems. The Christensen et al. (2004) study showed statistically and clinically significant improvement in both conditions at the end of treatment, with IBCT showing more consistent improvement throughout treatment. At a two-year follow-up, Christensen, Atkins, Yi, Baucum, and George (2006) found that approximately two thirds of couples were improved relative to pretreatment according to clinical significance criteria (69% of IBCT couples and 60% of TBCT couples). There were few significant differences between treatments, but the differences that did emerge tended to favor IBCT.

The current investigation follows the couples from the above clinical trial five years after treatment termination. As noted earlier, TBCT has performed poorly in previous long-term investigations of outcome. The Jacobson et al. (1987) investigation above found that 56% of couples were unchanged or deteriorated by the two-year follow-up. The Snyder et al. (1991) investigation described above showed that 58% of couples were unchanged or deteriorated at a four-year follow-up (with 38% divorced). IBCT was developed, in part, to address concerns about long-term maintenance of gains (Jacobson & Christensen, 1998) through a focus on emotional acceptance and an emphasis on natural contingencies. For example, rather than teaching couples the “right way” to communicate and reinforcing that communication, as in TBCT, IBCT therapists process partners’ reactions to each other’s communication, letting those responses (natural contingencies) shape each other’s behavior. We hypothesized that these strategies might bring about more durable change, especially in severely distressed couples.

The current investigation addresses (a) the long-term outcome of behavioral couple therapy in general, (b) the relative outcome of TBCT and IBCT, (c) the relative outcome for severely versus moderately distressed couples, and (d) the trajectory of change throughout treatment and follow-up periods. Because we assessed couples repeatedly throughout the follow-up period as well as during treatment, we can examine the trajectory of change throughout the approximately five-and-a-half-year period of treatment and follow-up.

Method

Participants

A total of 134 chronically and seriously distressed married couples (268 individuals) were recruited in Los Angeles (71 couples) and Seattle (63 couples) for a clinical trial of couple therapy. Spouses were typically in their early 40s (wives’ mean age = 41.6 [SD = 8.6]; husbands’ mean age = 43.5 [SD = 8.7]), were moderately educated (17.0 years [SD = 3.2] of education for wives and 17.0 years [SD = 3.2] for husbands, counting kindergarten), had been married an average of 10.0 years (SD = 7.6), and had children (68 of the 134 couples had children). The majority of participants were Caucasian, but more than 20% of both husbands and wives were of another ethnicity. All procedures were approved by the Institutional Review Boards of the respective universities; at their intake session, couples signed an informed consent form, which covered their participation in treatment and in assessments through the 2-year follow-up; at the 2-year follow-up assessment, they signed an additional consent form for continuing participation in follow-up assessments. (See Christensen et al., 2004, for more detail on the participants.)

Questionnaire Measures

Relationship satisfaction. The Dyadic Adjustment Scale (DAS; Spanier, 1976), is a 32-item, well-validated questionnaire used to assess marital satisfaction. A recent meta-analysis of the reliability of DAS involving 91 studies that included 128 samples consisting of 25,035 participants (Graham, Liu, & Jeziorski, 2006) revealed a mean alpha for the total score of .92; in the current sample, alphas were .89 and .87 for husbands and wives, respectively. Technically speaking, the DAS assesses dyadic adjustment rather than dyadic satisfaction, but because the field typically refers to this measure as a measure of satisfaction, we continue that practice here.

Activities promoted by therapy. The Marital Activities Questionnaire (Christensen, 1999) consists of 10 items designed to measure how often clients continue to do the activities they presumably learned in IBCT (five items) and TBCT (five items). Most items refer to individual behavior, and partners rated on separate 9-point scales the extent to which each member engaged in the behavior (e.g., a TBCT item about using specific communication skills such as “I statements” or an IBCT item about empathizing with the partner). For joint items, such as the TBCT item “Set aside specific times to communicate with each other about the relationship or about problems we face,” partners gave one rating of the extent to which they did the behavior. Internal consistency for TBCT behaviors was .73 for men and .68 for women; for IBCT behaviors, it was .61 for men and .74 for women.

Additional treatment. The Therapy Information Sheet was designed to assess additional treatment, such as psychotherapy.

Phone Assessment

Relationship status. A brief phone interview was designed to assess current relationship status and satisfaction. If the couple was still married and living together with their spouse, the DAS-7, which is described in the next section, was administered. If the

Participants

A total of 134 chronically and seriously distressed married couples (268 individuals) were recruited in Los Angeles (71 couples) and Seattle (63 couples) for a clinical trial of couple therapy. Spouses were typically in their early 40s (wives’ mean age = 41.6 [SD = 8.6]; husbands’ mean age = 43.5 [SD = 8.7]), were moderately educated (17.0 years [SD = 3.2] of education for wives and 17.0 years [SD = 3.2] for husbands, counting kindergarten), had been married an average of 10.0 years (SD = 7.6), and had children (68 of the 134 couples had children). The majority of participants were Caucasian, but more than 20% of both husbands and wives were of another ethnicity. All procedures were approved by the Institutional Review Boards of the respective universities; at their intake session, couples signed an informed consent form, which covered their participation in treatment and in assessments through the 2-year follow-up; at the 2-year follow-up assessment, they signed an additional consent form for continuing participation in follow-up assessments. (See Christensen et al., 2004, for more detail on the participants.)

Questionnaire Measures

Relationship satisfaction. The Dyadic Adjustment Scale (DAS; Spanier, 1976), is a 32-item, well-validated questionnaire used to assess marital satisfaction. A recent meta-analysis of the reliability of DAS involving 91 studies that included 128 samples consisting of 25,035 participants (Graham, Liu, & Jeziorski, 2006) revealed a mean alpha for the total score of .92; in the current sample, alphas were .89 and .87 for husbands and wives, respectively. Technically speaking, the DAS assesses dyadic adjustment rather than dyadic satisfaction, but because the field typically refers to this measure as a measure of satisfaction, we continue that practice here.

Activities promoted by therapy. The Marital Activities Questionnaire (Christensen, 1999) consists of 10 items designed to measure how often clients continue to do the activities they presumably learned in IBCT (five items) and TBCT (five items). Most items refer to individual behavior, and partners rated on separate 9-point scales the extent to which each member engaged in the behavior (e.g., a TBCT item about using specific communication skills such as “I statements” or an IBCT item about empathizing with the partner). For joint items, such as the TBCT item “Set aside specific times to communicate with each other about the relationship or about problems we face,” partners gave one rating of the extent to which they did the behavior. Internal consistency for TBCT behaviors was .73 for men and .68 for women; for IBCT behaviors, it was .61 for men and .74 for women.

Additional treatment. The Therapy Information Sheet was designed to assess additional treatment, such as psychotherapy.

Phone Assessment

Relationship status. A brief phone interview was designed to assess current relationship status and satisfaction. If the couple was still married and living together with their spouse, the DAS-7, which is described in the next section, was administered. If the
couple was either separated or divorced from their spouse, that status was determined. For couples who had dropped out of the program and wished to no longer be contacted, we conducted an extensive Internet search to determine their relationship status. We contacted online public sources for their names and addresses and online public sources for any of their marriage and divorce records in the county and state in which they had last resided. If this search revealed two separate addresses for the spouses or a record of legal separation or divorce, the couple was classified as separated or divorced, respectively. If this search revealed the same address for both spouses, no record of legal separation or divorce, or no record of a subsequent marriage for either, the couple was classified as still being together. We were able to obtain relationship status on all couples at a five-year follow-up through either direct phone contact with one or both members (n = 122 couples), reports of divorce in a previous assessment (n = 3 couples), or Internet search without phone contact (n = 9 couples).

**Relationship satisfaction.** In order to assess relationship satisfaction over the phone, we used a brief, 7-item version of the DAS, the DAS-7, which is based on an analysis of the subset of DAS items that best discriminated between distressed and adjusted marriages (Sharpley & Cross, 1982). Subsequent studies have demonstrated internal consistency and validity for the DAS-7 (Hunsley, Best, Lefebvre, & Vito, 2001; Hunsley, Pinsent, Lefebvre, James-Tanner, & Vito, 1995; Sharpley & Rogers, 1984). In the current sample, alpha values for the DAS-7 at 5 years were .86 for wives and .85 for husbands.

**Procedures**

**Treatments.** After initial assessments, the 134 couples in this project were randomly assigned, stratified by initial distress (moderately distressed—66 couples; severely distressed—68 couples), to one of two conditions: TBCT (66 couples) or IBCT (68 couples). In both conditions, couples could receive a maximum of 26 sessions. The treatment manual for TBCT was Jacobson and Margolin’s (1979) classic monograph, supplemented by a shorter, updated, more succinct treatment manual (Jacobson & Christensen, 1994). Couples in TBCT were also assigned readings from the communication guide by Gottman, Notarius, Gonis, and Markman (1977). The treatment manual for IBCT was the monograph by Jacobson and Christensen (1998); couples were also given a self-help book about IBCT by Christensen and Jacobson (2000).

At the end of the treatment program, couples were encouraged to build on the gains they made in treatment and were discouraged from further therapy in the near term. They were prohibited from further contact with their therapist for two years.

**Schedule of assessments.** The great majority of assessments were scheduled according to the length of time since intake rather than according to a couple’s progress through therapy. Since each couple was offered a maximum of 26 weekly sessions, which would ideally last six months if sessions occurred weekly, a midtherapy assessment occurred 13 weeks after intake and a second assessment was scheduled 26 weeks after intake when couples would normally have finished therapy. Some couples completed therapy prior to the 26-week assessment, but most completed it after the 26-week assessment; therefore, an assessment was scheduled for couples immediately after the end of treatment regardless of when that occurred (all therapy had to be completed within a year of intake). For two years after the 26-week assessments, follow-up assessments were scheduled every six months based on time since the 26-week assessment (the idealized end of treatment). After the two-year evaluation, assessments were going to be conducted yearly. However, after the three-year assessment, additional funding became available, and assessments were conducted every six months at three and a half years, four years, and four and a half years. However, some couples had already passed some assessment points when this funding became available. A final assessment was conducted when couples were five years past their therapy termination date, whenever that occurred. See Figure 1 for a flowchart of assessments during the last three years of follow-up (flowcharts for treatment and the first two years of follow-up were reported in previous articles). A couple is included in this chart at an assessment point (and not listed as a drop-out or missing) if at least one of the two members provided data at that point.

**Type of assessments.** In-person assessments were conducted at intake, 13 weeks, 26 weeks, and two-year follow-up. Mailed assessments were conducted at termination, six-month, 12-month, and 18-months follow-ups. Phone assessments were conducted at all subsequent assessments. These phone assessments were buttressed with a mailed assessment at the five-year follow-up. Whereas at least one member of most couples provided both phone and mailed information at the five-year assessment point, some provided only one form of information and some provided neither; Figure 1 provides separate data for these two forms of assessment at the five-year point. Christensen et al. (2004) examined the data throughout treatment; Christensen et al. (2006) examined the data through the first two years of follow-up.

**Results**

**Relationship Satisfaction From Pretherapy Through Five Years Posttherapy**

One of the primary aims of this clinical trial was to examine how both treatments affected relationship satisfaction during and following therapy. The DAS was completed in full every 3 months throughout treatment (at intake, 13-week, and 26-week assessments), at the end of treatment (termination assessments), and every six months for two years after the treatment assessments (six-month, 12-month, 18-month, and 24-month follow-up assessments). Starting at the three-year follow-up, DAS-7 assessments were completed by phone every six months through the five-year follow-up. In order to obtain a common measure for all of these assessments, we extracted the DAS-7 from the full DAS for all of the assessments from intake through two-year follow-up. To assess the reliability of such an extraction procedure, we called couples the night before their two-year follow-up appointment to remind them of their appointment and to obtain a DAS-7 even though we were going to obtain the full DAS the next day. The correlation between this DAS-7 phone assessment and the subsequent DAS-7 extracted from the full DAS was high for both husbands (r = .89, p < .001) and wives (r = .88, p < .001).

Means and standard deviations for the DAS-7 from pretreatment through five years following therapy are presented in Table 1. The means showed robust improvement in relationship satisfaction for both therapies during treatment (d = 0.90 and d = 0.71 for IBCT
and TBCT, respectively) and for those providing data at five years, similarly strong effect sizes are found relative to pretreatment (d = 1.03 and d = 0.92 for IBCT and TBCT, respectively). Following treatment, relationship satisfaction drops somewhat in the first year for IBCT and the first two years for TBCT, but both treatments show improvements and maintenance in Years 3–5. However, these descriptive statistics do not incorporate information on separations and divorces.

A three-level multilevel model was fit to the DAS-7 data, with repeated measures nested within individuals, nested within couples (Atkins, 2005). Therapy condition (IBCT = 0, TBCT = 1), initial distress (moderate = 0, high = 1), and relationship status at five years posttherapy (together = 0, divorced/legally separated = 1) were entered as Level 3 dummy variables, and gender (male = 0; female = 1) was entered as a Level 2 dummy variable. The trajectory of relationship satisfaction from pretreatment through 5 years posttherapy is highly nonlinear. Various transformations of time were explored to find the best fitting model, including piece-wise linear, cubic splines, and polynomials. To choose the best representation of time, we used the Bayesian information criterion (BIC; Raftery, 1995). BIC can be used to select between alternative models.

Figure 1. Flow-chart of couples throughout the study. A One couple was rejected early on for domestic violence (DV), but data were collected from this couple at the 5-year follow-up. B Because of design, funding patterns, and timing problems, these couples were not assessed. C One couple was dropped early on, but data were collected from this couple at the 5-year follow-up.

1 Effect sizes were calculated as mean differences in average couple DAS-7 scores divided by standard deviation at pretreatment for each therapy separately.
models and weights the selection by the sample size and number of parameters in the model (i.e., model complexity). Using this procedure, we found that a sixth-order polynomial provided the best fit (i.e., lowest BIC value). Thus, time was represented across six terms, each raised to a successive power. Random effects included up to a cubic transformation of time at Level 3, whereas Level 2 included only a random-intercept.\(^2\)

Main effect coefficients are presented and interpreted below. Because the coefficients from the model related to time are challenging to interpret, we present \(F\) statistics for cross-level interactions with time and we plot the predicted regression lines. On average, women reported slightly more distress than did men, \(B = -0.84, SE = 0.36, t(130) = -2.31, p = .02, 95\% CI = [-1.57, -0.11].\(^3\) Those couples who go on to separate or divorce were far more distressed overall, \(B = -4.52, SE = 1.47, t(130) = -3.07, p < .01, 95\% CI = [-8.10, -2.19].\) Initial distress (moderate vs. severe) was used as a stratification variable in the randomization process. Not surprisingly, those couples who were initially highly distressed reported lower marital satisfaction on average, \(B = -3.16, SE = 0.57, t(130) = -5.55, p < .01, 95\% CI = [-4.29, -2.03].\) The two treatments were not significantly different from one another on average, \(B = -0.15, SE = 0.56, t(130) = -0.26, p = .79, 95\% CI = [-1.29, 0.95].\)

The predicted regression lines (averaging over gender) are shown in Figure 2. There is a strong, nonlinear effect of time over the five and a half years of the study, \(F(6, 2194) = 21.2, p < .01.\) Moreover, there is a strong cross-level interaction between relationship status (i.e., separated/divorced) and time, \(F(6, 2194) = 11.4, p < .01,\) and also between initial severity and time, \(F(6, 2194) = 11.3, p < .01.\) Gender, \(F(6, 2194) = 1.5, p = .17,\) and therapy, \(F(6, 2194) = 1.7, p = .13,\) do not reveal significant cross-level interactions with time. As can be seen in Figure 2, couples who received IBCT reported higher marital satisfaction during the early phases of follow-up. Contrasts between the two therapies show significant differences between the therapies between 6-month and 24-month assessments. These differences drop away in later phases of therapy. Couples who go on to divorce start treatment somewhat more distressed, improve little, and show significant deterioration following the end of therapy. Note also that there are relatively few data points among the divorced group in the later phases of follow-up and, thus, the regression lines for these groups are only projected out to 225 weeks.

### Endpoint Analyses: Satisfaction and Relationship Status at Five Years

At the final endpoint, treatment differences were examined in relationship satisfaction using the full DAS, clinical significance based on the full DAS, and divorce status. At five years post-therapy, couples who had received IBCT reported an average of

---

\(^2\) Although this is a complex model, the data are clearly nonlinear, with more than a single bend. Moreover, close to 2,500 data points were included in the model, with up to 13 repeated measures per individual and 26 per couple. Orthogonal, polynomial contrasts were used to aid model fitting and convergence.

\(^3\) Because we used orthogonal polynomials, the current listing of main effects reflects differences at approximately the midpoint of the time course.
96.2 (SD = 18.9) on the full DAS, compared with 96.6 (SD = 20.3) for couples who had received TBCT, $B = 0.38$, $SE = 4.00$, $t(82) = 0.09$, $p = .92$. Using the methods described in Jacobson and Truax (1991), clinical significance outcomes from pretreatment to five years posttherapy were calculated for both treatments and presented in Table 2. Couples who divorced were considered deteriorated. Approximately one third of couples across both treatments remain recovered at five years posttherapy (i.e., significantly improved and in the functional range of relationship satisfaction). Roughly half of couples are either reliably improved or recovered, meaning roughly half are either unchanged or deteriorated, with the majority of deteriorated couples being divorced.

### Table 2

<table>
<thead>
<tr>
<th>Clinical significance</th>
<th>Therapy condition</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TBCT</td>
<td>IBCT</td>
</tr>
<tr>
<td>Deteriorated</td>
<td>24</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>39</td>
<td>36</td>
</tr>
<tr>
<td>Unchanged</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td>Improved</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>19</td>
</tr>
<tr>
<td>Recovered</td>
<td>20</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>33</td>
<td>31</td>
</tr>
<tr>
<td>Sum</td>
<td>61</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

*Note.* TBCT = traditional behavioral couple therapy; IBCT = integrative behavioral couple therapy.
Considering divorce directly, at five years posttherapy, 19 out of 68 (27.9%) couples who had received TBCT were divorced or legally separated, whereas 17 out of 66 (25.7%) couples who received IBCT were divorced or legally separated. Using a test of independent proportions, these divorce rates are not significantly different, $\chi^2(1) = 0.08, p = .78$.

**Termination Versus Endpoint Status**

Using the clinical significance categories above, we examined the link between relationship status at the termination of therapy and relationship status at the five-year point. Table 3 has cross-tabulations of clinical significance status at posttherapy (in rows) with clinical significance status at five years posttherapy. The diagonal elements of Table 3 show couples whose status is similar across follow-up, and a weighted kappa shows that the table reflects some association ($K = 0.21, Z = 3.74, p < .01$). Although kappa shows that there are greater than chance levels of agreement, there is notable disagreement. Cells below the main diagonal are much larger than those above, reflecting overall deterioration from posttherapy to five years posttherapy. To model this, we fit a quasi-symmetry model for square ordinal tables (Agresti, 2002), which provides an average odds ratio describing the directional drift from the main diagonal in a square table. This analysis, based on a logistic regression, found significant deteriorating drift ($B = 0.63, Z = 3.9, p < .01$), showing that, on average, couples were 1.9 times more likely to worsen in their clinical significance category than improve.

Table 3 also has the conditional probabilities of five-year status given termination status. A full 87% of couples classified as deteriorated at post were classified as deteriorated at five years, whereas only 7% of people classified as deteriorated at post were recovered at five years. A majority of couples (55.8%) who were clinically significantly improved at termination (reliably improved or recovered) were also clinically significantly improved at five years. Only 12% of those unchanged at termination are still unchanged at five years, whereas 28% are recovered and 44% deteriorated.

**Comparisons With Cookerly (1980) and Snyder et al. (1991)**

At five years posttherapy, Cookerly reported a 43.6% divorce rate for couples who had received conjoint couple therapy and a 70.2% divorce rate for couples who had received nonconjoint therapy. We compared the current divorce rate of 26.8% (averaging across treatments) with those estimates using a two-sample test of proportions (Agresti, 2002). For both estimates, the null hypothesis that the current divorce rate comes from a similar population is rejected, $\chi^2(1) = 14.6, p < .01$, and $\chi^2(1) = 118.2, p < .01$, for conjoint and nonconjoint divorce rates, respectively.

Using a similar strategy, we compared the current sample’s divorce rate with Snyder et al.’s (1991) estimates at four years posttherapy of 3% for insight-oriented marital therapy (IOMT) and 38% for behavioral marital therapy (BMT). Again, null hypotheses that the current sample estimates come from a similar population as Snyder et al.’s BMT condition, $\chi^2(1) = 7.1, p < .01$ or IOMT condition, $\chi^2(1) = 217.5, p < .01$ are rejected. Thus, the observed divorce rate in the current study is significantly below the BMT condition of Snyder et al. yet above their IOMT condition.

Snyder et al. (1991) used the Global Distress Scale of the Marital Satisfaction Inventory (Snyder, 1981) as a primary outcome measure, whereas in the current study, couples completed the Global Distress Scale of the Marital Satisfaction Inventory—Revised (Snyder, 1997), a similar and comparable measure. We compared Global Distress Scale scores in the IOMT condition at intake from Snyder et al.’s (1991) with Global Distress Scale scores from our IBCT and TBCT conditions at intake. Their IOMT condition was significantly less distressed ($M = 63.1, SD = 5.7$) than either of the current treatment samples: $M = 67.2, SD = 4.6$, $t(96) = 2.59, p < .01$ for TBCT and $M = 66.5, SD = 4.3$, $t(94) = 2.18, p = .01$, for IBCT. In their IOMT condition, one third (10/30) of couples would be classified as nondistressed at the start of treatment (a couple GDS score of $T < 59$), whereas in the current data, only 2.9% (2/68) of couples in TBCT and 1.5% (1/66) of couples in IBCT would be classified as nondistressed. Clearly the sample of couples in the IOMT condition of Snyder et al. (1991) is from a less distressed population of couples than those in the current study. Nonetheless, the outcome results (3% divorced at four years) in the Snyder et al. IOMT condition are impressive.

**Activities Promoted by Therapy at Five Years**

The MAQ consists of two subscales that measure marital behaviors promoted by TBCT and IBCT. At five years posttherapy, couples treated with IBCT and TBCT reported no significant differences in the level of TBCT behaviors (IBCT: $M = 3.1, SD = 1.2$; TBCT: $M = 3.1, SD = 1.5$) or in the level of IBCT behaviors (IBCT: $M = 4.1, SD = 1.2$; TBCT: $M = 4.1, SD = 1.5$) that they reported. There were significant differences in couples from both treatments across levels of clinical significance at five years posttherapy. Specifically, couples who were clinically recovered (as opposed to those classified as deteriorated, unchanged, or improved) at five years were more likely to report higher levels of IBCT behaviors ($M = 4.1, SD = 1.1$) and TBCT behaviors ($M = 3.1, SD = 1.4$) than those who were unchanged or deteriorated (IBCT behaviors: $M = 3.5, SD = 0.8, p < .01$; TBCT behaviors: $M = 2.6, SD = 1.0, p < .01$).

---

4 Conversely, the pretreatment Global Distress Scale means for Snyder et al.’s BMT condition are indistinguishable from either treatment condition in the current study at pretreatment: TBCT: $M = 4.0, SD = .66$; IBCT: $M = 3.9, SD = .62$. Thus, although Snyder et al.’s two treatment conditions were not significantly different from each other at pretreatment, and they used analysis of covariance to account for existing pretreatment differences, their IOMT condition was significantly less distressed than our treatment conditions at pretreatment, whereas their BMT condition was not.

5 We are grateful to Douglas Snyder for retrieving these data from archival records. Although the Global Distress Scales of the MSI and MSI-R differ, analyses during development of the revised Global Distress Scale indicate a correlation of .98 with the original, longer scale (Snyder, 1997).
Additional Therapy

During the five years following therapy, 25 couples reported further marital therapy\(^6\); 18 of these couples had been in the IBCT condition and 7 had been in the TBCT condition \((p < .02\) according to Fisher’s exact test). There was a moderate level of disagreement between husbands and wives about how much additional marital therapy the couple received. Of the 25 couples that reported additional marital therapy, five \((20\%)\) disagreed about whether they received any additional marital therapy. In examining therapy other than marital therapy, five \((20\%)\) disagreed about whether they received additional marital therapy. In examining therapy other than marital therapy \((e.g.,\) individual, family), \(65\%\) of individuals who received IBCT and \(41\%\) of individuals who received TBCT reported receiving some type of therapy in the five years following treatment termination. Using a mixed-effects logistic regression, we obtained an odds ratio of 2.68 comparing TBCT with IBCT, \(B = 0.99, n(132) = 3.19, p = .01, 95\%\) CI for odds ratio = [1.46, 4.95]) and an odds ratio of 0.62 comparing women with men, \(B = -0.48, n(132) = -2.05, p = .05, 95\%\) CI for odds ratio = [0.39, 0.98]. Specifically, these results indicate that individuals who received IBCT were significantly more likely to be involved in any type of therapy, and women were more likely than men to be involved in any type of therapy.

We also examined whether couples who had received marital therapy during the five-year follow-up phase reported differences in marital satisfaction relative to those who did not report receiving additional marital therapy. Because relatively few couples reported receiving additional marital therapy, we examined the data descriptively. At the end of therapy, there was very little difference between couples who received additional marital therapy during follow-up \((M = 95.6, SD = 15.3)\) and couples who did not seek additional marital therapy \((M = 95.6, SD = 20.1)\) based on the average DAS score for a couple. In addition, three of the 25 couples who sought additional marital therapy separated or divorced during the five years following therapy.

Finally, we examined differences in average DAS scores between couples where individuals reported receiving additional therapy other than marital therapy versus those who did not report receiving any additional therapy. Results of an ANOVA revealed a nonsignificant difference, \(F(1, 82) = .69, p < .42,\) between couples that reported receiving additional therapy other than marital therapy \((M = 93.3, SD = 13.9)\) and couples that did not report receiving any additional therapy other than marital therapy \((M = 96.9, SD = 21.7)\). Because some spouses reported individual and/or couple therapy of very short duration, five or more sessions of a given form of therapy was required in order for it to be counted as additional therapy in the analyses presented here.
half of couples (50.0% of IBCT couples and 45.9% of TBCT couples) showed clinically significant improvement (i.e., reliable improvement or recovery) at five-year follow-up compared with their status at pretreatment. Approximately one fourth of couples were separated or divorced at five-year follow-up (25.7% of IBCT couples and 27.9% of TBCT couples). Differences between TBCT and IBCT were not statistically significant. A small number of couples sought additional couple therapy, and a small number of individuals sought individual therapy during the follow-up period. There was no evidence that those who sought this additional therapy ended up with higher satisfaction scores. However, finding effects of additional therapy is complicated by (a) the small numbers who sought additional therapy and (b) the selection biases (those who seek additional therapy may be more distressed when they seek it).

The initial article on this clinical trial provided evidence that IBCT and TBCT were quite different treatments and that both were competently delivered (Christensen et al., 2004). Data at termination (Christensen et al., 2004) and 2-year follow-up (Christensen et al., 2006) further indicated statistical but not dramatic superiority in outcome for IBCT versus TBCT. The trajectories of satisfaction reported in this article also revealed significant differences between these treatments during the first two years of follow-up. It is perhaps not surprising that these differences between treatments evaporate after five years of follow-up assessments but with no intervening treatment. Perhaps a stronger treatment effect and a continuing superiority of IBCT over TBCT could have been maintained with booster sessions.

How can couple therapy be made more efficacious? Certainly, we cannot help all couples; some are better off apart. However, we have probably not reached a ceiling effect. The way to further improvement in couple therapy is probably through research on the mechanisms of change. As a number of others have pointed out (Doss, 2004; Kazdin, 1999), research on the mediators and ultimately the mechanisms of change offer the possibility of targeting the active ingredients in our multifaceted treatments and thus opening the way to making those ingredients even more potent. In previous research with the current sample, we have provided evidence for the mediating role of both behavior change and acceptance of behavior (Doss, Thum, Sevier, Atkins, & Christensen, 2005). Clearly more research on these likely mediators of change is needed. We have further shown with this sample that emotional arousal during problem-solving communication interacts with severity level to predict outcome (B. R. Baucom, Atkins, Simpson, & Christensen, 2009). Given that the emotional arousal of seriously distressed couples is often high and dysregulated, a further focus on arousal seems important. It is possible that the powerful factor of exposure, so important in the treatment of emotional disorders in general (Barlow, Allen, & Choate, 2004), is important in the treatment of couple problems. At the beginning of treatment, a couple may not be able to constructively address a “hot” issue, but through the safety and validation that couple therapy provides, this couple may become emotionally calmer in addressing this and other “hot” issues. Because we know so little about what in all we do actually has an impact on couples, further research in this area is likely to elucidate avenues for greater treatment efficacy.

Because there was no control group in this clinical trial, one might question whether the outcomes obtained with treatment and through five years of follow-up could have occurred without any intervention. That possibility seems unlikely at best, as a meta-analysis by D. H. Baucom, Hahlweg, & Kuschel (2003) has shown that control groups of couples in clinical trials of couple therapy do not change over time. Furthermore, the couples in this study were chronically and seriously distressed couples who might be less likely than other couples to change over time without intervention. To gauge the comparative power of this treatment versus other treatments for couples, we compared our effects with those from other studies. A meta-analysis by D. H. Baucom et al. (2003) based on 17 published investigations of randomized clinical trials of couple therapy estimated the mean, termination effect size on relationship satisfaction as $d = 0.82$. Shadish and Baldwin (2005) estimated an effect size of $d = 0.59$ based on a larger pool of 30 randomized clinical trials, which included both published investigations as well as unpublished dissertations. Certainly our termination effect sizes of $d = 0.90$ and $d = 0.71$ for IBCT and TBCT, respectively, compare quite favorably with these meta-analytic results.

There are only a handful of long-term follow-ups of couple therapy but all have included a common metric of relationship status, that is, number of couples separated or divorced at follow-up. Our separation–divorce rate of 26.8% (averaged across treatments) was substantially and statistically significantly lower than Cookerly’s (1980) 43.6% divorce rate at 5 years posttreatment for couples who had received conjoint couple therapy and their 70.2% divorce rate for couples who had received nonconjunct therapy. Similarly, our separation–divorce rates were significantly lower than Snyder et al.’s (1991) four-year separation–divorce rate of 38% for their BMT condition. In contrast, our separation–divorce rates were significantly higher than Snyder et al.’s impressive 3% for their IOMT condition, but couples in that condition were significantly less distressed at intake than couples in the present study. Thus, it seems that the outcomes obtained in the current study are as good as or better than those obtained in most other treatment studies.

We also examined the link between outcome at termination and outcome at five-year follow-up. Not surprisingly there was a statistically significant connection between those outcomes. Couples who were deteriorated at termination were most likely to remain deteriorated at five-year follow-up (87% stayed in the same category). A majority of couples who achieved a positive outcome at termination (reliable improvement or recovery) were in either of those positive categories at five-year follow-up. Given the power of inertia, one might think that the category of “unchanged” at termination would be the most stable state (if months of therapy do not change the couple, perhaps they are doomed to stable dissatisfaction). Interestingly, however, the unchanged category showed the most volatility: Perhaps the months of therapy and its unsatisfactory immediate outcome provoked some change up or down. Only 12% of those unchanged at termination are still unchanged at five years, whereas 28% are recovered and 44% deteriorated. When there was movement away from the termination outcome, it was more likely to be toward a worse outcome than a better outcome. This significant drift toward deterioration suggests again the possible value of booster session interventions to alter the downward slide evidenced by a number of couples.

The current study has a number of important methodological strengths. First, it is one of only a few studies to examine the
long-term outcome of couple therapy. Second, it is the first study to repeatedly assess couples so that the trajectory of change during treatment and follow-up can be examined. Third, the study was able to retain a substantial portion of couples by maintaining frequent contact with couples (approximately every six months through the five-year follow-up period) and paying them an increasing amount for mailed assessments. In over 90% of couples, at least one partner in that couple provided data at the five-year follow-up. Furthermore, using Internet sources and court records, we obtained relationship status at five years on all couples.

The study, of course, has its limitations as well. First and foremost is the reliance on self-report measures. Although observational measures were taken earlier in the study (e.g., Sevier, Eldridge, Jones, Doss, & Christensen, 2008), there were no observational measures taken during the last three years of the follow-up. While certainly desirable, the costs of such an effort, given the wide dispersion of our couples, were prohibitive. However, the primary outcome in studies of couple therapy is relationship satisfaction and status, which are normally assessed with self-report measures. A second limitation of this study concerns the timing of follow-up measures about 3–4 years into the study. A number of couples had passed the desired assessment points when additional funding finally became available. Thus, the three-and-a-half-year follow-up in particular is missing a number of couples who were then brought in for later assessments.

Nonetheless, this study is the first to show a detailed trajectory of change in relationship satisfaction over the course of couple therapy and long-term follow-up. Given that the sample was chosen because of its serious and chronic distress (nearly a hundred couples who desired treatment were excluded because they did not meet the criteria of serious and repeated distress; see Christensen et al., 2004), the outcomes are encouraging. Five years after the conclusion of therapy, half of the couples demonstrated clinically significant change compared with their pretreatment assessment. Only one fourth were separated and divorced. Both the immediate and long-term outcomes in this study were equal to or superior to those obtained by most other clinical trials of couple therapy.

References

Received June 29, 2009
Revision received October 7, 2009
Accepted October 12, 2009

Additional Journal Information

**Manuscripts:** Effective in January 2010, the Incoming Editor, Arthur M. Nezu, is receiving all new submissions to the journal. Submissions that are accepted will be published beginning in the 2011 volume. Authors are advised to consult the manuscript submission guidelines for JCCP (http://www.apa.org/pubs/journals/ccp/). Manuscripts that do not conform to the submission guidelines may be returned without review. All submissions, whether new, a first-submission manuscript, or a revision of a previously submitted manuscript must use the electronic submissions portal (http://www.apa.org/pubs/journals/ccp). If you encounter difficulties with submission, please e-mail Sara Young at syoung@apa.org or call 202-336-5859. General correspondence may be directed to the Editorial Office via e-mail at amn23@drexel.edu. The opinions and statements published are the responsibility of the authors, and such opinions and statements do not necessarily represent the policies of APA or the views of the Editor.

**Copyright and Permission:** Those who wish to reuse APA-copyrighted material in a non-APA publication must secure from APA written permission to reproduce a journal article in full or journal text of more than 800 cumulative words or more than 3 tables and/or figures. APA normally grants permission contingent on permission of the author, inclusion of the APA copyright notice on the first page of reproduced material, and payment of a fee of $25 per page. Libraries are permitted to photocopy beyond the limits of U.S. copyright law: (1) post-1977 articles, provided the per-copy fee in the code for this journal (0022-006X/10/$12.00) is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923; (2) pre-1978 articles, provided that the per-copy fee stated in the Publishers’ Fee List is paid through the Copyright Clearance Center. For more information along with a permission request form go to: www.apa.org/about/contact/copyright/index.aspx

**Electronic Access:** Individual subscribers to this journal have automatic access to all issues of the journal in the PsycARTICLES® full-text database. See http://www.apa.org/pubs/journals/ccp and click on View Table of Contents.

**Reprints:** Authors may order reprints of their articles from the printer when they receive proofs.

**Single Issues, Back Issues, and Back Volumes:** For information regarding single issues, back issues, or back volumes, visit www.apa.org/pubs/journals/subscription.aspx or write to Order Department, American Psychological Association, 750 First Street, NE, Washington, DC 20002-4242; call 202-336-5600 or 800-374-2721.

**Subscription Claims Information:** A claim form to assist members, institutions, and nonmember individuals who have a problem with their subscription is available at http://forms.apa.org/journals/subclaim/; or call 1-800-374-2721.

**Microform Editions:** For information regarding microform editions, write to University Microfilms, Ann Arbor, MI 48106.

**Change of Address:** Send change of address notice and a recent mailing label to the attention of Subscriptions Department, APA, 30 days prior to the actual change of address. APA will not replace undelivered copies resulting from address changes; journals will be forwarded only if subscribers notify the local post office in writing that they will guarantee periodicals forwarding postage.

**APA Journal Staff:** Susan J. A. Harris, Senior Director, Journals Program; Paige W. Jackson, Director, Editorial Services; Annie Hill, Senior Editorial Supervisor; Tyler Krupa, Lead Editor; Skip Maier, Director, Journal Services; Richard Pepple, Account Manager; Jodi Ashcraft, Director, Advertising Sales and Exhibits.