Individual Beck Depression Inventory (BDI) time courses of 38 clients receiving cognitive therapy (CT) and a modified form of CT were studied in order to investigate temporal changes during CT for depression. The primary aim was to determine if alternative methods of defining and computing gains occurring early in CT would alter the conclusions drawn in the current literature. Three types of gains were studied: sudden gains (previously studied sudden, substantial, and stable improvements in depression during 1 between-session interval after Session 2), first-session gains (occurring after first sessions), and pretreatment gains (occurring after pretreatment assessments). Positive outcomes were predicted by first-session gains and by sudden gains occurring in the first half of treatment, highlighting the importance of early change in CT for depression.

Cognitive Therapy (CT; Beck, Rush, Shaw, & Emery, 1979) is the most widely used and well researched of the empirically supported psychological treatments for depression, and its efficacy is well-established (DeRubeis & Crits-Christoph, 1998). However, the exact mechanisms by which CT accomplishes change are still unclear. In efforts to isolate active techniques and mediators of change in CT, several investigators have examined the temporal course of change during treatment.

One set of investigators has focused on rapid early response to treatment. For example, Rush, Kovacs, Beck, Weissburger, and Hollon (1981) found in their sample that over 60% of total symptom reduction occurred by Week 4, and Fennell and Teasdale (1987) found in their sample that more than 66% of overall improvement occurred by Week 3. Fennell and Teasdale also found that clients who experienced at least a 50% reduction in Beck Depression Inventory (BDI; Beck, Ward, Mendelsohn, Mock, & Erbaugh, 1961) scores in the first 2 weeks of treatment had a 100% recovery rate, while those who did not had an 11% recovery rate. Similarly, Blackburn and Bishop (1983) reported that outcomes at the end of treatment could be predicted from scores obtained in Weeks 3 and 4.

Ilardi and Craighead (1994), in a review of this literature, concluded that 60% to 80% of symptom reduction occurs in the first 4 weeks of CT. They reasoned that nonspecific factors must be responsible for these changes because the application of specific cognitive techniques does not occur until later in treatment. The therapy rationale, therapeutic alliance, and assignment of homework were among the nonspecific factors Ilardi and Craighead argued could play a role in this disproportionate early symptom reduction.

Tang and DeRubeis (1999a, 1999b) challenged Ilardi and Craighead’s (1994) conclusion. First, Tang and DeRubeis (1999a) pointed out that eight sessions often occur in the first 4 weeks of CT, making the rapid early response in CT seem less remarkable. Second, they cited several studies (DeRubeis & Feeley, 1990; Feeley, DeRubeis, & Gelfand, 1999) indicating that cognitive techniques are used substantially as early as Session 2. Thus, symptom reduction early in treatment should not be
interpreted as evidence only for the role of nonspecific factors. Third, they suggested that Ilardi and Craighead’s analysis of the group BDI time course was not appropriate because this method obscured significant individual session-to-session symptom reductions. Instead, Tang and DeRubeis argued that the individual time course for each client needs to be assessed in order to identify sessions that precede significant change. Then, particular sessions that precede change can be studied to determine what types of interventions were used.

Using this strategy, Tang and DeRubeis (1999a, 1999b) identified sudden gains: sudden, substantial, and stable improvements that occur in one between-session interval. For an interval to be called a sudden gain it had to meet three criteria: (a) a drop of 7 or more BDI points (the gain is large in absolute terms); (b) the drop is 25% or more of the pre-drop score (the gain is large in relative terms); and (c) a significant difference on an independent t test comparing the mean BDI score for three sessions before the gain to the mean score for the three sessions after the gain (the gain is stable). This third criterion resulted in the exclusion of changes occurring after the pretreatment assessment and immediately after the first session from the sudden gains analysis. As discussed later, the exclusion of these time points from the analysis has implications for the interpretation of Tang and DeRubeis’ findings.

Tang and DeRubeis (1999b) found that 39% of clients experienced sudden gains, the mean sudden gain magnitude was 11 BDI points, and the median pre-gain session was Session 5. In addition, sudden gains were preceded by client cognitive changes, were followed by improvements in the therapeutic alliance, and predicted better outcomes. Based on these findings, they presented a model of change in which sudden gains are triggered by critical sessions occurring relatively early in treatment, but after an initial “preparation stage” during which only minor symptom improvement occurs. They presented this as evidence that cognitive modification in CT is indeed important.

The sudden gains found by Tang and DeRubeis (1999b) were named as they are above simply to distinguish them from the gains investigated in the past, but it is important to note that they are just as sudden as sudden gains. We then explored the relationship

---

1 Tang and DeRubeis (1999b) did not include pretreatment BDI scores in their analysis. Thus, gains occurring after the pretreatment assessment and immediately after Session 1 were excluded from the analysis because the t test requires at least two data points in each group. Gains immediately after Session 2 were assessed but with the two pre-gain data points compared to the three post-gain data points.
between the occurrence of these three types of gains and acute treatment outcomes.

Method

Data Source

Data were obtained from a nonrandomized clinical trial that compared CT and a modified form of CT enhanced by techniques from Functional Analytic Psychotherapy (FAP) known as FAP-Enhanced Cognitive Therapy (FECT). Kohlenberg, Kanter, Bolling, Parker, and Tsai (2002) provide more details regarding this data set, the treatment conditions, and treatment outcome.

Participants

The current analyses were based on the 15 CT and 23 FECT clients who completed the study (treatment lasted 20 sessions, although 3 clients who completed 14, 17, and 18 sessions were classified as completers). Participants’ mean age was 42 years (SD = 9.7); 61% were female, 41% were married or living with someone, and 47% had graduated from a 4-year college. Eligibility criteria were a diagnosis of major depressive disorder according to the Structured Clinical Interview for the Diagnostic and Statistical Manual of Mental Disorders (Spitzer, Williams, & Gibbon, 1995) and a score of 18 or greater on the BDI. Exclusion criteria included a number of concurrent psychiatric disorders (bipolar or psychotic subtypes of depression, panic disorder, current alcohol or other substance abuse, past or present schizophrenia or schizophreniform disorder, organic brain syndrome, and mental retardation). Participants who were in some concurrent form of psychotherapy, who were receiving psychotropic medication, or who needed to be hospitalized because of imminent suicide potential or psychosis were also excluded.

Conditions

Treatment in both CT and FECT consisted of 20 sessions of CT as per the manuals provided by Beck and colleagues (1979) and J. Beck (1995). The FECT condition expanded on standard CT in that therapists were encouraged to pay particular attention to client problems and improvements as they occurred in the therapeutic relationship, as per guidelines described in FAP (Kohlenberg & Tsai, 1991; Kohlenberg et al., 2002, 2004). Specifically, therapists in FECT conducted standard CT with an added case conceptualization that described how client problems might occur in the relationship with the therapist. Therapists were encouraged to look for the occurrences of these problems and perform CT interventions on them. For example, a client with the dysfunctional belief “I’m a loser” might think that the therapist thinks she is a loser. The therapist would then conduct cognitive restructuring interventions on this manifestation of the belief as it occurred live in the therapeutic relationship. The principles and theory of FAP (Kohlenberg & Tsai, 1991) suggest that these interventions might be particularly powerful in creating change. There were no specific changes in the temporal sequence of interventions in FECT compared to CT.

Adherence analyses using a modified form of the Collaborative Study Psychotherapy Rating Scale—Version 6 (CSPRS-6; Hollon et al., 1987) showed that FECT resulted in an increased focus on the therapeutic relationship compared to CT. However, there were no significant differences in the amount or types of cognitive interventions performed in CT and FECT (Kohlenberg et al., 2002). In other words, while the FECT condition added a focus on the therapeutic relationship, cognitive techniques as per Beck and colleagues (1979) were still the primary intervention.

Therapists

Four experienced cognitive therapists provided therapy for both conditions. All therapists had been in practice for at least 10 years and had served as CT research therapists on prior clinical trials. Three of the therapists attended a 2-day workshop by Dr. Steve Hollon as a refresher in preparation for the FECT study, and two were board certified by the Academy of Cognitive Therapy. Dr. Keith Dobson rated 24 CT and 28 FECT sessions for competency on the Cognitive Therapy Scale (CTS; Vallis, Shaw, & Dobson, 1986) and ongoing feedback based on these ratings was provided to the therapists. There were no significant differences in CTS total scores between therapists or between conditions, nor was there a significant Therapist × Condition interaction. The mean CTS total score was 43.58 (SD = 6.00), which is considered adequate and comparable to other studies (Shaw, 1984).

Measurement

The BDI was given at the pretreatment assessment (1 to 2 weeks before the first therapy session) and weekly throughout treatment. The final session BDI (usually Session 20) was used as the end-of-treatment BDI.

The sudden gains criteria of Tang and DeRubeis (1999b) were altered to determine pretreatment and first-session gains. Like sudden gains, both pretreatment gains and first-session gains should be large in both absolute and relative terms, so the first two of Tang and DeRubeis’ criteria were maintained (the drop in BDI must be 7 or more points
and the size of the gain must be at least 25% of the pre-gain BDI score). Applying Tang and DeRubeis’ third criterion (a significant difference on an independent t test comparing the mean of the three sessions before the gain to the mean of the three sessions after the gain) to these intervals was impossible because there were no sessions before the gain in the case of pretreatment gains and only one session before the gain in the case of first-session gains. For the pretreatment gains, we used the criteria developed by Gaynor and colleagues (2003), which required that 50% of the gain be maintained for at least two sessions following the gain. For the first-session gains, we conducted the same independent t test analysis as Tang and DeRubeis (1999b), but used the pretreatment assessment BDI and first session BDI as the two before-gain data points.

Our criteria for sudden gains (described above), gain reversals (an increase in BDI at any point after the gain that erases half or more of the gain), and recovery (an end-of-treatment BDI of 10 or less) were identical to those of Tang and DeRubeis (1999b).

There are two minor differences (other than the existence of the FECT condition) between our data set and that of Tang and DeRubeis (1999b). First, in calculating total BDI change over the course of therapy, we used pretreatment BDIs as our first data point while they used first session BDIs. This was done in order to facilitate comparisons between sudden, pretreatment, and first-session gains. Second, in our data set the application of Tang and DeRubeis’ third criterion to the interval between the second-to-last and the last session was not possible because we did not have a BDI for the week following the last session. However, this did not affect the current analysis because no client in our sample met the first two criteria for that final interval.

Results

Comparison of CT and FECT
As described in Kohlenberg and colleagues (2002), there were no significant differences between CT and FECT in pretreatment BDIs, end-of-treatment BDIs, percentage of responders (at least 50% reduction in BDI from pretreatment to end-of-treatment), or percentage of those classified as “recovered” (end-of-treatment BDI of 10 or less).

Our current analyses showed that there were no significant differences between CT and FECT in the percentage of clients with pretreatment gains (CT = 20%, FECT = 13%), first-session gains (CT = 20%, FECT = 13%), or sudden gains (CT = 33%, FECT = 47%). Similarly, the average magnitudes of pretreatment gains (CT M = 12.33, SD = 4.62; FECT M = 11.00, SD = 5.29), first-session gains (CT M = 11.33, SD = 5.13; FECT M = 14.00, SD = 6.00), and sudden gains (CT M = 9.20, SD = 3.35; FECT M = 9.27, SD = 2.80) did not differ between conditions. Finally, the median pre-sudden-gain session was Session 10 for both CT and FECT. While significant differences were found between CT and FECT in other areas (e.g., interpersonal functioning; Kanter, Schildcrout, & Kohlenberg, 2005; Kohlenberg et al., 2002), these analyses suggest that the treatments were not differentially effective on the BDI and did not produce differential patterns of pretreatment, first-sesson, and sudden gains. Therefore, CT and FECT clients were combined into one group for all subsequent analyses.

Pretreatment Gains and First-Session Gains
Six clients (16%) experienced pretreatment gains, six (16%) experienced first-session gains, and 26 (68%) experienced neither a pretreatment gain nor a first-session gain. The first three rows of Table 1 show the mean BDI scores for these three groups. Two clients who experienced pretreatment gains also experienced sudden gains, and two clients who experienced first-session gains also experienced sudden gains. The mean pretreatment gain was 11.66 BDI points (SD = 4.13) and accounted for 66% of the total recovery from pretreatment to the last session. The mean first-session gain was 12.66 BDI points (SD = 4.23) and accounted for 50% of the total recovery from assessment to the last session. Three of the six clients with pretreatment gains and three of the six clients with first-session gains experienced reversals.

An analysis of covariance (ANCOVA, controlling for pretreatment scores) revealed that end-of-treatment scores were significantly different, F(2, 34) = 8.70, p = .001. Follow-up contrasts revealed that those with first-session gains had lower end-of-treatment scores than those with neither a first-session nor a pretreatment gain (p < .001) and those with pretreatment gains (p = .021).

Four of six clients (67%) with pretreatment gains, six of six clients (100%) with first-session gains, and 12 of 26 clients (46%) with neither a pretreatment gain nor a first-session gain recovered at end-of-treatment. Recovery rates for those with

---

2 The analyses reported below were also run using first session scores instead of pretreatment scores as the first data point and these analyses did not produce different conclusions regarding first session and sudden gains.
pretreatment gains, first-session gains, and neither were significantly different ($\chi^2 = 6.02, p = .049$). Fisher exact tests revealed that significantly more of those with first-session gains recovered than those with neither a first-session nor a pretreatment gain ($p < .024$). There was no significant difference in recovery rates between those with pretreatment gains and those with neither kind of gain.

**Sudden Gains**

A sudden gain was experienced by 16 of the 38 clients (42%) and no client had more than one. Of the 22 clients who did not experience a sudden gain, 4 experienced a pretreatment gain and 4 experienced a first-session gain. The median sudden gain occurred after Session 10 and ranged from after Session 3 to after Session 18. The average course of the three sessions before and the three sessions after sudden gains is shown in Figure 1. As shown in Table 1, the average sudden gain was 9.25 BDI points ($SD = 2.86$) and accounted for 51% of the recovery from pretreatment to end-of-treatment. Seven of the 16 (44%) clients who experienced sudden gains experienced reversals. Eight of the 16 (50%) clients who experienced sudden gains were recovered at end-of-treatment.

**Comparison of clients with sudden gains to clients with no sudden gains.** Using an independent samples $t$ test, pretreatment BDI scores of clients with sudden gains were not significantly different from those of clients without sudden gains (last two rows of Table 1). To analyze change in BDI over the course of treatment, we conservatively controlled for pretreatment scores using an ANCOVA and found no significant differences in last session scores between groups. Finally, using a chi-square analysis, the recovery rate of those who did not have sudden gains (14/22, 64%) was not significantly different from the recovery rate of those who had sudden gains (8/16, 50%). It should be noted that of those clients classified as not having sudden gains, 8 experienced pretreatment or first-session gains.

The above results are at odds with the earlier sudden gains findings in two ways. First, our median sudden gain occurred later in treatment (after Session 10). Past studies of sudden gains in CT samples (that have reported the median session) have found median sessions ranging from Session 5 (e.g., Tang & DeRubeis, 1999b) to Session 8 (Tang et al., 2005). Second, our sudden gains did not predict improved outcomes, while those in most previous studies did. We thought that these two discrepancies might have been related, in that the inability to predict outcomes in our sample was due to the occurrence of sudden gains later in therapy. To explore this, we conducted a post-hoc analysis comparing outcomes of clients with sudden gains occurring during the first half of treatment (pre-gain sessions of 2 through 10) to those of clients with sudden gains occurring during the second half of treatment (pre-gain sessions of 11 through termination). Those who experienced sudden gains in the first half of treatment had a mean BDI of 26.56 ($SD = 7.02$) at pretreatment and 5.56 ($SD = 4.59$) at end-of-treatment, while those who experienced sudden gains in the second half had a mean BDI of 29.00 ($SD = 7.66$) at pretreatment and 14.43 ($SD = 8.14$) at end-of-treatment. Using an

---

**Table 1**

<table>
<thead>
<tr>
<th>Gain type</th>
<th>$n$</th>
<th>Pretreatment BDI</th>
<th>End-of-treatment BDI</th>
<th>Overall change</th>
<th>Gain size</th>
<th>Percent of total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretreatment gain</td>
<td>6</td>
<td>28.00 8.51</td>
<td>10.33 6.38</td>
<td>17.7 5.75</td>
<td>11.66 4.13</td>
<td>65.99</td>
</tr>
<tr>
<td>First-session gain</td>
<td>6</td>
<td>27.83 7.22</td>
<td>2.50 2.88</td>
<td>25.3 6.89</td>
<td>12.66 4.23</td>
<td>50.04</td>
</tr>
<tr>
<td>Neither</td>
<td>26</td>
<td>24.81 5.22</td>
<td>10.81 7.78</td>
<td>14.0 5.34</td>
<td>— —</td>
<td>— —</td>
</tr>
<tr>
<td>Sudden gain</td>
<td>16</td>
<td>27.63 7.16</td>
<td>9.44 7.64</td>
<td>18.19 8.73</td>
<td>9.25 2.86</td>
<td>50.85</td>
</tr>
<tr>
<td>No sudden gain</td>
<td>22</td>
<td>24.45 4.97</td>
<td>9.41 7.63</td>
<td>15.05 4.97</td>
<td>— —</td>
<td>— —</td>
</tr>
</tbody>
</table>

*Note.* Overall Change = Change in BDI from pretreatment to end-of-treatment. Percent of Total = Gain size/overall change. Neither = Clients with neither a pretreatment gain nor a first-session gain.

---

**FIGURE 1** The average sudden gain. BDI scores are shown for pretreatment, the three sessions before the sudden gain, the three sessions after the sudden gain, and end-of-treatment.
independent samples $t$ test, pretreatment scores were found to be not significantly different. However, using ANCOVA (again controlling for pretreatment scores), those who experienced sudden gains during the first half of therapy had significantly lower BDI scores at their last session, $F(1, 13) = 6.51, p = .024$. In addition, six of the nine (67%) clients with first-half sudden gains recovered, while only two of the seven (29%) clients with second-half sudden gains recovered.

Discussion

In this sample, sudden gains were similar in size and frequency to those reported in the literature. Our average gain was 9.25 BDI points and 42% of our sample experienced sudden gains, which accounted for 51% of total change in BDI. However, our median pre-gain session was Session 10, which is later than commonly reported, and our sudden gainers—as a group (early sudden gainers are discussed below)—did not evidence improved outcomes at end-of-treatment compared to non-sudden-gainers. We also found that a substantial minority of clients (12 of 38, 32%) experienced significant pretreatment and first-session gains and these gains accounted for a majority of symptom improvement (66% for pretreatment gains, 50% for first-session gains). Unlike sudden gains, most clients who experienced pretreatment or first session gains (10/12, 83%) were recovered at end-of-treatment. Those with first-session gains did particularly well, evidencing a 100% (6/6) recovery rate. These findings collectively point to the importance of very early changes in CT. Significant gains occurred before Session 2 and these gains should not be excluded from future investigations of temporal change in depression. First-session gains may be particularly important.

The debates between proponents of nonspecific versus specific factors (Ilardi & Craighead, 1994; Tang & DeRubeis, 1999a; Weinberger, 1995), cognitive versus behavioral techniques (Jacobson et al., 1996; Jacobson & Gortner, 2000), and the search for reliable cognitive mediators within CT for depression (e.g., Whisman, 1993) remain active and unresolved. In this study, the existence of pretreatment and first-session gains that predict outcome speaks to the potential importance of nonspecific factors in CT for depression. Certainly no case for specific factors can be made for those with pretreatment gains, although one can still make such a case for those with first-session gains, as cognitive therapists are encouraged to demonstrate the cognitive model in the first session. Typically, the presentation of this rationale has been interpreted as a nonspecific factor (Ilardi & Craighead, 1994); however, some cognitive theorists (e.g., Whisman, 1999) have suggested that the effect of such nonspecific factors on early depression nonetheless may be mediated by changes in cognitive variables, such as hope (Snyder et al., 2000) or “depression about depression” (Fennell & Teasdale, 1987; Teasdale, 1985). It seems that the earlier in therapy the gain occurs, the harder it becomes to make a case for cognitive techniques as the causative factors, although cognitive changes may still occur. We note that our methodology did not allow us to draw any conclusions for or against the role of specific factors across the full range of therapy sessions, only for the role of nonspecific factors early in treatment. Our data suggest that sudden gains, pretreatment gains, and first-session gains all may be important to outcome in CT.

Our findings also add to accumulating evidence on the timing and importance of sudden gains. A fairly reliable finding is that sudden gains occur early in therapy and predict outcome; in our sample, the sudden gains that occurred later in therapy appeared unrelated to outcome. Thus, although sudden gains occurred later in therapy in our sample than in previous investigations, this difference is rendered less important by these gains’ irrelevance to outcome. Similarly, Stiles and colleagues (2003) found that sudden gains that occurred before Session 16 were associated with better outcomes while those that occurred after Session 16 were not (although that study was not an analysis of CT for depression), and Kelly and colleagues (2004) found that only sudden gains occurring in the first third of treatment were positively related to outcome.

Another way of interpreting our findings is that if a “true” sudden gain occurs early in therapy and predicts outcome, then our sample contained several “false” sudden gains. This may raise an important issue about the methodology used to identify sudden gains, particularly the third criterion, which is meant to establish pre-gain and post-gain BDI stability. In fact, sudden gains may be identified as such even in the face of considerable instability in pre-gain and post-gain BDI scores. This is because a significant $t$ test between two groups of three data points is only an approximate measure of stability. For example, Figure 2 presents three of the sudden gains identified in our sample, two of which had poor outcomes. In each case, the gain looks similar to the average sudden gain...
depicted in Figure 1 if only the three sessions before and after the gain are depicted. However, a larger context of considerable BDI variability over the course of therapy appears to be important in determining the relevance of the sudden gain in each case. In the face of this variability, we caution against overinterpreting these graphs, other than as examples of considerable variability.

An additional consideration is that presenting only the average sudden gain rather than individual graphs has the additional effect of smoothing out variability in the three pre-gain and post-gain scores, possibly overemphasizing the clarity of a sudden gain in any particular case (Stiles et al., 2003). In this sense, a conservative conclusion is that some gains may be best described as random, or at least causally unknown, variability in depression mistakenly identified as sudden gains. These “false” gains would be expected to occur randomly throughout therapy, resulting in the identification of late sudden gains that do not predict outcome, and introducing error variance that limits the statistical ability to relate “true” early sudden gains and outcome. Why our sample demonstrated more of these “false” gains than other samples remains unclear, but it does not seem to be related to differences between FECT and CT or to differences between CT in this sample and in other samples.

An additional methodological consideration is that our decision to measure pretreatment and first-session gains with the same, or as similar as possible, metric as sudden gains may not have been the best choice. We chose this method largely to facilitate comparison with sudden gains in terms of size. However, this resulted in the identification of pretreatment and first-session gains that are much more common than sudden gains. Specifically, 32% of our sample had pretreatment or first-session gains in these two between-session intervals, while sudden gains occurred in 42% of our sample, but they were spread over 17 sessions. Thus, our method is useful if the goal is to identify early gains similar in size to sudden gains, but not useful if the goal is to identify early gains similar in uniqueness to sudden gains. Future investigators may consider alternate methodologies for the identification of these early gains that only identify unusually unique gains.

The current study was limited by sample sizes, especially in regard to pretreatment and first-session gains. This makes it difficult to generalize conclusions about these early gains. It should be noted that although the FECT study was a nonrandomized trial, the CT was performed by experienced and well-trained therapists, the therapy was rated as competent by an expert rater, the inclusion and exclusion criteria were identical to those of larger trials, and our response rates (66% of treatment completers) compare favorably to larger trials (Westen & Morrison, 2001). Thus, we believe our sample of depressed clients and therapists is representative of samples and therapists in clinical trials. Generalization to the larger community remains an open question (however, see Stiles et al., 2003).

References


**RECEIVED:** November 3, 2004  
**ACCEPTED:** April 22, 2005  
Available online 28 February 2006