Acceptance and Commitment Therapy: A Meta-Analytic Review

Mark B. Powers  Maarten B. Zum Vörde Sive Vörding  Paul M.G. Emmelkamp
Faculty of Social and Behavioral Sciences, Clinical Psychology, University of Amsterdam, Amsterdam, The Netherlands

Key Words
Acceptance and commitment therapy  ·  Meta-analytic review

Abstract

Background: There are now a substantial number of controlled trials investigating the efficacy of acceptance and commitment therapy (ACT). This meta-analysis combined multiple well-controlled studies to help clarify the overall impact of ACT relative to waiting lists, psychological placebos, treatment as usual, and established therapies. Method: A comprehensive literature search produced 18 randomized controlled trials (n = 917) that were included in the final analyses. Effect size was computed with Hedges’s g which can be interpreted with Cohen’s convention of small (0.2), medium (0.5), and large (0.8) effects. Results: There was a clear overall advantage of ACT compared to control conditions (effect size = 0.42). The average ACT-treated participant was more improved than 66% of the participants in the control conditions. Analyzed separately ACT was superior to waiting lists and psychological placebos (effect size = 0.68) and treatment as usual (effect size = 0.42). However, ACT was not significantly more effective than established treatments (effect size = 0.18, p = 0.13). Also, ACT was not superior to control conditions for the distress problems (anxiety/depression: effect size = 0.03, p = 0.84). Conclusions: The results reveal that ACT is more effective than control conditions for several problem domains, but there is no evidence yet that ACT is more effective than established treatments.

A growing number of studies have accumulated over the past two decades supporting the use of acceptance and commitment therapy (ACT) for several disorders [1]. The earliest version of ACT was called comprehensive distancing [2]. More recently, Hayes et al. [1] has defined ACT as ‘… a psychological intervention based on modern behavioral psychology, including relational frame theory, that applies mindfulness and acceptance processes, and commitment and behavior change processes, to the creation of psychological flexibility’. ACT includes the six core treatment processes of acceptance, defusion, contact with the present moment, self as context, values, and committed action. These treatment processes are then applied through various exercises. For example, patients may be instructed to carry a self-monitoring index card with them. The patients may then consult the card whenever they encounter a difficult situation. In addition, this exercise may be completed in session by examining a previous or hypothetical situation. On one side is the fear
algorithm that requires the patients to look at the following areas: fusion with their thoughts, evaluation of their experiences, avoidance of their experiences, and reason giving for their behavior. This helps remind the patients what barriers they face in achieving their goals and living consistent with their values. On the other side of the index card is the ACT algorithm that instructs them to: accept their reactions and be present, choose a valued direction, and take action. ACT has been modified for specific applications such as anxiety disorders [3] and depression [4]. Controlled trials have shown that ACT is beneficial for anxiety and depression [5–8], depression alone [2, 9], physical health problems [10–14], and other mental disorders [15–21].

A previous meta-analysis of ACT in 12 randomized controlled trials showed a mean controlled effect size for the primary outcome measures of $d = 0.48$ after treatment and $d = 0.63$ at follow-up [1]. Another meta-analysis of ACT in 13 randomized controlled trials reported a mean controlled effect size for primary outcome measures of Hedges’s $g = 0.68$ [22]. However, 5 additional ACT studies have been completed since that time. In addition, neither Hayes et al. [1] nor Öst [22] separately reported mean effect sizes for individual classes of disorders or for secondary outcome measures.

Therefore, the aim of the current meta-analysis was to present an up-to-date quantitative review of ACT studies across disorders and outcome measures. We derived several hypotheses from the extant literature. First, overall we expected that ACT would outperform control conditions when all outcome variables, target problems, and time points (after treatment and follow-up) were combined (hypothesis 1). Second, we did not expect that these results would be moderated by publication year (hypothesis 2).

**Method**

**Study Selection**

We selected controlled randomized trials of acceptance-based treatments for mental and physical health disorders using a comprehensive search strategy. We searched the following databases: PsycINFO 1840 to March 2008, MEDLINE 1966 to March 2008, SCOPUS 1841 to March 2008, and the Cochrane Central Register of Controlled Trials for 2008. The searches included the following terms: ‘acceptance and commitment therapy’ or ‘acceptance’, alone and in combination with ‘random’, ‘randomly’, ‘randomise’, ‘randomize’, ‘randomized’, ‘randomized’, ‘clinical trial’, or ‘trial’. These words were searched as key words, title, abstract, and MeSH subject heading terms. Also, we examined citation maps and used the ‘cited by’ search tools. These findings were cross-referenced with references from reviews and previous meta-analyses. These initial search strategies identified 329 potential articles. Studies meeting the following inclusion criteria were selected for the meta-analysis: (a) at least one ACT-based treatment; (b) random or consecutive assignment; (c) either an active or inactive control group; (d) human participants, and (e) published in English. Studies meeting the following exclusion criteria were not selected for the meta-analysis: (a) single-case studies; (b) studies of non-treatment-seeking participants, and (c) studies without data necessary to compute effect sizes (e.g. means, standard deviations). However, authors of selected studies were contacted directly when there were insufficient data provided in their articles to include in the meta-analysis.

Twelve studies were excluded due to a lack of a control group [23–34]. Seven studies were excluded because they focused on nonclinical or non-treatment-seeking samples. Follow-up comparison data were unavailable in 6 studies that were included [12, 14–16, 19, 21]. In Zettle [8], follow-up comparison data were reported for all measurements except the WRAT3.

Eighteen studies with a total sample size of 917 participants met the final inclusion criteria and were included in the meta-analysis.

**Software**

All analyses were completed with comprehensive meta-analysis [35]. Comprehensive meta-analysis is a program funded by the National Institutes of Health SBIR program.

**Procedure**

Data on the following variables were collected: target problem (e.g. anxiety and depression, depression only, physical health, other), number of participants per condition, and year of publication. Dependent variables were classified into two categories including: primary outcome measures (e.g. BDI, pain intensity) and secondary outcome measures for general subjective distress (e.g. GAF, SCL-90) and impairment/disability (e.g. SDS, ASI-Family/Social). This was completed separately for each study. For example, in a depression study the BDI would be considered a primary outcome measure, whereas in a pain study the BDI would be considered a secondary outcome measure. Treatment conditions were categorized into the following ‘types’: ACT; cognitive behavioral therapy (CBT); supportive therapy; cognitive therapy (CT); systematic desensitization; twelve-step facilitation; treatment as usual; psychological placebo, and waiting list.

**Effect Size Calculation**

Between-group effect sizes for each study were computed using Hedges’s $g$ [36]. When the necessary data were available Hedges’s $g$ was computed with a pooled standard deviation using the following formula:

$$g = \frac{X_T - X_C}{S_p}$$

where $X_T$ is the mean of the treatment group, $X_C$ is the mean of the comparison group, and $S_p$ is the pooled standard deviation. All effect sizes were corrected for small sample sizes according to Hedges and Olkin [37]. These Hedges’s $g$ controlled effect sizes may then be conservatively interpreted with Cohen’s convention.
of small (0.2), medium (0.5), and large (0.8) effects [38]. When there were multiple outcomes per domain they were combined according to Borenstein et al. [39]. The overall mean effect size for all of the studies combined was computed using the following formula:

$$g = \frac{\sum w_j g_j}{\sum w_j}$$

where $w_j$ is the weight for each study and $g_j$ is the effect size for each study.

### Results

Across studies, the most common active control condition was treatment as usual accounting for 9 out of 18 studies (table 1). In addition, a waiting list control group was included in 3 studies, and a psychological placebo control intervention was included in 1 study [13]. Finally, established interventions were included in 8 studies consisting of problem solving [5], systematic desensitization [8], CBT [7], CT [2, 6, 9], mental pain control [14], and twelve-step facilitation [20] (table 1).

### Homogeneity

A heterogeneity analysis of all studies and variables showed no significant heterogeneity ($Q = 28.55, p = 0.16$). Although this would suggest the possibility of reporting fixed effects we elected to report random effect analyses to better generalize beyond the studies included.

### Hypothesis 1: ACT Studies Compared to Control Conditions

Consistent with prediction, ACT outperformed control conditions on primary outcome measures with all target disorders and time variables (after treatment and at follow-up) pooled. This analysis showed a medium effect size favoring ACT over control conditions (Hedges’s $g = 0.42$; $SE = 0.10$, $95\% CI = 0.23–0.60$; fig. 1).

The average ACT-treated participant was more improved than 66% of the participants in the control conditions. Similar results were found when ACT was compared to control conditions on secondary outcome measures (Hedges’s $g = 0.59$; $SE = 0.14$, $95\% CI = 0.31–0.86$). Likewise, when control conditions were analyzed separately ACT outperformed waiting list and psychological placebos pooled (Hedges’s $g = 0.68$; $SE = 0.24$, $95\% CI = 0.22–1.15$), and treatment as usual (Hedges’s $g = 0.42$;
Table 1. ACT studies included in the meta-analysis

<table>
<thead>
<tr>
<th>Study</th>
<th>Target</th>
<th>n</th>
<th>Treatment type</th>
<th>Primary outcome measures</th>
<th>Coding</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bach and psychosis</td>
<td>40</td>
<td>ACT + TAU</td>
<td>Rehospitalization, Delusions</td>
<td>other mental disorders</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Boyd and anxiety and depression</td>
<td>24</td>
<td>ACT</td>
<td>Beck Depression Inventory, Job satisfaction/monotony</td>
<td>ACT + TAU &gt; TAU</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Dab and stress and chronic pain</td>
<td>11</td>
<td>ACT</td>
<td>ACT + TAU</td>
<td>ACT = CI</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Forman</td>
<td>55</td>
<td>ACT</td>
<td>Beck Depression Inventory, Clinical Global Impressions</td>
<td>ACT + TAU &gt; TAU</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Gaudiano and depression</td>
<td>15</td>
<td>ACT</td>
<td>Beck Depression Inventory, Clinical Global Impressions, Stroop</td>
<td>ACT + TAU &gt; TAU</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Gifford smoking cessation</td>
<td>21</td>
<td>ACT + TAU</td>
<td>Smoking cessation</td>
<td>ACT + TAU &gt; TAU</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Gifford smoking cessation</td>
<td>90</td>
<td>ACT + CBT + TAU</td>
<td>ACT + TAU &gt; TAU</td>
<td>ACT + CBT + TAU &gt; TAU</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Gratz and BPD</td>
<td>12</td>
<td>ACT</td>
<td>Beck Depression Inventory, Hamilton Depression Rating Scale</td>
<td>ACT + TAU &gt; TAU</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Gregg diabetes self-management</td>
<td>43</td>
<td>ACT + TAU</td>
<td>Self-reported diabetes self-care, Glycated hemoglobin</td>
<td>ACT + TAU &gt; TAU</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Hayes polysubstance abuse</td>
<td>42</td>
<td>ACT + TAU, TSF + TAU</td>
<td>Self-reported use, Urinalysis</td>
<td>ACT + TAU &gt; TAU</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Lappalainen general distress</td>
<td>14</td>
<td>ACT</td>
<td>Beck Depression Inventory, Symptom Checklist-90</td>
<td>ACT + CBT</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Leili and weight control</td>
<td>40</td>
<td>ACT</td>
<td>ACT, CRB and ACT, weight control</td>
<td>ACT + CBT</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Lundgren and drug use</td>
<td>38</td>
<td>ACT</td>
<td>ACT, CRB and ACT, weight control</td>
<td>ACT + CBT</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Vowles chronic low back pain</td>
<td>26</td>
<td>ACT + TAU</td>
<td>ACT + TAU &gt; TAU, Self-reported use, St. J. Hair Pulling Scale</td>
<td>ACT + TAU</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Wood and trichotillomania</td>
<td>12</td>
<td>ACT</td>
<td>ACT, CRB and ACT, weight control</td>
<td>ACT + CBT</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Zettle and depression</td>
<td>9</td>
<td>ACT</td>
<td>ACT, CRB and ACT, weight control</td>
<td>ACT + CBT</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Zettle and depression</td>
<td>11</td>
<td>ACT</td>
<td>ACT, CRB and ACT, weight control</td>
<td>ACT + CBT</td>
<td></td>
</tr>
</tbody>
</table>

BPD = Borderline personality disorder; CBT = cognitive behavioral therapy; CT = cognitive therapy; TAU = treatment as usual; IPP = innovation promotion program; WL = waiting list; TSF = twelve-step facilitation; ST = supportive therapy; psychPL = psychological placebo; SysD = systematic desensitization.
Meta-Analytic Review of Acceptance and Commitment Therapy

SE = 0.12, 95% CI = 0.20–0.65). However, ACT was not significantly more effective than established treatments (Hedges’s g = 0.18; SE = 0.12, 95% CI = –0.06 to 0.42; p = 0.13).

The superiority of ACT over control conditions was similar after treatment (Hedges’s g = 0.41; SE = 0.12, 95% CI = 0.18–0.64) and at follow-up (Hedges’s g = 0.41; SE = 0.11, 95% CI = 0.19–0.63). The studies varied, however, in their length of time to follow-up assessments (3–12 months).

In order to analyze each target problem separately, the data from all control conditions were pooled. Results showed that ACT was superior to control conditions for depression (a total of 2 studies: Hedges’s g = 0.76; SE = 0.33, 95% CI = 0.11–1.40), physical health (a total of 5 studies: Hedges’s g = 0.39; SE = 0.16, 95% CI = 0.08–0.70), and other mental health conditions (a total of 7 studies: Hedges’s g = 0.60; SE = 0.15, 95% CI = 0.31–0.88). However, ACT was not superior to control conditions for the distress problems (anxiety/depression, a total of 4 studies: Hedges’s g = 0.03; SE = 0.14, 95% CI = –0.25 to 0.31; p = 0.84).

Hypothesis 2: Effect Size as a Function of Publication Year

Consistent with prediction, a meta-regression analysis showed there was not a significant relationship between publication year and effect size (β = –0.02, p = 0.24).

Publication Bias: ‘The File Drawer Problem’

Several authors suggest that there may be a potential discrepancy between the number of published trials and the total number that are completed [40–43]. Therefore, any meta-analysis of published studies may be missing nonsignificant studies and therefore overestimate the overall effect size. Rosenthal [36] and others have called this confound ‘the file drawer problem’. A conservative method of addressing this problem is to assume that the effect sizes of all current or future unpublished studies are equal to 0 and to compute the number of such studies it would require to reduce the overall effect size to a nonsignificant level exceeded 5K + 10, which in this study would be 100. Analyses revealed that it would require more than 134 current or future unpublished studies with an effect size of 0 to bring the overall effect size of the primary analyses within the nonsignificant range, suggesting that the findings in this meta-analysis are robust.

Study Quality

Study quality was assessed as recommended by Jadad et al. [45]. Each study was assigned a quality score between 0 and 5 based on randomization, masking, and study flow reporting (see table 1 for scores). A meta-regression analysis showed there was not a significant relationship between study quality and effect sizes (β = –0.06, p = 0.50).

Discussion

Major Findings

The meta-analysis of 18 (n = 917) randomized controlled ACT trials largely supported the study hypotheses. Consistent with prediction, ACT outperformed control conditions on both primary and secondary outcome measures after treatment and at follow-up. Analyzed separately ACT was also superior to all control conditions except for treating distress problems (anxiety/depression). However, 3 out of the 4 studies for this analysis compared ACT to an established treatment. The one ACT study for the distress problems (anxiety/depression) that was compared to waiting list showed a small effect size favoring ACT. However, this effect size was still not significant (p = 0.28). Also, ACT was not significantly more effective than established treatments overall (e.g. CBT, CT). Effect sizes were not associated with publication year and a fail-safe N analysis showed the results may be considered robust. It is also interesting to note that we did not find a significant dose-response relationship. Short one-time ACT workshops showed similar effect sizes to longer-term therapy.

The magnitude of the effect sizes for ACT in this study was similar to previous meta-analyses [1, 22], which included substantially fewer studies. However, the meta-analysis by Hayes et al. [1] showed that ACT was superior to structured interventions after treatment (d = 0.48) and at follow-up (d = 0.63). Also, Öst [22] found that ACT was superior to established treatments with a moderate effect size. When ACT was compared to established treatments alone in this study there was no significant difference. The
discrepancy with the study by Hayes et al. [1] was associated with two factors. First, there are more studies in this meta-analysis. Second, Hayes et al. [1] included treatment as usual in addition to established treatments in his analysis. It should be noted that one of the studies with negative findings had participants with a relatively minor math anxiety problem [8]. However, there was still no difference between ACT and established treatments with this study removed. The discrepancy with the Öst [22] finding was due to several factors. First, this meta-analysis included 3 more studies for this comparison. Second, Öst [22] included the study by Gifford et al. [18] in his analysis. However, we classified the control condition in Gifford (nicotine replacement therapy) as treatment as usual to be consistent with other studies involving a standard pharmacotherapy condition. Nevertheless, even when we included this study classified as in Öst [22] ACT was still not significantly more effective than established treatments (Hedges’s $g = 0.20$, $p = 0.17$). Third, we considered the study by Bond and Bunce [5] in this analysis while Öst [22] did not. The reason for inclusion in this analysis is that we classified the innovation promotion program as an established treatment, whereas Öst [22] classified it as a psychological placebo. The innovation promotion program condition was a problem-solving intervention that included: (a) problem identification; (b) brainstorming for creative solutions; (c) development of an action plan, and (d) implementation and evaluation. We classified this condition as an established treatment as problem solving has been part of the cognitive behavioral movement since the 1960s and has been found to be effective for numerous disorders and problems [46]. Nevertheless, when we removed this study from our analysis ACT was still not significantly more effective than established treatments (Hedges’s $g = 0.24$, $p = 0.14$). The lack of difference between ACT and established treatments is not surprising. There is often no significant difference among established treatments either. For example, a meta-analysis of social phobia trials showed there was no significant difference between exposure, CT, and CBT [47]. Therefore, although ACT did not outperform established treatments, we do not see this as a relative weakness.

Limitations

Several limitations deserve comment. First, the studies could not be coded by specific disorders as most studies to date focus on diffuse problems across mental and physical health domains. However, this may be a function of the very philosophical basis of these methods. Hayes [48] defines acceptance-based approaches as oriented toward broadly defined problems. For example, the Zettle [8] study invited participants with mathematics anxiety. However, the overall effect size was not significantly different even with this study removed (Hedges’s $g = 0.45$; SE = 0.09, 95% CI = 0.27–0.63). In addition, even though there are only a couple of studies per disorder class the effect sizes are generally positive. For example, the studies of smoking cessation ($g = 0.40$) and psychosis ($g = 0.44$) were in the moderate range with effect sizes smallest for anxiety ($g = 0.39$) and highest for depression ($g = 0.74$). Nevertheless, further studies are currently under way examining acceptance-based treatment for specific disorders such as generalized anxiety disorder [29]. Second, the most common comparison condition was treatment as usual. Although this is a relative strength for external validity, without a waiting list and/or psychological placebo comparison in the same study it is difficult to assess the efficacy of the treatment as usual conditions. Effect sizes for active treatments are typically larger when compared to a waiting list than to treatment as usual [49]. However, response to treatment still varies considerably from study to study, so a direct comparison to multiple control conditions is preferred [50]. Therefore, effect sizes for ACT interventions compared to treatment as usual are difficult to interpret. For example, we do not know if these treatment as usual conditions were superior to waiting list. We would recommend inclusion of waiting list and psychological placebo conditions in future trials to avoid this concern. This concern is also true for comparisons of ACT with established treatments. Third, due to the inclusion/exclusion criteria this meta-analysis excluded 5 trials that were included in the study by Hayes et al. [1] for the following reasons: 2 of the studies did not provide the data necessary to compute the effect sizes, 2 studies did not include treatment-seeking participants (burnout among substance abuse counselors), and 1 study was a laboratory biological challenge study. Finally, this meta-analysis did not include an analysis of the putative treatment mediators. However, the proposed mediators are supported in the literature [6, 51]. For example, Forman et al. [6] found that changes in ‘observing’ and ‘describing’ mediated outcome for CT but ‘experiential avoidance’, ‘acting with awareness’, and ‘acceptance’ mediated outcomes for ACT.

Summary and Conclusion

In sum, ACT (18 studies) outperformed control conditions (waiting list, treatment as usual, psychological placebo) across target problems and outcome domains after...
treatment and at follow-up. The average ACT-treated participant was more improved than 66% of the participants in the control conditions. However, ACT did not outperform established treatments. Therefore, there was no distinct advantage of using ACT over existing established treatments. Further studies are needed in which ACT is compared with empirically supported treatments (e.g., CT, CBT, interpersonal therapy) for specific DSM-IV disorders before widespread application in routine clinical care is recommended. Overall, the effects of ACT are promising, particularly given this intervention is now only two decades old.

References