Journal of Contextual Behavioral Science 18 (2020) 181–192

Contents lists available at ScienceDirect

Journal of Contextual Behavioral Science

journal homepage: www.elsevier.com/locate/jcbs

Review Articles

The empirical status of acceptance and commitment therapy: A review of meta-analyses☆

Andrew T. Gloster [a,*], Noemi Walder [a], Michael E. Levin [b], Michael P. Twohig [b], Maria Karekla [c]

[a] University of Basel, Division of Clinical Psychology and Intervention Science, Switzerland
[b] Utah State University, U.S.A
[c] University of Cyprus, Cyprus

ARTICLE INFO

Keywords:
Acceptance and commitment therapy
Review
Meta-analysis
Evidence

ABSTRACT

The efficacy of Acceptance and Commitment Therapy (ACT) has been evaluated in many randomized controlled trials investigating a broad range of target conditions. This paper reviews the meta-analytic evidence on ACT. The 20 included meta-analyses reported 100 controlled effect sizes across n = 12,477 participants. Controlled effect sizes were grouped by target conditions and comparison group. Results showed that ACT is efficacious for all conditions examined, including anxiety, depression, substance use, pain, and transdiagnostic groups. Results also showed that ACT was generally superior to inactive controls (e.g. waitlist, placebo), treatment as usual, and most active intervention conditions (excluding CBT). Weaknesses and areas for future development are discussed.

1. Introduction

Acceptance and Commitment Therapy (ACT) aims to decrease suffering and increase well-being via six core processes of change (Hayes, Strosahl, & Wilson, 2012). In the thirty years since the first study on ACT was published (Zettle & Hayes, 1986), over 325 randomized controlled trials have been conducted (Hayes, 2019). From its seeds in North America, the proliferation of ACT trials has resulted in empirical studies from South America, Europe, Asia, Africa, and Australia. Such impressive growth is matched by positive results, with most studies reporting results that favor ACT. To date, no counterindications or iatrogenic effects have been reported to our knowledge, though they have not been extensively studied in an explicit manner. Nevertheless, some studies have reported that ACT performed less well compared to a control group in some comparisons. For example, five studies found that outcomes were not significantly different compared to either treatment as usual, cognitive behavioral therapy (CBT), befriending, or waitlist control (Craske et al., 2014; Plumb Vilardaga, 2013; Shawyer et al., 2012; Wetherell et al., 2011; White et al., 2011). Other studies showed different change trajectories between ACT and the control condition. In one study, ACT was superior to CBT at posttreatment but not significantly different at a 3-months follow-up timepoint (Avdagic, Morrissey, & Boschen, 2014) and in another, ACT was inferior at post-treatment but superior to CBT at 6-months follow-up (Lanza, García, Lamelas, & González-Menéndez, 2014). Furthermore, the quality of studies within the ACT literature varies greatly, a fact criticized in the literature (Linardon, Gleeson, Yap, Murphy, & Brennan, 2019; Öst, 2014). Thus, there is a need to systematically examine the current literature and, further, to assess the methodological quality of this evidence.

Matching the development of randomized controlled studies is the growth of reviews and meta-analyses that have examined ACT. To date, over 60 such papers have examined ACT within various topics ranging from clinical psychology to behavioral health. Many of the reviews and meta-analyses examine ACT in combination with other interventions such as dialectic behavioral therapy, mindfulness-based cognitive therapy, or behavior activation depending on the purpose of the study. This fact makes it difficult to determine the efficacy of ACT in isolation. Furthermore, many reviews and meta-analyses examine the effect of ACT in a single group of diagnoses (e.g., depression, anxiety, psychosis, etc.). Whereas this is common in the literature, the theoretical basis of ACT is transdiagnostic and thus it is important to systematically examine

☆ This research was supported by the Swiss National Science Foundation (SNF Grant # PP00P1_163716/1 & PP00P1_190082).
* Corresponding author. University of Basel Department of Psychology Division of Clinical Psychology and Intervention Science Missionsstrasse, 62 A 4055, Basel, Switzerland.
E-mail address: andrew.gloster@unibas.ch (A.T. Gloster).

https://doi.org/10.1016/j.jcbs.2020.09.009
Received 13 March 2020; Received in revised form 18 September 2020; Accepted 24 September 2020
Available online 29 September 2020

the full breadth of studies that exist in order to determine if ACT is equally efficacious across diagnoses or if ACT is less efficacious for some conditions. Furthermore, the theoretical basis of ACT suggests that outcomes of interest in intervention studies should not focus (exclusively) on symptoms or diagnoses, as has been done traditionally in the larger psychotherapy literature, but rather measure the degree to which ACT improves participants' functioning and well-being.

Summary claims of ACT's efficacy – as with any intervention – are also relative, in that the reported effect sizes are impacted by the comparison group used to determine the effect size. For example, an uncontrolled effect size (i.e., within group, pre-post comparison) will almost always be larger than controlled effect sizes (i.e., comparison to changes seen in participants in an alternate condition). Likewise, the between-group effect sizes differ as a function of the comparison group. It is therefore necessary to systematically compare ACT across various diagnostic categories and comparison groups in order to determine their impact on observed effect sizes.

With these considerations in mind, the aim of the present study was to answer the question: what are the aggregated effect sizes of ACT vs. control groups, and by target conditions, across published meta-analyses. Towards this end, we reviewed the existing meta-analytic evidence of effect sizes for ACT factoring in control group and target condition. Specifically, we only included meta-analyses reporting between-condition analyses (controlled effect sizes) and where ACT was tested in isolation (e.g., not grouped together with other therapies).

2. Method

2.1. Selection of meta-analyses

A systematic literature search was conducted by the second author on August 30th, 2019 to identify meta-analyses of ACT. In electronic databases (Ovid Medline®, PsycArticles, PsycInfo, Web of Science, and the compilation on the webpage of the Association of Contextual Behavioral Science (ACBS, 2020)) the following search terms were used: "acceptance", "commitment", "therapy", "meta", and "analysis". These searches yielded 53 results. An additional 14 meta-analyses were identified in reference lists or via hand search. Overall the literature search yielded 67 results. Inclusion criteria for this review were: 1) written in English, 2) included meta-analytic analyses of randomized controlled trials comparing ACT to active and/or inactive control conditions, and 3) published in a peer-reviewed journal.

During a first screening process of the 67 initial manuscripts, 44 were excluded because 2 meta-analyses were not written in English, 3 were comments on or author's responses to a published meta-analysis, 16 were reviews and did not report any controlled meta-analytic effect sizes, 16 did not include at least one effect size for ACT compared to another condition, 2 were not peer-reviewed, 2 meta-analyses investigated psychological flexibility processes exclusively in the lab (rather than RCTs), and 2 only described the protocol of the meta-analysis (for more details see flowchart in Fig. 1). After the extraction (see next section) three more meta-analyses were excluded because they did not report effect sizes for ACT alone compared to inactive/active conditions (instead they combined ACT with other mindfulness-based and modern cognitive behavioral treatments).

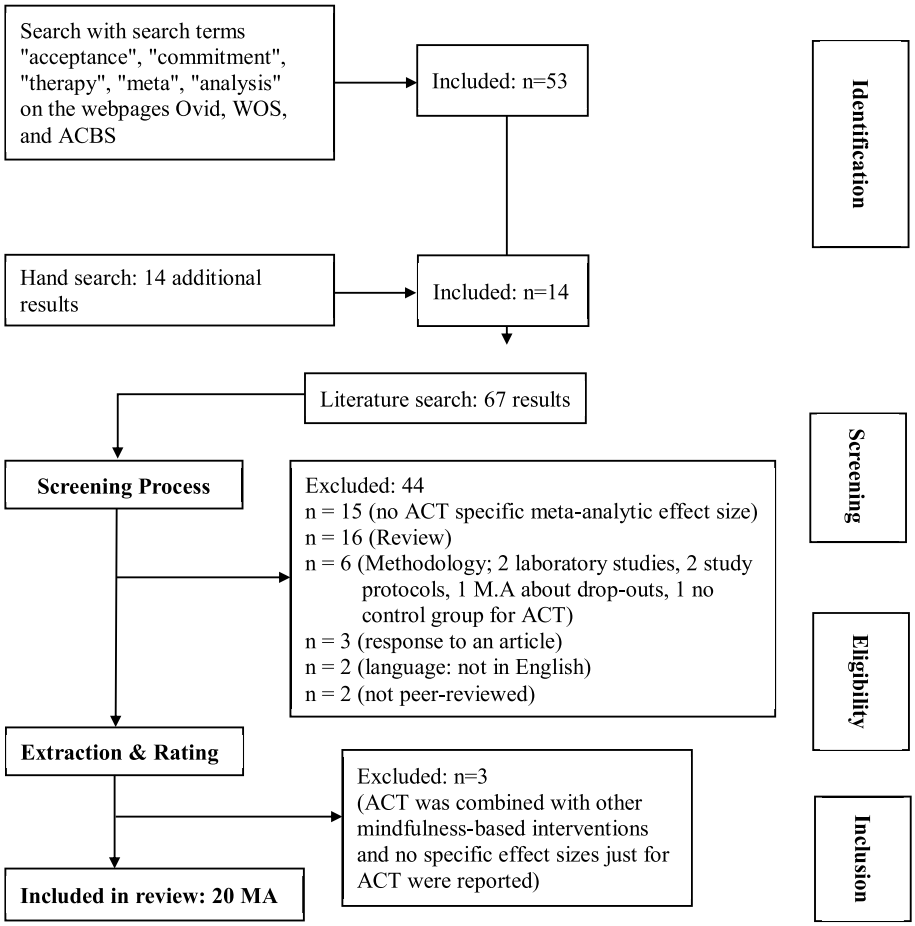

Fig. 1. Flowchart describing literature search and processing.

2.2. Extraction and rating

After the first screening process, the information of each remaining meta-analysis was extracted. We extracted effect sizes for different outcome measures, over different control conditions, and the number of comparisons these effect sizes included. In some cases, different effect sizes were given for the same comparison (e.g., ACT compared to waitlist). In these instances, the smallest effect size was chosen. For example, when outliers were omitted, the effect size without the outliers was the one extracted if it had a smaller effect size than the one with outliers.

In a first step, we grouped the effect sizes according to the investigated target condition. If a meta-analysis combined studies looking at different target conditions, we classified the reported effect size as transdiagnostic. This resulted in the following target conditions: depression (n = 15), anxiety (n = 11), substance abuse (n = 6), chronic pain (n = 8), transdiagnostic combinations of conditions (n = 24), all other conditions (n = 10), and other outcomes such as quality of life (n = 26).

In a second step, three independent raters (first, third and last author) rated the intervention or control condition ACT was compared to, for all effect sizes reported within the identified meta-analyses. Comparison groups were waitlist (WL), cognitive behavior therapy (CBT), active treatments not including CBT (active), treatment as usual (TAU), placebo, or a combination of different non-active control groups that includes WL, TAU, and placebo when they were not analyzed separately in the meta-analyses (combined control conditions). After this rating, differences were examined, discussed, and a consensus grouping was reached. All raters agreed with the final rating. The aim of the rating was to cluster the effect sizes based on their comparison group and to see how ACT performs compared to different control conditions. Within all included meta-analyses 100 comparisons were identified,

Table 1
Results of the Assessing the Methodological Quality of Systematic Reviews 2 (AMSTAR 2) quality assessment.

Meta-Analysis	AMSTAR 2 Items															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Öst (2008)	No	No	Yes	Partial yes	No	No	No	Partial Yes	Yes	No	Yes	No	Yes	Yes	Yes	No
Powers, Zum Vörde Sive Vörding, and Emmelkamp (2009)	No	No	No	Partial yes	No	No	No	Partial Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No
Veehof et al. (2011)	Yes	No	No	Partial yes	Yes	Yes	No	Partial Yes	Yes	No	Yes	No	Yes	No	Yes	Yes
(Ruiz (2012))	Yes	No	Yes	Partial yes	No	No	Yes	Partial Yes	No	No	Yes	No	Yes	Yes	Yes	No
Bluett et al. (2014)	No	No	No	Partial yes	No	No	No	Partial Yes	No	No	Yes	No	Yes	No	No	No
Öst (2014)	No	No	Yes	Partial yes	No	No	No	Partial Yes	Yes	No	Yes	No	Yes	No	Yes	No
A-Tjak et al. (2015)	No	No	No	Yes	Yes	No	No	Partial Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No
(Hacker et al. (2016))	Yes	No	No	Partial yes	Yes	Yes	No	Partial Yes	Yes	No	Yes	No	No	No	Yes	Yes
Lee, An, Levin, and Twohig (2015)	Yes	No	No	Partial yes	No	Yes	No	Partial Yes	No	No	Yes	No	Yes	No	Yes	Yes
Brown, Glendenning, Hoon, and John (2016)	Yes	No	No	Partial yes	Yes	No	No	Partial Yes	Yes	No	Yes	No	Yes	No	No	Yes
Spijkerman et al. (2016)	Yes	No	No	Partial yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No
Tonarelli et al. (2016)	Yes	Partial yes	No	Partial yes	Yes	No	No	Partial Yes	No	No	Yes	No	Yes	No	No	No
Veehof et al. (2016)	No	Partial yes	No	Partial yes	Yes	No	No	Yes	Yes	No	Yes	No	Yes	No	Yes	Yes
French et al. (2017)	Yes	No	No	Yes	Yes	No	No	Partial Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No
Hughes et al. (2017)	Yes	Yes	No	Yes	No	No	Yes	Yes	Yes	No	Yes	No	Yes	Yes	No	No
Rogers, Ferrari, Mosely, Lang, and Brennan (2017)	Yes	No	No	Partial yes	Yes	Yes	No	Yes	No	No	Yes	No	Yes	Yes	Yes	Yes
Reeve et al. (2018)	Yes	No	Yes	Partial yes	No	No	No	Partial Yes	Yes	No	Yes	No	Yes	Yes	Yes	No
Howell and Passmore, (2019)	Yes	No	No	Partial yes	No	No	No	Partial Yes	No	No	Yes	No	Yes	No	Yes	No
Ii et al. (2019)	Yes	Yes	No	Partial yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes	Yes	No	Yes
Linardon et al. (2019	Yes	No	No	Partial yes	No	No	No	Partial Yes	Yes	No	Yes	No	Yes	No	Yes	Yes

Notes: Item 1) Did the research questions and inclusion criteria for the review include the components of PICO?; Item 2) Did the meta-analysis contain an explicit statement that the methods were established prior to the conduct of the meta-analysis and did the meta-analysis justify any significant deviations from the protocol?; Item 3) Was an explanation about the selection of the study designs for inclusion included in the meta-analysis?; Item 4) Was a comprehensive literature search strategy used?; Item 5) Was the study selection performed in duplicate?; Item 6) Was the data extraction performed in duplicate?; Item 7) Was a list of excluded studies with justification for the exclusions provided?; Item 8) Were the included studies described in adequate detail?; Item 9) Did the authors of the meta-analysis use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?; Item 10) Did the authors of the meta-analysis report on the sources of funding for the studies included in the review?; Item 11) Were appropriate methods used for statistical combination of results?; Item 12) Was the potential impact of RoB in individual studies on the results of the meta-analysis assessed?; Item 13) Did the authors of the meta-analysis account for RoB in individual studies when interpreting/discussing the results of the review?; Item 14) Was a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review provided?; Item 15) Did the authors of the meta-analysis carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?; Item 16) Did the authors of the meta-analysis report any potential sources of conflict of interest, including any funding they received for conducting the review?.

which were split as follows: n = 12 comparison to CBT, n = 22 comparisons to another active intervention, n = 13 comparisons to TAU, n = 11 comparisons to WL, n = 3 comparisons to placebo, and n = 39 comparisons to combined control conditions.

For the final sample of 20 meta-analyses, the third and fourth author independently performed a quality assessment using the validated Assessing the Methodological Quality of Systematic Reviews 2 (AMSTAR-2) checklist. The AMSTAR-2 checklist includes 16 items focusing on the use of PICO as inclusion criteria, the prior registration of the review designs, how studies were selected and excluded, how the data was extracted, how the authors accounted for biases in their selected studies, the statistical analyses and the funding of the review as well as conflicts of interests (Shea et al., 2017). All items can be found in the notes of Table 1.

Outcomes were determined by means of various standardized interviews, questionnaires, behavioral or biological measurements. A detailed list of all measures across all comparisons is provided in Table 2.

All effect sizes in this review are reported in hedges g or are otherwise indicated. We first extracted the effect sizes as they were in the original meta-analyses. Effect sizes originally reported in Cohen's d were transformed into Hedge's g using the 'esc' package in R (Lüdecke, 2018) to increase comparability between effect sizes from different meta-analyses. Cohen's d and Hedge's g are very similar, however, in small sample sizes Hedge's g outperforms Cohen's d (Ellis, 2010). To simplify the interpretation of the results, U_3 scores are also provided. U scores were introduced by Cohen (1988) as a measure of nonoverlap. A U_3 score describes the percentage of the control group (e.g., CBT, Active, TAU, WL) that is exceeded by the upper half of the experimental group (ACT). Each U_3 score corresponds to a specific effect size. For example, an effect size of 0 would correspond to a U_3 score of 50% and an effect size of 1 would correspond to a U_3 score of 84%. To illustrate the meaning of a U_3 score, a U_3 score of 84% signifies that the outcome of an average ACT patient is superior to the outcome of 84% of the patients in the control group. In each result section the effect sizes as well as the range of the U_3 scores are given. Consider also Table 3, to see how a single effect size is expressed as a U_3 score.

To further illustrate the results, we report an overall mean effect size of the different effect sizes described within each target condition or comparison condition. The overall mean effect size was determined by the arithmetic mean of the individual effect sizes. These numbers should be read with caution, since they could not be weighted for number of participants as is done in primary meta-analyses.

3. Results

3.1. Sample

The final sample consisted of 20 meta-analyses, which were based on 133 studies and 12,477 participants. The individual studies that were reviewed in the meta-analyses spanned from 1986 (Zettle & Hayes, 1986) to 2018 (Grégoire, Lachance, Bouffard, & Dionne, 2018). Some studies were used in more than one meta-analysis. In order to understand the extent that individual studies were used in multiple meta-analyses ("double-dipping"), we examined all included studies in each meta-analysis and reviewed how many times each constituent study was used across all the included meta-analyses. More than half of the studies were included in only one meta-analysis. A third of the studies were used in two to four meta-analyses, a few were used five to seven times, and one study (Lundgren, Dahl, Melin, & Kies, 2006) was used ten times. The amount of unique studies in each meta-analysis varied greatly. Some of the more current meta-analyses report up to 85% unique studies (Reeve, Tickle, & Moghaddam, 2018), though one reported no unique studies (Ii et al., 2019). Newer meta-analyses have a greater chance of including unique studies (because new studies are continually being reported), while some of the older metanalyses no longer contain any unique studies as a function of their age (Tonarelli, Pasillas, Alvarado, Dwivedi, & Cancellare, 2016; Veehof, Oskam, Schreurs, & Bohlmeijer, 2011; Öst, 2008).

The methodological quality of the included meta-analyses were assessed using the AMSTAR-2 checklist (Shea et al., 2017). All or nearly all of the included meta-analyses reported on information assessed in the checklist with respect to: literature search (item 4), details of the included studies (item 8), appropriate statistical methods (item 11), and accounted for risk of bias in interpretation (item 13). None or next to none of the meta-analyses included information about: details of excluded studies (item 7) and information on the funding source (item 10). The other information assessed by the checklist was included in some to most of the included meta-analyses (range 4–15 of the meta-analyses) (see Table 1 for details).

Over all comparisons analyzed in this review, only four comparisons resulted in U_3 scores that were below 50%, meaning that for these four comparisons the outcome of an average patient in the ACT condition is superior to the outcome of less than 50% of the patients in the control condition. In 19 comparisons U_3 scores ranging from 50.0% to 59.9% were found, 44 comparisons had U_3 scores from 60.0% to 69.9%, 24 comparisons had U_3 scores from 70.0% to 79.9% and 4 comparisons indicated U_3 scores higher than 80.0%. For the remaining 5 comparisons, the effect sizes were given in risk ratios that could not be translated into U_3 scores from the information provided.

3.2. Outcomes of symptom reduction by target conditions

The findings are presented for symptom reduction measures by condition.

Depression (15 effect sizes). Nine meta-analyses were included in this review that reported on the effects of ACT for depression. Most (6 of 9) presented with significant effect sizes favoring ACT (range of ES g = 0.24- 0.76; small to medium ES) compared to active (e.g., TAU, all active psychological interventions except CBT) and inactive (e.g., waitlist, placebo) conditions. The overall mean ES was small, $g = 0.33$. Two of the meta-analytic studies favored the control condition, however both were non-significantly better than ACT for depression (Reeve et al., 2018 compared to combined control groups; Veehof, Trompetter, Bohlmeijer, & Schreurs, 2016 compared to CBT). The U_3 scores for depression ranged from 39.7% to 79.7%.

Anxiety (11 effect sizes). Seven meta-analyses were included that reported effects of ACT for anxiety spectrum disorders. Six of these presented with significant effect sizes favoring ACT with small to medium ES (g = 0.18-0.57) compared to comparison conditions. Only one meta-analysis favored active control conditions (Hacker, Stone, & Macbeth, 2016), however the effect was negligible and non-significant ($g = 0.04$, $p > .05$). The overall mean ES was small, $g = 0.24$. The U_3 scores for anxiety ranged from 48.4% to 71.6%.

Substance use (6 effect sizes). Only three meta-analyses were included that reported effects of ACT for substance use. Two of these significantly favored ACT with small effect sizes (g = 0.40-0.45) compared to other active interventions. The overall mean ES was small, $g = 0.41$. None of the studies favored the control conditions. The U_3 scores for substance use ranged from 63.3% to 67.4%.

Chronic Pain (8 effect sizes). Two meta-analyses were included in this review that evaluated ACT for chronic pain, both of which focused on studies comparing ACT to active interventions, CBT, and a combination of inactive control conditions. There was a significant and large effect favoring ACT for one meta-analysis (Hughes, Clark, Colclough, Dale, & McMillan, 2017; $g = 0.83$), whereas for the other meta-analysis the effects were non-significant (Veehof et al., 2016). The overall mean ES was small, $g = 0.44$. The U_3 scores for pain ranged from 49.2% to 82.6%.

Transdiagnostic combinations of conditions (24 effect sizes). Five meta-analyses were included that examined the effects of ACT transdiagnostically across a range of conditions compared to active and

Table 2
Meta-analyses of acceptance and commitment therapy outcome measures.

Meta-analysis	Number of effect sizes in Comparisons	Outcome cluster	Comparison group	Timepoint	Outcome Measures
Öst (2008)	8	Transdiagnostic	Active	Post	Specific measures not listed
	5	Transdiagnostic	TAU	Post	Specific measures not listed
	2	Transdiagnostic	WL	Post	Specific measures not listed
Powers et al. (2009)	9	Transdiagnostic	TAU	Post	BEST, CGI, DERS, Delusions, DSHI, Glycated hemoglobin, Hallucinations, Pain, Rehospitalization, Self-reported diabetes self-care, Smoking cessation, Stress symptoms hemoglobin
	4	Transdiagnostic	WL	Post	BDI, BMI, Hairs pulled, Job satisfaction/motivation, MGH-HS, Weight Stigma Questionnaire
	8	Transdiagnostic	Active	Post	ASI, BAI, BDI, HDRS, Job satisfaction/motivation, MARS, Pain, Self-reported use, SCL-90, TAI, Urine analysis
	2	Depression	Combined	Post	BDI, HDRS
	5	Other Conditions: Physical Health	Combined	Post	BMI, Glycated hemoglobin, Pain, Self-reported diabetes self-care, Seizure frequency, Seizure index, Stress symptoms, Weight Stigma Questionnaire,
	7	Transdiagnostic	Combined	Post	ASI, BEST, CGI, DERS, Delusions, DSHI, Hallucinations, Hairs pulled, MGH-SH, Rehospitalization, Self-reported use, Smoking cessation, Urinanalysis
	4	Transdiagnostic	Combined	Post	BAI, BDI, Job satisfaction/motivation, MARS, SCL-90, TAI
Veehof et al. (2011)	2	Pain	Combined	Post	HADS, Pain, PDI, SWLS
(Ruiz (2012))	16	Transdiagnostic	CBT	Post and FU	BAI, BDI, CSR, FACT-Breast, FNE, FQ, HRSD, Mood Visual Scale, QOLI, QOLS, SASS, SIAS, SF-36, SPS, SUD, VSLS, WILL, Y-BOCS
	10	Depression	CBT	Post and FU	BDI, HRSD, Mood Visual Scale, SASS
	9	Anxiety	CBT	Post and FU	BAI, CSR, FNE, FQ, HAS, MARS, PSWQ, SIAS, SPS, STAI, SUD, TAI, WILL, Y-BOCS,
	11	Other Outcomes: Quality of Life	CBT	Post and FU	FACT-Breast, QOLS, QOLI, SF-36, VSLS
Bluett et al. (2014)	7	Anxiety	Active	Post	BAI, DASS, GAI, HADS, SCL-90-Anxiety, STAI-S
	5	Anxiety	CBT	Post	BAI, DASS, GAI, HADS, SCL-90-Anxiety, STAI-S
Öst (2014)	16	Transdiagnostic	WL	Post	Specific measures not listed
	4	Transdiagnostic	Placebo	Post	Specific measures not listed
	14	Transdiagnostic	TAU	Post	Specific measures not listed
	30	Transdiagnostic	Active	Post	Specific measures not listed
	22	Transdiagnostic	CBT	Post	Specific measures not listed
	7	Transdiagnostic	WL FU	Follow-up	Specific measures not listed
	3	Transdiagnostic	Placebo	Follow-up	Specific measures not listed
	7	Transdiagnostic	TAU	Follow-up	Specific measures not listed
	23	Transdiagnostic	Active	Follow-up	Specific measures not listed
	17	Transdiagnostic	CBT	Follow-up	Specific measures not listed
A-Tjak et al. (2015)	9	Transdiagnostic	WL	Post and FU	AAQ, Average hairs pulled per day, BMI, Clinician severity rating, DASS, GHQ, Hours of viewing pornography, Mental health difficulties, MGH-HS, NIMH-TIS, PDI, Physical activity, PSWQ, Weekly Pain, Weight Stigma Questionnaire, Stress, THI
	5	Transdiagnostic	Placebo	Post and FU	Confidence in coping with command hallucinations, Confidence to resist command hallucinations, LDQ, OMPQ, PANSS, Seizure frequency, THI, Y-BOCS
	12	Transdiagnostic	TAU	Post and FU	ADAMS, BDI, BDI-II, Believability ratings, BEST, BPI, BPRS, BSQ, DSHI, Drug test, Drug use self-report, EDE-Q, FDI, ISS, GHQ, HbA1C, Hallucinations, Number of glucose control, MIDAS, MPQ-SF, PAIRS, Rehospitalization, Smoking Cessation Quit Rate, Self-management, SBEQ, THI, Understanding, VABS
	9	Transdiagnostic	CBT	Post and FU	ASI, BAI, BDI, BPI, Drug test, Dysphoria, FQ, HDRS, Negative affect, Negative self, PSWQ, RADS-2, Somatic, THI
	30	Other Outcomes: Secondary Outcome	CBT	Post and FU	Specific secondary outcome measures not listed
	19	Other Outcomes: Quality of Life	CBT	Post and FU	Specific secondary outcome measures not listed
	23	Other Outcomes Process Measures	Combined	Post and FU	Specific secondary outcome measures not listed
	8	Transdiagnostic	TAU	Post and FU	Specific secondary outcome measures not listed
	8	Substance Abuse	TAU	Post and FU	BDI-II, Drug Test, Drug use self-report, ISS, LDQ, Smoking Cessation Quit Rate
	15	Other Conditions: Somatic Complaints	Combined	Post and FU	BMI, BSQ, CECS, COPE, GHQ, HbA1C, Mental health difficulties, MIDAS, MPQ-SF, Number in glucose control, OMPQ, Physical activity, POMS, Self-management, Seizure frequency/duration, THI, Understanding, Weekly pain
(Hacker et al. (2016))	15	Depression	Active	Post	BDI, CES-DC, DASS-D, HADS-D,
	10	Anxiety	Active	Post	ASI, BAI, CSR, DASS-A, HADS-A, PASS, PSWQ, STAI
	28	Anxiety	WL	Post	BAI, DASS-A, HADS-A, PAI-A, PSWQ, STAI, STAI-T
	39	Depression	WL	Post	BDI, CES-D, DASS-D, GDS-10, HADS-D, PAI-D, PHQ-9, RADS-2
Lee et al. (2015)	10	Substance Abuse	Active	Follow-up	Substance abstinence

(*continued on next page*)

Table 2 (*continued*)

Meta-analysis	Number of effect sizes in Comparisons	Outcome cluster	Comparison group	Timepoint	Outcome Measures
	3	Substance Abuse	Active (CBT)	Follow-up	Substance abstinence
	5	Substance Abuse: Smoking	Active	Post	Substance abstinence
	5	Substance Abuse: Drugs	Active	Post	Substance abstinence
Brown et al. (2016)	10	Depression	Combined	Post	BDI, CES-D, DASS, HADS, MADRS-S
	7	Anxiety	Combined	Post	BAI, DASS, HADS
	8	Other Outcomes: Quality of Life	Combined	Post	GHQ-12, MHC-SF, QOLI, SCL-90
Spijkerman et al. (2016)	5	Depression	Combined	Post	BDI-II, CES-D, DASS-D, HADS-D, PHQ-9-D, POMS-D
	5	Anxiety	Combined	Post	BAI, DASS-A, HADS-A, POMS-A
	2	Other Conditions: Stress	Combined	Post	CSOSI, DASS-S, PSS, PSQ
	4	Other Outcomes: Well-Being	Combined	Post	MHC-SF, QOLI, SWLS, WHO-5,
	2	Other Outcomes: Mindfulness	Combined	Post	CAMS-R, FFMQ, FMI, MAAS
Tonarelli et al. (2016)	2	Other Conditions: Psychosis (Negative Symptoms)	TAU	Post	PANNS +
	2	Other Conditions: Psychosis (Positive Symptoms)	TAU	Post	PANNS -
	3	Other Conditions: Schizophrenia	TAU	Post	Delusions, Emotional dysfunction, Hallucinations
	3	Other Conditions: Schizo-affective	TAU	Post	Delusions, Emotional dysfunction, Hallucinations
	2	Other Outcomes: Rehospitalization	TAU	Post	Rehospitalization Rate at 4-month follow-up
Veehof et al. (2016)	2	Pain	CBT	Post	BPI-SF
	2	Depression	CBT	Post	BDI
	3	Pain	Active	Post	MPI, NRS, VAS
	3	Depression	Active	Post	HADS
	2	Other Outcomes: Disability	Active	Post	OMPQ, SF-36 PCS
	2	Other Outcomes: Quality of Life	Active	Post	QOLI
French et al. (2017)	9	Depression	Combined	Post	BDI-II, CES-D, DASS-21, CMDI, HADS
	8	Depression	WL	Post	BDI-II, CES-D, DASS-21, CMDI, HADS
	8	Anxiety	Combined	Post	BAI, DASS-21, HADS
	8	Anxiety	WL	Post	BAI, DASS-21, HADS
	10	Other Outcomes: Psychological Flexibility	Combined	Post	AAQ-II, AFQ-Y, AIS, CPAQ, PIPS, TAQ
	8	Other Outcomes: Psychological Flexibility	WL	Post	AAQ-II, CPAQ, PIPS
Hughes et al. (2017)	3	Pain: Acceptance	Combined	Post	BPCI-A, CPAQ
	6	Other Outcomes: Quality of Life	Combined	Post	LSQ, QOLI, QOLS, SF-36, SWLS
	3	Other Outcomes: Quality of Life	Combined	Follow-up	LSQ, QOLI, QOLS, SF-36, SWLS
	5	Other Outcomes: Functioning	Combined	Post	PAIRS, PDI, RMDQ
	4	Other Outcomes: Functioning	Combined	Follow-up	PAIRS, PDI
	4	Anxiety	Combined	Post	HADS, STAI-S
	3	Anxiety	Combined	Follow-up	HADS, STAI-S
	5	Depression	Combined	Post	BDI, DASS, HADS, PHQ-9
	4	Depression	Combined	Follow-up	BDI, DASS, HADS, PHQ-9
	2	Other Outcomes: Psychological Flexibility	Combined	Post	PIPS
	2	Other Outcomes: Psychological Flexibility	Combined	Follow-up	PIPS
	6	Pain: Intensity	Combined	Post	MPI, NRS, PIR, VAS
	4	Pain: Intensity	Combined	Follow-up	MPI, NRS, PIR, VAS
	2	Pain	Active	Post	BPI-S, NRS
	2	Pain	Active	Follow-up	BPI-S, NRS
	2	Other Outcomes: Quality of Life	Active	Post	SF-12 PCS, SWLS
	2	Other Outcomes: Quality of Life	Active	Follow-up	SF-12 PCS, SWLS
	2	Other Outcomes: Functioning	Active	Post	BPI-I, OMPQ, PDI
	2	Other Outcomes: Functioning	Active	Follow-up	BPI-I, OMPQ, PDI
	2	Depression	Active	Post	BDI-II, HADS
Rogers et al. (2017)	3	Other Outcomes: Quality of Life	Combined	Post	IWQOL-Lite, QOLI, WHOQOL
Reeve et al. (2018)	3	Depression	Combined	Post	MBI, SSQ
	2	Depression	Combined	Follow-up	MBI, SSQ

(*continued on next page*)

Table 2 (*continued*)

Meta-analysis	Number of effect sizes in Comparisons	Outcome cluster	Comparison group	Timepoint	Outcome Measures
	3	Other Conditions: Stress	Combined	Post	DSI, GHQ-12, GHQ-28, PANAS, WEMWBS
	2	Other Conditions: Stress	Combined	Follow-up	GHQ-12, GHQ-28, WEMWBS
	3	Other Outcomes: Psychological Flexibility	Combined	Post	AAQ-II, COPE, SSVQ, VLQ, WBSI
	2	Other Outcomes: Psychological Flexibility	Combined	Follow-up	AAQ-II, SSVQ, VLQ, WBSI
Howell and Passmore, (2019)	5	Other Outcomes: Well-Being	Combined	Post	ABS, MHC-SF, WBMMS
Ii et al. (2019)	3	Substance Abuse	TAU	Post	Substance discontinuation
Linardon et al. (2019)	3	Other Conditions: Eating Disorders	WL	Post	Binge eating frequency, BSQ, DEBQ, EAT, EDE, PEWS

Notes: AAQ = Acceptance and Action Questionnaire, ABS = Affect Balance Scales, ADAMS = Anxiety, Depression, and Mood Scale, AFQ = Avoidance and Fusion Questionnaire for Youth, AIS = Avoidance and Inflexibility Scale, ASI = Addiction Severity Index, BAI = Beck Anxiety Inventory, BDI = Becks Depression Inventory, BEST = Borderline evaluation of severity over time, BMI = Body Mass Index, BPCI = Brief Pain Coping Inventory, BPI = Brief Pain Inventory, BPRS = Brief Psychiatric Rating Scale, BSQ = Body Shape Questionnaire, CAMS-R = Cognitive and Affective Mindfulness Scale Revised, CECS = Courtland Emotional Control Scale, CES-DC = Centre for Epidemiological Studies Depression Scale for Children, CGI = Clinical Global Impression, CMDI = Chicago Multi-scale Depression Inventory, COPE = Assessment of coping, CPAQ = Chronic Pain Acceptance Questionnaire, CSOSI = Calgary Symptoms of Stress Inventory, CSR = Clinical Severity Ratings, DASS = Depression Anxiety Stress Scales, DEBQ = Dutch Eating Behavior Questionnaire, DERS = Difficulties in Emotion Regulation Scale, DSHI = Deliberate Self-harm Inventory, DSI = Daily Stress Inventory, EAT = Eating Attitudes Test, EDE-Q = Eating Disorders Examination Questionnaire, FACT = Functional Assessment of Cancer Therapy, FDI = Functional Disability Inventory, FFMQ = Five Facet Mindfulness Questionnaire, FNE = Fear of Negative Evaluation Scale, FQ= Fear Questionnaire, GAI = Geriatric Anxiety Inventory, GHQ = General Health Questionnaire, HADS = Hospital Anxiety and Depression Scale, HDRS = Hamilton Depression Rating Scale, ISS = Internalized Shame Scales, IWQOL = Impact of Weight on Quality of Life, LDQ = Leeds Dependence Questionnaire, LSQ = Life Satisfaction Questionnaire, MAAS = Mindful Attention Awareness Scale, MARS = Mathematics Anxiety Rating Scale, MBI = Maslach Burnout Inventory, MGH-HS = Massachusetts General Hospital Hairpulling Scale, MHC = Mental Health Continuum, MIDAS = Migraine Disability Assessment Scale, MPQ = McGill Pain Questionnaire, NIMH-TIS = NIMH-Trichotillomania Impairment Scale, NRS = Numerical Rating Scale, OMPQ = Örebro Muscoloskeletal Pain Questionnaire, PAI = Personality Assessment Inventory, PAIRS = Pain Impairment Relationship Scale, PANAS = Positive and Negative Affect Scale, PANSS = Positive and Negative Symptoms Scale, PASS = Pain Anxiety Symptom Scale, PCS = Physical Component Summary, PDI = Pain Disability Index, PEWS = Pediatric Early Warning Score, PHQ = Patient Health Questionnaire, PIPS = Psychological Inflexibility in Pain Scale, POMS = Profile of Mood States, PSS = Perceived Stress Scale, PSWQ = Penn State Worry Questionnaire, QOLI = Quality of Life Inventory, QOLS = Quality of Life Scale, RADS-2 = Reynolds Adolescent Depression Scale, RMDQ = Life Satisfaction Questionnaire, SASS = Social Adaptation Self-Evaluation Scale, SBEQ = Subjective Binge Eating Questionnaire, SCL-90 = Symptom Checklist-90, SF-36 = Short Form Health Survey, SIAS = Social Interaction Anxiety Scale, SPS = Social Phobia Scale, SSQ = Staff Stressor Questionnaire, SSVQ = Support Staff Values Questionnaire, STAI = State Trait Anxiety Inventory, SUD = Subject Units of Discomfort, SWLS = Satisfaction with Life Scale, TAI = Test Anxiety Inventory, TAQ = Tinnitus Acceptance Questionnaire, THI = Tinnitus Handicap Inventory, VABS = Vineland Adaptive Behaviors Scales, VAS = Visual Analogue Scale, VLQ = Valued Living Questionnaire, VSLS = Visual Scale Life Satisfaction, WBMMS = Well-Being Manifestations Measure Scale, WBSI = White Bear Suppression Inventory, WEMWBS = Warwick-Edinburgh Mental Well-Being Scale, WHO-5 = 5 Item World Health Organization Well-Being Index, WHOQOL = World Health Organization Quality of Life, WILL = Willingness Scale, Y-BOCS = Yale-Brown Obsessive-Compulsive Scale.

inactive control groups. All resulted in significant small to large effect sizes in favor of ACT (g = 0.17-0.96). The overall mean ES was small, g = 0.46. The U_3 scores investigating transdiagnostic conditions ranged from 51.2% to 83.1%.

Other conditions (10 effect sizes). Individual meta-analyses were included that reported on other conditions, such as eating disorders (n = 1), psychosis (n = 1), stress (n = 2), somatic complaints (n = 1) and physical conditions (n = 1). Five of these reported significant small to medium ES for ACT compared to control conditions (g = 0.29- 0.64). In only one meta-analysis specifically for positive psychosis symptomatology there was a non-significant negligible effect in favor of the TAU control group compared to ACT (Tonarelli et al., 2016). The U_3 scores investigating other conditions ranged from 44.0% to 73.9%.

Other Outcomes (26 effect sizes). Regarding quality of life as an outcome of the interventions tested, 6 meta-analyses were found and all reported effects in favor of ACT compared to active and inactive control groups. For three of these meta-analyses the effects were significant and medium ES (g = 0.37-0.45). For the rest (3 studies) there were non-significant differences between ACT and control conditions on quality of life. The overall mean ES was small g = 0.48. The U_3 scores range from 52.0% to 94.0%.

Three meta-analyses examined intervention effects on psychological flexibility. Two of the studies significantly favored ACT compared to active and inactive control conditions with small to large ES (g = 0.32-0.83). In the (Reeve et al. (2018)) meta-analysis, ACT did not significantly differ from other control conditions on psychological flexibility. The overall mean ES was small g = 0.42. The U_3 scores range from 52.8% to 79.7%.

Some meta-analyses (total n = 7) utilized different outcome measures (e.g., well-being, rehospitalization, physical health, mindfulness, functioning, and disability). Five of them presented with significant small to medium ES (g = 0.29-0.67) in favor of ACT. For the other two studies, ACT was not found to significantly differ on these outcomes (well-being; Spijkerman, Pots, & Bohlmeijer, 2016; and disability; Veehof et al., 2016). The overall mean ES was medium g = 0.57. The U_3 scores range from 56.7% to 99.4%.

3.3. Findings by comparison conditions

WL. Eleven effect sizes from seven meta-analyses compared ACT to WL. All 11 comparisons favored ACT and all were reported to be statistically significant. Effect sizes were calculated for outcomes of depression, anxiety, eating disorders, transdiagnostic conditions, and psychological flexibility. These outcomes were measured at post and follow up time points. The 11 meta-analytic effects comparing ACT to WL reported ESs ranging from small (g = 0.35; French, Golijani-Moghaddam, & Schröder, 2017) to large (g = 0.82; A-Tjak et al., 2015). The mean overall ES comparing ACT to WL corresponded to a medium effect (g = 0.57). The U_3 scores range from 63.7% to 83.1%.

Placebo. Three effect sizes from two meta-analyses compared ACT to placebo. Two of the three comparisons were reported to be statistically significant. All the effect sizes for comparisons between ACT and placebo were for transdiagnostic conditions. These outcomes were measured at post and follow up time points. All three meta-analytic effects reported for comparisons with placebo were medium effects sizes ranging from g = .51 (A-Tjak et al., 2015) to g = 0.59 (Öst, 2014).

Table 3
Effect sizes of included meta-analyses of Acceptance and Commitment Therapy Outcomes.

Meta-analysis	Number of Comparisons	Outcome cluster	Comparison group	Timepoint of comparison	ES	Significance	U_3 (%)
Öst (2008)	8	Transdiagnostic	Active	Post	0.53	Significant	70.2
	5	Transdiagnostic	TAU	Post	0.79	Significant	78.5
	2	Transdiagnostic	WL	Post	0.96	Significant	83.1
Powers et al. (2009)	9	Transdiagnostic	TAU	Post	0.42	Significant	66.3
	4	Transdiagnostic	WL	Post	0.68	Significant	75.2
	8	Transdiagnostic	Active	Post	0.18	Not significant	57.1
	2	Depression	Combined	Post	0.76	Significant	77.6
	5	Other Conditions: Physical Health	Combined	Post	0.39	Significant	65.2
	7	Transdiagnostic	Combined	Post	0.60	Significant	72.6
	4	Transdiagnostic	Combined	Post	0.03	Not significant	51.2
Veehof et al. (2011)	2	Pain	Combined	Post	0.28	Not significant	61.0
(Ruiz (2012))	16	Transdiagnostic	CBT	Timepoints combined (Post and FU)	0.40	Significant	65.5
	10	Depression	CBT	Timepoints combined	0.27	Not significant	60.6
	9	Anxiety	CBT	Timepoints combined	0.14	Not significant	55.6
	11	Other Outcomes: Quality of Life	CBT	Timepoints combined	0.22	Not significant	58.7
Bluett et al. (2014)	7	Anxiety	Active	Post	0.02	Not significant	50.8
	5	Anxiety	CBT	Post	0.00	Not significant	50.0
Öst (2014)	16	Transdiagnostic	WL	Post	0.63	Significant	73.6
	4	Transdiagnostic	Placebo	Post	0.59	Not significant	72.2
	14	Transdiagnostic	TAU	Post	0.55	Significant	70.9
	30	Transdiagnostic	Active	Post	0.22	Significant	58.7
	22	Transdiagnostic	CBT	Post	0.16	Not significant	56.4
	7	Transdiagnostic	WL	Follow-up	0.39	Significant	65.2
	3	Transdiagnostic	Placebo	Follow-up	0.53	Not significant	70.2
	7	Transdiagnostic	TAU	Follow-up	0.48	Significant	68.4
	23	Transdiagnostic	Active	Follow-up	0.17	Significant	56.7
	17	Transdiagnostic	CBT	Follow-up	0.06	Not significant	52.4
A-Tjak et al. (2015)	9	Transdiagnostic	WL	Timepoints combined	0.82	Significant	79.4
	5	Transdiagnostic	Placebo	Timepoints combined	0.51	Significant	69.5
	12	Transdiagnostic	TAU	Timepoints combined	0.64	Significant	73.9
	9	Transdiagnostic	CBT	Timepoints combined	0.32	Not significant	62.6
	30	Other Outcomes: Secondary Outcome	CBT	Timepoints combined	0.30	Significant	61.8
	19	Other Outcomes: Quality of Life	CBT	Timepoints combined	0.37	Significant	64.4
	23	Other Outcomes: Process Measures	Combined	Timepoints combined	0.56	Significant	71.2
	8	Transdiagnostic	TAU	Timepoints combined	0.37	Significant	64.4
	8	Substance Abuse	TAU	Timepoints combined	0.40	Significant	65.5
	15	Other Conditions: Somatic Complaints	Combined	Timepoints combined	0.58	Significant	71.9
(Hacker et al. (2016))	15	Depression	Active	Post	0.26	not significant	60.3
	10	Anxiety	Active	Post	−0.04	not significant	48.4
	28	Anxiety	WL	Post	0.45	Significant	67.4
	39	Depression	WL	Post	0.54	Significant	70.5
Lee et al. (2015)	10	Substance Abuse	Active	Follow-up	0.43	Significant	66.6
	3	Substance Abuse	CBT	Follow-up	0.34	Not significant	63.3
	5	Substance Abuse: Smoking	Active	Post	0.42	Significant	66.3
	5	Substance Abuse: Drugs	Active	Post	0.45	Significant	67.4
Brown et al. (2016)	10	Depression	Combined	Post	0.24	Significant	59.5
	7	Anxiety	Combined	Post	0.18	Significant	57.1
	8	Other Outcomes: Quality of Life	Combined	Post	0.06	Not significant	52.4
Spijkerman et al. (2016)	5	Depression	Combined	Post	0.40	Significant	65.5
	5	Anxiety	Combined	Post	0.37	Significant	64.4
	2	Other Conditions: Stress	Combined	Post	0.34	not significant	63.3
	4	Other Outcomes: Well-Being	Combined	Post	0.17	not significant	56.7

(*continued on next page*)

Table 3 (*continued*)

Meta-analysis	Number of Comparisons	Outcome cluster	Comparison group	Timepoint of comparison	ES	Significance	U_3 (%)
	2	Other Outcomes: Mindfulness	Combined	Post	0.39	Significant	65.2
Tonarelli et al. (2016)	2	Other Conditions: Psychosis (Positive Symptoms)	TAU	Post	−0.15	Not significant	44.0
	2	Other Conditions: Psychosis (Negative Symptoms)	TAU	Post	0.64	Significant	73.9
	3	Other Conditions: Schizophrenia	TAU	Post	RR = 1.03	Not significant	
	3	Other Conditions: Schizo-affective	TAU	Post	RR = 0.73	Not significant	
	2	Other Outcomes: Rehospitalization	TAU	Post	RR = 0.54	Significant	
Veehof et al. (2016)	2	Pain	CBT	Post	−0.02	Not significant	49.2
	2	Depression	CBT	Post	−0.25	Not significant	40.1 0.1
	3	Pain	Active	Post	0.94	Not significant	82.6
	3	Depression	Active	Post	0.83	Not significant	79.7
	2	Other Outcomes: Disability	Active	Post	2.52	Not significant	99.4
	2	Other Outcomes: Quality of Life	Active	Post	1.55	Not significant	93.9
French et al. (2017)	9	Depression	Combined	Post	0.28	Significant	61.0
	8	Depression	WL	Post	0.40	Significant	65.5
	8	Anxiety	Combined	Post	0.30	Significant	61.8
	8	Anxiety	WL	Post	0.35	Significant	63.7
	12	Other Outcomes: Psychological Flexibility	Combined	Post	0.32	Significant	62.6
	8	Other Outcomes: Psychological Flexibility	WL	Post	0.52	Significant	69.8
Hughes et al. (2017)	3	Pain: Acceptance	Combined	Post	0.52	Significant	69.8
	6	Other Outcomes: Quality of Life	Combined	Post	0.05	Not significant	52.0
	3	Other Outcomes: Quality of Life	Combined	Follow-up	0.26	Not significant	60.3
	5	Other Outcomes: Functioning	Combined	Post	0.45	Significant	67.4
	4	Other Outcomes: Functioning	Combined	Follow-up	0.41	Significant	65.9
	4	Anxiety	Combined	Post	0.57	Significant	71.6
	3	Anxiety	Combined	Follow-up	0.32	Not significant	62.6
	5	Depression	Combined	Post	0.52	Significant	69.8
	4	Depression	Combined	Follow-up	0.52	Significant	69.8
	2	Other Outcomes: Psychological Flexibility	Combined	Post	0.83	Significant	79.7
	2	Other Outcomes: Psychological Flexibility	Combined	Follow-up	0.64	Significant	73.9
	6	Pain: Intensity	Combined	Post	0.26	Not significant	60.3
	4	Pain: Intensity	Combined	Follow-up	0.29	Not significant	61.4
	2	Pain	Active	Post	0.83	Significant	79.7
	2	Pain	Active	Follow-up	0.42	Significant	66.3
	2	Other Outcomes: Quality of Life	Active	Post	0.39	Significant	65.2
	2	Other Outcomes: Quality of Life	Active	Follow-up	0.45	Significant	67.4
	2	Other Outcomes: Functioning	Active	Post	0.67	Significant	74.9
	2	Other Outcomes: Functioning	Active	Follow-up	0.35	Significant	63.7
	2	Depression	Active	Post	0.35	Significant	63.7
Rogers et al. (2017)	3	Other Outcomes: Quality of Life	Combined	Post	0.66	Not significant	74.5
Reeve et al. (2018)	4	Depression	Combined	Post	−0.26	Not significant	39.7
	3	Depression	Combined	Follow-up	0.05	Not significant	52.0
	4	Other Conditions: Stress	Combined	Post	0.29	Not Significant	61.4
	3	Other Conditions: Stress	Combined	Follow-up	0.09	Not significant	53.6
	3	Other Outcomes: Psychological Flexibility	Combined	Post	0.07	Not significant	52.8
	2	Other Outcomes: Psychological Flexibility	Combined	Follow-up	0.16	Not significant	56.4
Howell and Passmore, (2019)	5	Other Outcomes: Well-Being	Combined	Post	0.29	Significant	61.4
Li et al. (2019)	3	Substance Abuse	TAU	Post			

(*continued on next page*)

Table 3 (*continued*)

Meta-analysis	Number of Comparisons	Outcome cluster	Comparison group	Timepoint of comparison	ES	Significance	U_3 (%)
Linardon et al. (2019)	3	Other Conditions: Eating Disorders	WL	Post	RR = 1.34 0.5	Not significant Significant	69.1

Notes: The table presents the outcome clusters as we reported them in the result section. The category "other conditions" includes all psychological disorders other than anxiety, depression, substance abuse, and pain. The psychological diagnoses are specified behind the respective colon. The category "other outcomes" includes different secondary outcomes such as well-being, psychological flexibility, and quality of life. The effect sizes are if not other specified given in hedge's *g*. The significance of the effect size was determined by the authors of the original meta-analysis by indicating a p-value below .05, or 0.01 or 0.001, or by reporting the confidence interval. CBT (Cognitive Behavior Therapy), TAU (Treatment as Usual), WL (Waitlist), FU (Follow-up).

The mean overall effect size comparing ACT to placebo corresponded to a medium effect (g = 0.54). The U_3 scores range from 69.5% to 72.2%.

Treatment as Usual (TAU). Thirteen effect sizes from six meta-analyses compared ACT to TAU. Twelve of the 13 comparisons favored ACT and 8 of the 13 comparisons were reported to be statistically significant. The one comparison that favored TAU was a non-significant effect for negative symptomatology in psychosis. Effect sizes were calculated for outcomes of substance abuse, psychosis (positive and negative symptoms), re-hospitalization, and quality of life. These outcomes were measured at post and follow up time points. The thirteen meta-analytic effects comparing ACT to TAU reported effects sizes ranging from no effect $g = -0.15$ (Tonarelli et al., 2016) to medium $g = 0.79$ (Öst, 2014). The mean overall effect size comparing ACT to TAU corresponded to a small effect (g = 0.46). The U_3 scores range from 44.0% to 78.5%.

Active Interventions (other than CBT). Twenty-two effect sizes from eight meta-analyses compared ACT to active interventions. Twenty-one of the 22 comparisons favored ACT and 14 of the 22 comparisons were reported to be statistically significant. The one comparison that favored the active condition was a non-significant effect for anxiety (Hacker et al., 2016). Effect sizes were calculated for outcomes of anxiety, depression, chronic pain, substance abuse, transdiagnostic conditions, functioning, disability, and quality of life. These outcomes were measured at post and follow up time points. The 22 meta-analytic effects comparing ACT to active interventions other than CBT reported effects sizes ranging from no effect in anxiety $g = -0.04$ (Hacker et al., 2016) to large in disability $g = 2.52$ (Veehof et al., 2016). The mean overall effect size comparing ACT to active interventions corresponded to a medium effect ($g = 0.57$). The U_3 scores range from 48.4% to 99.4%.

Cognitive Behavioral Therapy (CBT). Twelve effect sizes from five meta-analyses compared ACT to CBT. Ten of the 12 comparisons favored ACT and 3 of the 12 comparisons were reported to be statistically significant. The two comparisons that favored CBT were non-significant effects for depression and chronic pain (Veehof et al., 2016). Effect sizes were calculated for outcomes of anxiety, depression, chronic pain, quality of life and secondary outcomes. These outcomes were measured at post and follow up time points. The 12 meta-analytic effects comparing ACT to CBT reported effects sizes ranging from no effect in anxiety $g = 0.00$ (Bluett, Homan, Morrison, Levin, & Twohig, 2014) to small in transdiagnostic outcomes $g = 0.40$ (Ruiz, 2012). The mean overall effect size comparing ACT to active interventions corresponded to a negligible effect (g = 0.16). The U_3 scores range from 40.1% to 65.5%.

Combined Control Conditions. Thirty-nine effect sizes from 10 meta-analyses compared ACT to a combination of control conditions (e.g. placebo, waitlist, TAU). Thirty-eight of the 39 comparisons favored ACT and 27 of the 39 comparisons were reported to be statistically significant. The one comparison that favored the combined control conditions was a non-significant effect for depression (Reeve et al., 2018). Effect sizes were calculated for outcomes of anxiety, depression, other mental conditions, chronic pain, somatic complaints, stress, mindfulness, psychological flexibility, quality of life, well-being, functioning, and other process measures. These outcomes were measured at post and follow up time points. The 39 meta-analytic effects comparing ACT to combined conditions reported effect sizes ranging from no effect in depression $g = -0.26$ (Reeve et al., 2018) to a large effect for psychological flexibility $g = 0.83$ (Hughes et al., 2017). The mean overall effect size comparing ACT to combined control conditions corresponded to a small effect (g = 0.33). The U_3 scores range from 39.7% to 79.7%.

4. Discussion

As evidenced across 20 meta-analyses, 133 studies, and 12,477 participants, ACT is efficacious. The evidence suggests that ACT is efficacious across a broad range of intervention targets (e.g., diagnoses of mental disorders and health conditions such as chronic pain), with largely equivalent results across these areas. As expected, effect sizes were larger when compared to inactive control groups and smaller when compared to active control groups. Importantly, in this review we exclusively extracted and reported on controlled effect sizes (i.e., between-condition comparisons in RCTs) because these are the most conservative estimates. U3 scores, a measure of nonoverlap, were reported as well to illustrate the effect sizes. The scores ranged from 40% comparing ACT to CBT to over 90% comparing ACT to another active intervention.

The literature of treatment outcome studies has traditionally been organized around specific diagnoses, and meta-analyses have followed suit. In the present review we found multiple meta-analyses showing that ACT is associated with controlled effect sizes ranging from small to medium (with mean effect sizes in the small range) for target conditions of depression, anxiety, substance abuse, and chronic pain. Multiple meta-analyses also found that ACT is efficacious transdiagnostically for a range of conditions, again with small controlled effect sizes. Single meta-analyses further found evidence for eating disorders, stress, somatic complaints, and physical conditions, with small to medium controlled effect sizes. The consistent small to medium sized controlled effects across all target conditions suggests that ACT's effects are largely uniform. The results of this review are consistent with the transdiagnostic theoretical basis of ACT. Nevertheless, in order to more fully test the transdiagnostic assumptions of ACT, future studies are needed. One type of study that is needed are meta-analyses that expand the types of disorders examined in order to continue to examine whether less common targets or populations profit as much as the targets examined in meta-analyses to date. Related, studies are needed that explicitly test multiple types of diagnoses and targets simultaneously (as opposed to in isolation and then combining at the meta-analytic level) in order to more thoroughly test the degree to which ACT can successfully be applied transdiagnostically. This later point first needs to be examined in outcome studies before being examined in meta-analyses.

Given that effects observed in all studies and meta-analyses are dependent on multiple factors and conditions, we further examined the controlled effect sizes with respect to functional outcomes and not simply symptom-based outcomes. From the onset, ACT authors have stipulated that the goal of ACT is not reduction of internal states (although that may happen) but promoting functioning and well-being. This is predicated on the fact that mental health and well-being are not

simply the opposite of symptoms (Keyes, 2005). Thus, while they are partially related, it is possible that an individual can have a high level of internal symptoms and high level of well-being just as it is possible that one can be anxiety-free and have low levels of well-being. Furthermore, ACT theory explicitly states that successful treatment promotes psychological flexibility (Hayes et al., 2012; Hayes, Luoma, Bond, Masuda, & Lillis, 2006). Based on 20 meta-analyses, this review found that controlled effect sizes for ACT are small to medium on quality of life, small to large on psychological flexibility (though one meta-analysis did not report superior meta-analytic effects for psychological flexibility), and small to medium on measures of well-being, functioning, and disability. The somewhat higher controlled effect sizes observed for these outcomes in comparison to outcomes of symptoms can be interpreted as consistent with theory. It should be noted, however, that although these are controlled effect sizes, the magnitude of the difference was significant in 50% of the comparisons. It remains an open theoretical and empirical question as to the best way to define, assess, and capture successful intervention change. We remain mindful of the fact that these effects are often based on questionnaires and thus are subject to various biases. That said, this is common across studies and meta-analyses and so it can be assumed that these effects are held largely constant.

We also examined the controlled effect sizes with respect to control conditions, with the assumption that effect sizes vary as a function of the comparison condition. When examining this contextual factor, we found small to large effect sizes for ACT compared to non-active control (e.g., waitlist), passive interventions (e.g., placebo), or a combination. With a few exceptions, ACT was either non-significantly different to or superior to other active interventions including treatment as usual, and a combination of various active interventions. ACT was generally not statistically different from CBT, although ACT was found to be more efficacious than CBT in a minority of meta-analyses (e.g., Ruiz, 2012). These results are consistent with previous studies (A-Tjak et al., 2015). Although different types of comparison groups are used for different purposes, we agree with others that testing and isolating processes of change (e.g., psychological flexibility, etc.) are more pressing research priorities than comparative trials testing two different treatments to determine if one is more efficacious. Naturally, these types of studies are not mutually exclusive, but future studies need to focus on common and unique processes, contexts, and procedures irrespective of the type of group design or even single-subject experimental designs (Gloster et al., 2017; Levin, Hildebrandt, Lillis, & Hayes, 2012; Villanueva et al., 2019; Villatte et al., 2016).

This review is subject to some important limitations. First, as aforementioned, there was a "double dipping" issue, where some studies were used for several comparisons and in more than one meta-analysis. Thus, some of the effects from individual studies may factor more than others. This was more likely to occur with older meta-analyses because newer ones have a wider range of published ACT trials from which to include studies. It remains unclear whether this "double-dipping" results in an overestimation, underestimation, or has a negligible effect on meta-analytic evidence. Future meta-analyses are encouraged to carefully consider this issue when selecting studies. A second limitation is that there were differences in terms of quality between the meta-analyses included, but we were not able to balance these differences. Some meta-analyses had inconsistencies regarding how many studies were included in a comparison. In several, it was unclear which studies were included, and in some the outcome measures were not listed. As reflected in the AMSTAR-2 assessment, several study details were missing and, in some case, incomplete. It is possible that some of these details were implemented in the studies, but not reported. Irrespective, future meta-analyses are strongly encouraged to be explicit about these methodological issues. Finally, meta-analyses are not without problems. Although meta-analyses allow summarization of effects, the observed effect sizes need to be contextualized. In this study we attempted to do this by examining the heterogeneity of effects across categories of target conditions, outcome variable, comparison group, and "double dipping." Other factors impacting the heterogeneity of effect sizes are probable and future research should try to better capture these. Finally, individual meta-analyses constituted their groups differently (e.g., what patients make up "transdiagnostic" or which treatment is used in "TAU") such that observed differences between meta-analyses within these labels may differ in part due to these contextual factors.

These limitations notwithstanding, the present review found that ACT is efficacious for a wide range of intervention targets and outcomes. Further, ACT can be considered as efficacious as traditional CBT and more efficacious than other active comparisons. Future studies are strongly recommended to examine change processes including different trajectories of change and include (additional) outcomes of functioning and well-being.

Declaration of competing interest

None.

This research was supported by the Swiss National Science Foundation (SNF Grant # PP00P1_163716/1 & PP00P1_190082). The views expressed in this paper are those of the authors and not necessarily those of the funder. The funding body in no way influenced the authors in writing the manuscript. The authors declare that they have no competing interests.

References

A-Tjak, J. G. L., Davis, M. L., Morina, N., Powers, M. B., Smits, J. A. J., et al. (2015). A meta-analysis of the efficacy of acceptance and commitment therapy for clinically relevant mental and physical health problems. *Psychotherapy and Psychosomatics, 84* (1), 30–36. https://doi.org/10.1159/000365764

Association for Contextual Behavioral Science. (2020). *ACT randomized controlled trials since 1986*. Retrieved from https://contextualscience.org/ACT_Randomized_Controlled_Trials.

Avdagic, E., Morrissey, S. A., & Boschen, M. J. (2014). A randomised controlled trial of acceptance and commitment therapy and cognitive-behaviour therapy for generalised anxiety disorder. *Behaviour Change, 31*(2), 110–130. https://doi.org/10.1017/bec.2014.5

Bluett, E. J., Homan, K. J., Morrison, K. L., Levin, M. E., & Twohig, M. P. (2014). Acceptance and commitment therapy for anxiety and OCD spectrum disorders: An empirical review. *Journal of Anxiety Disorders, 28*(6), 612–624. https://doi.org/10.1016/j.janxdis.2014.06.008

Brown, M., Glendenning, A., Hoon, A. E., & John, A. (2016). Effectiveness of web-delivered acceptance and commitment therapy in relation to mental health and well-being: A systematic review and meta-analysis. *Journal of Medical Internet Research, 18* (8), 1–14. https://doi.org/10.2196/jmir.6200

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York, NY, US: Lawrence Erlenbaum Associates.

Craske, M. G., Niles, A. N., Burklund, L. J., Wolitzky-Taylor, K. B., Vilardaga, J. C. P., Arch, J. J., et al. (2014). Randomized controlled trial of cognitive behavioral therapy and acceptance and commitment therapy for social phobia: Outcomes and moderators. *Journal of Consulting and Clinical Psychology, 82*(6), 1034–1048. https://doi.org/10.1037/a0037212

Ellis, P. D. (2010). *The essential guide to effect sizes. Statistical power, meta-analysis, and the interpretation of research results*. Cambridge: Cambridge University Press.

French, K., Golijani-Moghaddam, N., & Schröder, T. (2017). What is the evidence for the efficacy of self-help acceptance and commitment therapy? A systematic review and meta-analysis. *Journal of Contextual Behavioral Science, 6*(4), 360–374. https://doi.org/10.1016/j.jcbs.2017.08.002

Gloster, A. T., Klotsche, J., Ciarrochi, J., Eifert, G., Sonntag, R., Wittchen, H. U., et al. (2017). Increasing valued behaviors precedes reduction in suffering: Findings from a randomized controlled trial using ACT. *Behaviour Research and Therapy, 91*, 64–71. https://doi.org/10.1016/j.brat.2017.01.013

Grégoire, S., Lachance, L., Bouffard, T., & Dionne, F. (2018). The use of acceptance and commitment therapy to promote mental health and school engagement in university students: A multisite randomized controlled trial. *Behavior Therapy, 49*(3), 360–372. https://doi.org/10.1016/j.beth.2017.10.003

Hacker, T., Stone, P., & Macbeth, A. (2016). Acceptance and commitment therapy - do we know enough? Cumulative and sequential meta-analyses of randomized controlled trials. *Journal of Affective Disorders, 190*, 551–565. https://doi.org/10.1016/j.jad.2015.10.053

Hayes, S. C. (2019). State of the ACT evidence | association for contextual behavioral science. Retrieved http://contextualscience.org/state_of_the_act_evidence. (Accessed 29 January 2020).

Hayes, S. C., Luoma, J. B., Bond, F. W., Masuda, A., & Lillis, J. (2006). Acceptance and commitment therapy: Model, processes and outcomes. *Behaviour Research and Therapy, 44*(1), 1–25. https://doi.org/10.1016/j.brat.2005.06.006

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2012). Acceptance and commitment therapy. In *The process and practice of mindful change* (2nd ed.). New York, NY, US: The Guilford Press. https://doi.org/10.1017/CBO9781107415324.004.

Howell, A. J., & Passmore, H. A. (2019). Acceptance and commitment training (ACT) as a positive psychological intervention: A systematic review and initial meta-analysis regarding ACT's role in well-being promotion among university students. *Journal of Happiness Studies, 20*(6), 1995–2010. https://doi.org/10.1007/s10902-018-0027-7

Hughes, L. S., Clark, J., Colclough, J. A., Dale, E., & McMillan, D. (2017). Acceptance and commitment therapy (ACT) for chronic pain. *The Clinical Journal of Pain, 33*(6), 552–568. https://doi.org/10.1097/AJP.0000000000000425

Ii, T., Sato, H., Watanabe, N., Kondo, M., Masuda, A., Hayes, S. C., et al. (2019). Psychological flexibility-based interventions versus first-line psychosocial interventions for substance use disorders: Systematic review and meta-analyses of randomized controlled trials. *Journal of Contextual Behavioral Science, 13*(July), 109–120. https://doi.org/10.1016/j.jcbs.2019.07.003

Keyes, C. L. M. (2005). Mental illness and/or mental health? Investigating axioms of the complete state model of health. *Journal of Consulting and Clinical Psychology, 73*(3), 539–548. https://doi.org/10.1037/0022-006X.73.3.539

Lanza, P. V., García, P. F., Lamelas, F. R., & González-Menéndez, A. (2014). Acceptance and commitment therapy versus cognitive behavioral therapy in the treatment of substance use disorder with incarcerated women. *Journal of Clinical Psychology, 70*(7), 644–657. https://doi.org/10.1002/jclp.22060

Lee, E. B., An, W., Levin, M. E., & Twohig, M. P. (2015). An initial meta-analysis of Acceptance and Commitment Therapy for treating substance use disorders. *Drug and Alcohol Dependence, 155*, 1–7. https://doi.org/10.1016/j.drugalcdep.2015.08.004

Levin, M. E., Hildebrandt, M. J., Lillis, J., & Hayes, S. C. (2012). The impact of treatment components suggested by the psychological flexibility model: A meta-analysis of laboratory-based component studies. *Behavior Therapy, 43*(4), 741–756. https://doi.org/10.1016/j.beth.2012.05.003

Linardon, J., Gleeson, J., Yap, K., Murphy, K., & Brennan, L. (2019). Meta-analysis of the effects of third-wave behavioural interventions on disordered eating and body image concerns: Implications for eating disorder prevention. *Cognitive Behaviour Therapy, 48*(1), 15–38. https://doi.org/10.1080/16506073.2018.1517389

Lüdecke, D. (2018). *Effect size computation for meta analysis*. Retrieved from https://cran.r-project.org/package=esc.

Lundgren, T., Dahl, J., Melin, L., & Kies, B. (2006). Evaluation of acceptance and commitment therapy for drug refractory epilepsy: A randomized controlled trial in South Africa - a pilot study. *Epilepsia, 47*(12), 2173–2179. https://doi.org/10.1111/j.1528-1167.2006.00892.x

Öst, L. G. (2008). Efficacy of the third wave of behavioral therapies: A systematic review and meta-analysis. *Behaviour Research and Therapy, 46*(3), 296–321. https://doi.org/10.1016/j.brat.2007.12.005

Öst, L. G. (2014). The efficacy of Acceptance and Commitment Therapy: An updated systematic review and meta-analysis. *Behaviour Research and Therapy, 61*, 105–121. https://doi.org/10.1016/j.brat.2014.07.018

Plumb Vilardaga, J. C. (2013). Acceptance and commitment therapy for longstanding chronic pain in a community-based outpatient group setting. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 74* (5-B(E)), No-Specified. Retrieved from http://gateway.proquest.com/openurl?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&res_dat=xri:pqm&rft_dat=xri:pqdiss:3550275%5Cnhttp://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc10&NEWS=N&AN=2013-99220-469.

Powers, M. B., Zum Vörde Sive Vörding, M. B., & Emmelkamp, P. M. G. (2009). Acceptance and commitment therapy: A meta-analytic review. *Psychotherapy and Psychosomatics, 78*(2), 73–80. https://doi.org/10.1159/000190790

Reeve, A., Tickle, A., & Moghaddam, N. (2018). Are acceptance and commitment therapy-based interventions effective for reducing burnout in direct-care staff? A systematic review and meta-analysis. *Mental Health Review Journal, 23*(3), 131–155. https://doi.org/10.1108/MHRJ-11-2017-0052

Rogers, J. M., Ferrari, M., Mosely, K., Lang, C. P., & Brennan, L. (2017). Mindfulness-based interventions for adults who are overweight or obese: A meta-analysis of physical and psychological health outcomes. *Obesity Reviews, 18*(1), 51–67. https://doi.org/10.1111/obr.12461

Ruiz, F. J. (2012). Acceptance and commitment therapy versus traditional cognitive behavioral therapy: A systematic review and meta-analysis of current empirical evidence. *International Journal of Psychology and Psychological Therapy, 12*(3), 333–357.

Shawyer, F., Farhall, J., Mackinnon, A., Trauer, T., Sims, E., Ratcliff, K., et al. (2012). A randomised controlled trial of acceptance-based cognitive behavioural therapy for command hallucinations in psychotic disorders. *Behaviour Research and Therapy, 50*(2), 110–121. https://doi.org/10.1016/j.brat.2011.11.007

Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., Moran, J., et al. (2017). Amstar 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ, 358*, 1–9. https://doi.org/10.1136/bmj.j4008

Spijkerman, M. P. J., Pots, W. T. M., & Bohlmeijer, E. T. (2016). Effectiveness of online mindfulness-based interventions in improving mental health: A review and meta-analysis of randomised controlled trials. *Clinical Psychology Review, 45*, 102–114. https://doi.org/10.1016/j.cpr.2016.03.009

Tonarelli, S. B., Pasillas, R., Alvarado, L., Dwivedi, A., & Cancellare, A. (2016). Acceptance and commitment therapy compared to treatment as usual in psychosis: A systematic review and meta-analysis. *Journal of Psychiatry, 19*(3), 2–5. https://doi.org/10.4172/2378-5756.1000366

Veehof, M. M., Oskam, M. J., Schreurs, K. M. G., & Bohlmeijer, E. T. (2011). Acceptance-based interventions for the treatment of chronic pain: A systematic review and meta-analysis. *Pain, 152*(3), 533–542. https://doi.org/10.1016/j.pain.2010.11.002

Veehof, M. M., Trompetter, H. R., Bohlmeijer, E. T., & Schreurs, K. M. G. (2016). Acceptance- and mindfulness-based interventions for the treatment of chronic pain: A meta-analytic review. *Cognitive Behaviour Therapy, 45*(1), 5–31. https://doi.org/10.1080/16506073.2015.1098724

Villanueva, J., Meyer, A. H., Rinner, M. T. B., Firsching, V. J., Benoy, C., Brogli, S., et al. (2019). "Choose change": Design and methods of an acceptance and commitment therapy effectiveness trial for transdiagnostic treatment-resistant patients. *BMC Psychiatry, 19*(1), 173. https://doi.org/10.1186/s12888-019-2109-4

Villatte, J. L., Vilardaga, R., Villatte, M., Plumb Vilardaga, J. C., Atkins, D. C., & Hayes, S. C. (2016). Acceptance and Commitment Therapy modules: Differential impact on treatment processes and outcomes. *Behaviour Research and Therapy, 77*, 52–61. https://doi.org/10.1016/j.brat.2015.12.001

Wetherell, J. L., Afari, N., Rutledge, T., Sorrell, J. T., Stoddard, J. A., Petkus, A. J., et al. (2011). A randomized, controlled trial of acceptance and commitment therapy and cognitive-behavioral therapy for chronic pain. *Pain, 152*(9), 2098–2107. https://doi.org/10.1016/j.pain.2011.05.016

White, R., Gumley, A., McTaggart, J., Rattrie, L., McConville, D., Cleare, S., et al. (2011). A feasibility study of Acceptance and Commitment Therapy for emotional dysfunction following psychosis. *Behaviour Research and Therapy, 49*(12), 901–907. https://doi.org/10.1016/j.brat.2011.09.003

Zettle, R. D., & Hayes, S. C. (1986). Dysfunctional control by client verbal behavior: The context of reason-giving. *The Analysis of Verbal Behavior, 4*(1), 30–38. https://doi.org/10.1007/bf03392813