

<https://doi.org/10.1038/s41746-026-02466-z>

# Identifying what works in mental health apps through meta-regression analyses of 169 trials

Check for updates

Jannis Kraiss<sup>1</sup>✉, Felix Fiß<sup>1</sup>, Farid Chakhssi<sup>1,2</sup>, Fatma Betül Aktas<sup>1,3</sup>, Jurriijn Alexander Koelen<sup>1</sup> & Jorge Piano Simões<sup>1</sup>

This meta-analysis aimed to code active cognitive behavioral elements in mental health apps and to examine the association between these elements and improvements in depression and anxiety. Trials evaluating mental health apps were coded based on 34 pre-registered elements. 169 trials with 1137 timepoints were included ( $N = 41,807$ ; mean age = 34.3 years; 72.9% female). Psychoeducation, relaxation, mindfulness, and self-monitoring were used most frequently. Bivariate mixed-effect meta-regression models showed that many elements were weakly to moderately effective. Desensitization, stimulus control, and activity scheduling were most strongly and robustly associated with improvements in depression and exposure-based elements with improvements in anxiety. Ineffective elements included graded tasks and personal strengths, but in sum, there was considerable variation in the frequency and impact of active elements. Interventions incorporating a greater number of elements were more effective. This meta-analysis provides insight into how active elements in mental health apps are associated with therapeutic change, informing future interventions.

The widespread adoption of smartphones and the convenience they offer have positioned mental health applications (apps) as promising solutions to address the growing global mental health crisis, including challenges in mental health care access and a growing demand for services<sup>1</sup>. Mental health apps are highly scalable, allowing a large number of individuals to receive support<sup>2,3</sup>. Additionally, app-based interventions provide users with the discretion to seek care privately, reducing stigma, and the flexibility to engage with therapeutic tools anytime and anywhere<sup>4–6</sup>. Depressive and anxiety disorders, as well as subthreshold levels of these conditions, are among the most commonly experienced mental health issues worldwide and are associated with substantial economic costs, decreased quality of life, and disease burden<sup>7,8</sup>. Given the high prevalence and burden of these conditions and the significant gaps in mental health service availability<sup>9–11</sup>, mental health apps hold great potential for supporting people with these conditions and overcoming many barriers that limit access to care.

However, to realize this potential optimally, it is crucial to identify which features of mental health apps are most effective in addressing symptoms of depression and anxiety. Scrutinizing the evidence for mental health apps is essential, given that apps focusing on mental health are among the most widely used health apps<sup>12,13</sup>. Several meta-analyses have investigated the efficacy of app-based interventions on symptoms of depression and anxiety<sup>14–18</sup>. The most recent and comprehensive meta-analysis

synthesized findings from 176 randomized controlled trials (RCTs) and reported small to moderate effect sizes for app-based interventions in improving symptoms of depression and anxiety<sup>12</sup>, indicating that app-based cognitive behavioral therapy (CBT) interventions are effective in addressing mental health challenges.

While this research indicates that apps can be effective, existing meta-analyses often lack a detailed analysis of specific *active elements* within mental health apps. Previous meta-analyses<sup>14,15,19,20</sup>, including the one by Linardon and colleagues<sup>12</sup>, categorized interventions using low-resolution labels such as “cognitive behavioral therapy” or “acceptance-based therapies”. We argue that it is essential to move beyond these broad categorizations and rather examine which specific active elements of app-based interventions are effective. Active elements are the procedures or techniques delivered by the app in order to produce therapeutic change<sup>21</sup>. The most established set of active elements stems from the treatment model of CBT and involves elements such as relaxation, exposure, or behavioral activation. In addition, active elements may be based on so-called “third-wave” CBT interventions<sup>22</sup>, including acceptance and commitment therapy (ACT)<sup>23</sup>, mindfulness-based cognitive therapy (MBCT)<sup>24</sup>, and dialectical behavior therapy (DBT)<sup>25</sup>. Examining active elements in a more granular manner is essential because interventions with the same label may employ different active elements or combinations thereof<sup>21</sup>. For example, two “CBT”

<sup>1</sup>Department of Psychology, Health, and Technology, University of Twente, Enschede, the Netherlands. <sup>2</sup>Thubble, Dimence Groep, Deventer, the Netherlands.

<sup>3</sup>Department of Guidance and Psychological Counseling, Ibn Haldun University, Istanbul, Turkey. ✉e-mail: [j.t.kraiss@utwente.nl](mailto:j.t.kraiss@utwente.nl)

interventions might differ significantly in the active elements they entail. Furthermore, interventions not labeled as “CBT” might still incorporate CBT active elements (e.g., an ACT intervention may integrate relaxation exercises). Complicating this matter, the provision of active elements in trials is not limited to experimental conditions, but comparator groups in trials may also integrate active elements. For example, a relaxation app may be employed in the comparison group. Ideally, meta-analyses should also consider the active elements provided in the comparator groups to derive more precise estimates regarding the efficacy of active elements.

Understanding the efficacy of specific active elements requires a higher-resolution approach than the common practice of labeling interventions into broad treatment models. This granular approach can be achieved by coding specific active elements (in both the experimental and comparator groups) based on theory-driven active elements, grounded in established treatment models. Manual coding of active elements has recently been used in health psychology to identify effective behavior change techniques, particularly in the context of smoking cessation<sup>26,27</sup>. However, in the field of digital mental health interventions, this approach remains novel. Applying it in this context could provide more granular insights into what active elements are effective in app-based mental health interventions, offering a clear advantage over traditional meta-analytic methods that rely on broader classifications.

The aim of the current study is to code active cognitive behavioral elements in randomized controlled trials evaluating the efficacy of mental health apps and meta-analytically examine the association between each active element and changes in symptoms of depression and anxiety.

## Results

### Study selection

In total, 11,606 records were identified. After removal of duplicates and exclusion of titles and abstracts, 558 full texts were assessed for eligibility. Eventually, 206 papers were included, of which 25 papers were identified after additional citation searching. A reference list of all included trials and an overview of study characteristics can be found in the Supplementary Information (Tables S2 and S3). Finally, 38 individual trials and 118 groups were removed because the groups were not well-described ( $n = 51$ ) or relevant information for the calculation of effect sizes was missing ( $n = 67$ ). For the current meta-analysis, 169 trials with 340 groups and 1137 time-points were included. Figure 1 shows the PRISMA flowchart of study selection.

### Study and sample characteristics

The 169 trials included in the analysis included 41,807 participants (weighted mean age = 34.3, 72.9% female), of which 20,704 were allocated to a control group and 21,103 to an experimental group. Of the 340 groups, 173 were experimental, and 167 were comparator groups. The average weighted dropout rate at post-intervention was 28.0% in the experimental and 23.7% in the comparator groups. Most trials were conducted in the United States ( $n = 51$ ), Australia ( $n = 20$ ), China ( $n = 14$ ), and the United Kingdom ( $n = 12$ ). Most trials included nonclinical ( $n = 71$ ) or subclinical samples ( $n = 64$ ), and most trials included samples without comorbidity ( $n = 155$ ). The mental health app was an adjunctive treatment in only a few trials ( $n = 18$ ). The average time to post-intervention was 6.02 weeks. Most experimental interventions were CBT ( $n = 55$ ), multidisciplinary ( $n = 45$ ), or third-wave ( $n = 45$ ) interventions, while most control groups were waitlist ( $n = 60$ ), usual care ( $n = 23$ ), or 'no intervention' groups ( $n = 20$ ).

### Active elements

On average, groups used 3.82 (SD = 4.77) active elements. Experimental groups used 6.84 active elements (SD = 4.80, range 0–28). Two experimental groups used zero active elements, which is due to the fact that none of the 34 previously defined active elements occurred in these groups. These two groups comprised a body dissatisfaction app targeting avoidance tendencies and a simplified version of an ACT intervention that asked participants to regularly indicate whether their current behavior is directed towards or away

from their values. Trials assessing symptoms of depression as treatment outcome used 6.76 active elements in experimental conditions (SD = 4.79, range 0–28), while trials assessing symptoms of anxiety used 6.69 active elements in experimental groups (SD = 4.83, range 0–22). On average, control groups used 0.69 active elements (SD = 1.80, range 0–14). Trials assessing symptoms of depression used 0.74 active elements in control groups (SD = 1.88, range 0–14), and trials assessing symptoms of anxiety used 0.47 active elements in control groups (SD = 0.92, range 0–7). Only 9.6% of the comparator groups used more than one active element, while 80.9% of experimental groups used between 0 and 10 active elements. Figure 2 shows the distribution of active elements separately for experimental and comparator groups.

Psychoeducation ( $n = 152$ , 44.7%), relaxation ( $n = 104$ , 30.6%), and mindfulness ( $n = 100$ , 29.4%) were the most frequently used active elements. When examining trials examining symptoms of depression or anxiety separately, these three were also the most frequent elements (Table 1). In experimental groups, the most frequently used active elements were also psychoeducation ( $n = 117$ , 67.6%), relaxation ( $n = 99$ , 57.2%), and mindfulness ( $n = 96$ , 55.5%). The most frequent active elements in control groups were psychoeducation ( $n = 35$ , 21%), self-monitoring ( $n = 21$ , 12.6%), goal-setting ( $n = 6$ , 3.6%), and problem-solving ( $n = 6$ , 3.6%).

Correlations showed weak to moderate correlations between active elements, suggesting weak to moderate co-occurrence of most combinations of active elements (Fig. S1).

### Association between active elements and changes in symptoms of depression

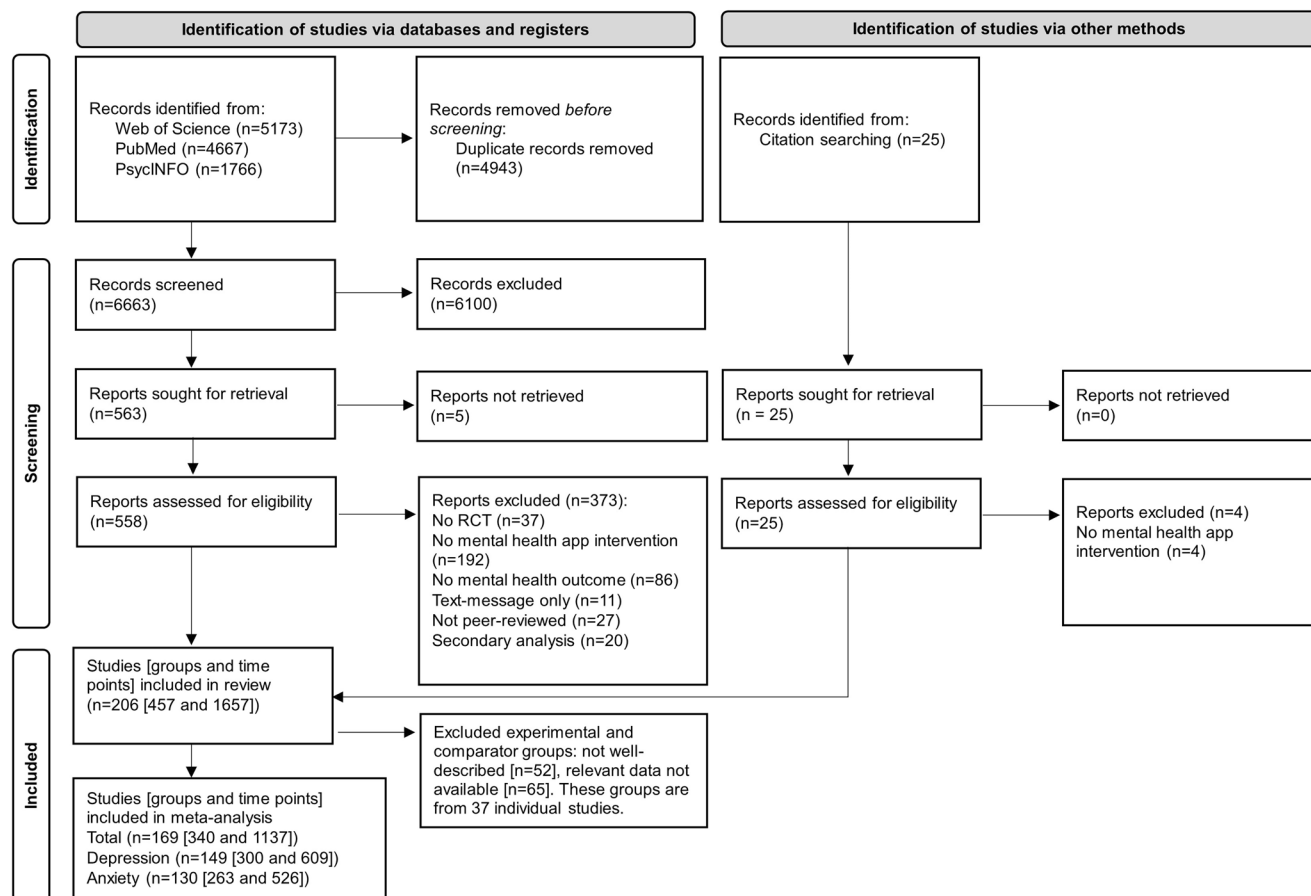
Findings from bivariate mixed-effect meta-regression models on the association between active elements and changes in symptoms of depression can be found in Table 2. Most active elements were found to be significantly associated with improvements in symptoms of depression. However, exposure in vivo and desensitization were not found to be associated with changes in symptoms of depression. Substantial differences in associations were found, ranging from elements strongly related to improvements in symptoms of depression, such as imagery-based exposure ( $B = -0.76$ ,  $p < 0.001$ ) or self-reinforcement ( $B = -0.42$ ,  $p < 0.001$ ), to elements less associated with improvements in symptoms, such as social skills ( $B = -0.22$ ,  $p < 0.01$ ) or self-compassion ( $B = -0.21$ ,  $p < 0.01$ ).  $I^2$  values ranged from 10.7% to 20.3% (level 2) and from 78.6% to 88.2% (level 3), indicating that heterogeneity in effect sizes is largely explained by between-study differences Table 2.

Most of the estimates were relatively robust to the removal of outliers. However, some less frequently used active elements became statistically insignificant after removing outliers. These included functional analysis, imagery-based exposure, interoceptive exposure, graded tasks, optimism and gratitude. Desensitization became significant after removal of outliers ( $B = -0.22$ ,  $p < 0.01$ ). Conducting the analyses with covariates included did not substantially change the results, and no significant moderation was found with the source of coding for any of the active elements.

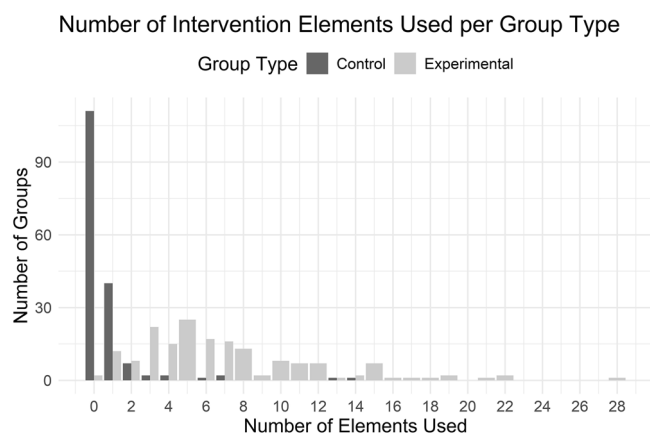
### Association between active elements and changes in symptoms of anxiety

For symptoms of anxiety (Table 3), exposure in vivo and desensitization were found to be significantly related to improvements in symptoms. However, many other elements that were significantly associated with improvements in symptoms of depression were not found to be significant, for example, journaling, self-reinforcement or social skills.  $I^2$  values ranged from 54.2% to 74.6% (level 2) and from 22.8% to 42.9% (level 3), suggesting a larger proportion of within-study heterogeneity compared to the results on depression.

Most of the estimates were relatively robust to the removal of outliers, but some of the active elements became statistically significant after removing outliers. These included functional analysis, goal-setting, journaling, self-reinforcement, social skills, mental imagery, externally-focused attention, values, committed action, gratitude, and savoring. The magnitude



**Fig. 1 | PRISMA flowchart of study selection.** This figure shows the study selection process, including records and reports excluded at each stage and reasons for exclusion of full-text reports.



**Fig. 2 | Frequency of active elements used in experimental and comparator groups.** This figure shows the frequency of active elements used in experimental and comparator groups across all 169 trials included in the meta-regression analysis.

of the effect size did not substantially change in most cases. Conducting the analyses with covariates included did not substantially change the results. We did find a significant moderation with source of coding for exposure in vivo ( $B = 0.46, p < 0.05$ ), but not for any of the other active elements.

**Moderation by clinical status**

Problem-solving, values, and savoring were more strongly associated with improvements in symptoms of depression in clinical samples (problem-solving:  $B = -0.27, p < 0.05$ ; values:  $B = -2.14, p < 0.001$ ; savoring:

$B = -2.11, p < 0.001$ ). Relaxation, behavioral activation, and psychoeducation were more strongly associated with improvements in symptoms of anxiety in clinical groups (relaxation:  $B = -0.44, p < 0.001$ ; behavioral activation:  $B = -0.45, p < 0.05$ ; psychoeducation:  $B = -0.32, p < 0.01$ ).

**Association between the number of active elements and changes in symptoms**

We found that a higher number of active elements was significantly associated with improvements in symptoms of depression ( $B = -0.03, p < 0.001$ ) and anxiety ( $B = -0.03, p < 0.001$ ), suggesting that the more active elements an app integrated, the greater the improvement in depression and anxiety. As can be seen in Fig. 3, fitting cubic splines suggested that the relationship between the number of elements and improvements in depression was largely linear. However, for symptoms of anxiety, the cubic splines showed an U-shaped relationship, suggesting that in groups that utilized 15 or more active elements, interventions were becoming less effective. It should be noted though that for symptoms of anxiety, only 14 groups (5%) utilized 15 or more active elements. Hence, this result needs to be interpreted with caution.

**Risk of bias**

The majority of the 169 trials (96%) had low risk for bias from in- and exclusion criteria, random sequence generation (79%) and blinding of outcome assessment (97%). Sixty-seven percent of trials reported an analysis approach adhering to intent-to-treat principles. However, only 39% of trials were scored with low risk for allocation concealment, 30% for blinding of participants or personnel, and 47% for risk of bias from drop-out. Only seven trials were scored with low risk for four or more items. A table with the risk of bias assessment per trial can be found in the Supplementary

**Table 1 | Frequency of active elements across groups included in the meta-analysis**

Active element	Number of groups (%)		
	Total	Depression	Anxiety
Psychoeducation	152 (44.7)	134 (44.7)	115 (43.7)
Relaxation	104 (30.6)	94 (31.3)	83 (31.6)
Mindfulness	100 (29.4)	90 (30)	78 (29.7)
Self-monitoring	99 (29.1)	88 (29.3)	72 (27.4)
Present-moment focus	94 (27.6)	84 (28)	74 (28.1)
Cognitive restructuring	69 (20.3)	62 (20.7)	46 (17.5)
Behavioral activation	57 (16.8)	54 (18)	35 (13.3)
Problem-solving	53 (15.6)	47 (15.7)	39 (14.8)
Acceptance	52 (15.3)	45 (15)	39 (14.8)
Goal-setting	48 (14.1)	44 (14.7)	35 (13.3)
Self-compassion	42 (12.4)	35 (11.7)	35 (13.3)
Activity scheduling	36 (10.6)	35 (11.7)	20 (7.6)
Social skills	36 (10.6)	33 (11)	25 (9.5)
Cognitive defusion	33 (9.7)	28 (9.3)	28 (10.6)
Gratitude	33 (9.7)	28 (9.3)	26 (9.9)
Journaling	26 (7.7)	28 (9.3)	18 (6.8)
Mental imagery	26 (7.7)	25 (8.3)	22 (8.4)
Savoring	24 (7.1)	18 (6)	21 (8)
Values	24 (7.1)	22 (7.3)	18 (6.8)
Personal strengths	21 (6.2)	20 (6.7)	13 (4.9)
Behavior experiment	18 (5.3)	14 (4.7)	14 (5.3)
Exposure in vivo	18 (5.3)	11 (3.7)	14 (5.3)
Thought record	18 (5.3)	16 (5.3)	15 (5.7)
Functional analysis	16 (4.7)	12 (4)	12 (4.6)
Optimism	16 (4.7)	15 (5)	8 (3)
Externally-focused attention	14 (4.1)	10 (3.3)	12 (4.6)
Committed action	14 (4.1)	13 (4.3)	8 (3)
Self reinforcement	14 (4.1)	13 (4.3)	8 (3)
Stimulus control	11 (3.2)	10 (3.3)	10 (3.8)
Graded tasks	10 (2.9)	9 (3)	9 (3.4)
Interoceptive exposure	8 (2.4)	4 (1.3)	7 (2.7)
Desensitization	5 (1.5)	2 (0.7)	4 (1.5)
Imagery-based exposure	4 (1.2)	3 (1)	4 (1.5)
Worry exposure	4 (1.2)	4 (1.3)	3 (1.1)

Information (Table S4) as well as a summary of risk of bias across all trials (Fig. S2).

## Discussion

This systematic review and meta-analysis coded active cognitive behavioral elements in randomized controlled trials evaluating the efficacy of mental health apps for improving symptoms of anxiety and depression. Its primary aim was to examine how each active element was associated with changes in symptoms. This meta-analysis is the first to examine the fine-grained impact of active elements in app-based mental health interventions targeting symptoms of depression and anxiety.

An interesting finding is the considerable variation in how frequently each active element was used in mental health apps. The most frequently used elements were psychoeducation, relaxation, mindfulness, and self-monitoring. In contrast, elements from ACT and positive psychology were

used less frequently. This pattern is generally consistent with a recent review that found that ACT and positive psychology interventions were less commonly included in mental health apps<sup>29</sup>. Interestingly, that review reported that only about 2% of interventions explicitly claimed to use ACT, whereas we found that ACT elements were used much more frequently, for example, cognitive defusion was deployed in about 10% of interventions. This might suggest that ACT elements are often embedded within interventions, even if the intervention is not explicitly labeled as ACT-based. Broader consumer familiarity with CBT and mindfulness-based approaches may make developers more likely to label interventions explicitly as such, whereas ACT elements may be incorporated under more broader mindfulness or CBT labels. The frequencies of active elements observed in this study should be interpreted in light of potential underreporting, which is a common issue in trials on behavioral interventions<sup>29,30</sup> and likely leads to an underestimation of the true prevalence of active elements. While coding interventions for this meta-analysis, we observed considerable variation in the quality of intervention reporting. We therefore urgently encourage authors of future trials to report active elements more consistently and thoroughly. This could be achieved, for example, by requiring authors to complete an intervention reporting checklist as a condition for publication in academic journals.

Most active elements were associated with improvements in depressive symptoms, also after excluding outliers. Many of the more frequently used active elements, such as problem-solving, cognitive restructuring, and behavioral activation, appear to have a relatively similar strength of association with improvements in depressive symptoms. This suggests that these elements are similarly effective and may therefore all be comparably effective when integrated into mental health apps. However, there might still be moderators of their efficacy, which we did not investigate in the current study. These may include other app features, such as guidance, or sample characteristics, such as age or gender. It could, for example, be that certain active elements are more effective in non-clinical or subthreshold groups, while others are more effective in clinical populations. Elements such as mood tracking may be particularly suited for individuals with milder symptoms, where increasing awareness and self-monitoring can play a preventive role or serve as early intervention. In contrast, more intensive elements like cognitive restructuring or exposure-based techniques may be necessary to produce meaningful change in clinical populations with higher symptom severity. However, the effectiveness of such intensive elements may also depend on the presence of therapeutic guidance, particularly in clinical populations. Our exploratory analyses by clinical status indeed suggest that some active elements are more effective in clinical groups. However, results from these analyses should be interpreted with caution, as the meta-analysis was not powered for this purpose, and as their lower impact in non-clinical and subclinical groups may be attributable to floor effects. Notably, the number of trials including clinical groups was relatively low in this meta-analysis, suggesting that the evidence base for these interventions in clinical populations remains limited. This may be due to ethical concerns around withholding standard treatment from clinical participants, challenges in recruiting individuals with more severe symptoms, or the additional resources required to conduct rigorous trials in clinical settings. Consequently, further research is needed to establish the effectiveness of specific active elements in these relevant groups.

It is striking that the strength of the effect sizes for many active elements is comparable to that found in the large meta-analysis by Linardon and colleagues<sup>12</sup>. They found an effect size of 0.28 across all trials using low-resolution labels. This overlap might be explained by the fact that the active elements (at least partially) co-occur in the apps. Therefore, the observed effect sizes in the current meta-analysis may partly reflect the influence of an underlying factor shared by these elements. This underlying factor might be “core” CBT, including elements such as relaxation, cognitive restructuring, and exposure, which are the most frequent active elements and are likely to be combined regularly within one intervention. A complementary hypothesis could be that engagement, including related factors such as gamification and personalization, might be a common factor in digital

**Table 2 | Results of bivariate mixed-effect meta-regression models for the association between individual active elements and improvements in symptoms of depression**

Active element	Associations ( $\beta$ , 95% CI)	
	Complete data	Outliers excluded
Functional analysis	-0.31 (-0.55, -0.06)*	-0.11 (-0.32, 0.10)
Exposure in vivo	-0.22 (-0.46, 0.02)	-0.19 (-0.39, 0.01)
Imagery-based exposure	-0.76 (-1.20, -0.32)***	-0.25 (-0.67, 0.17)
Interoceptive exposure	-0.71 (-1.09, -0.33)***	-0.34 (-0.71, 0.02)
Desensitization	-0.44 (-0.92, 0.05)	-0.47 (-0.90, -0.05)*
Problem-solving	-0.30 (-0.41, -0.19)***	-0.25 (-0.34, -0.16)***
Goal-setting	-0.27 (-0.38, -0.15)***	-0.20 (-0.30, -0.09)***
Self-monitoring	-0.24 (-0.33, -0.15)***	-0.21 (-0.29, -0.13)***
Journaling	-0.38 (-0.55, -0.20)***	-0.27 (-0.42, -0.12)***
Cognitive restructuring	-0.29 (-0.38, -0.21)***	-0.25 (-0.33, -0.18)***
Thought record	-0.26 (-0.44, -0.08)**	-0.24 (-0.39, -0.08)**
Self-reinforcement	-0.42 (-0.62, -0.22)***	-0.29 (-0.47, -0.11)**
Social skills	-0.22 (-0.35, -0.08)**	-0.19 (-0.31, -0.08)***
Activity scheduling	-0.36 (-0.49, -0.23)***	-0.31 (-0.43, -0.20)***
Behavior experiment	-0.32 (-0.54, -0.10)**	-0.16 (-0.35, 0.03)
Self-compassion	-0.21 (-0.34, -0.08)**	-0.14 (-0.25, -0.03)*
Mental imagery	-0.27 (-0.41, -0.13)***	-0.19 (-0.31, -0.06)**
Worry exposure	-0.47 (-0.84, -0.11)*	-0.20 (-0.54, 0.15)
Relaxation	-0.30 (-0.36, -0.23)***	-0.27 (-0.33, -0.22)***
Behavior activation	-0.31 (-0.40, -0.22)***	-0.27 (-0.35, -0.18)***
Psychoeducation	-0.31 (-0.38, -0.23)***	-0.28 (-0.34, -0.21)***
Graded tasks	-0.30 (-0.60, -0.05)*	-0.10 (-0.33, 0.13)
Stimulus control	-0.39 (-0.62, -0.16)***	-0.37 (-0.57, -0.18)***
Externally-focused attention	-0.35 (-0.61, -0.08)*	-0.33 (-0.54, -0.11)**
Cognitive defusion	-0.27 (-0.41, -0.14)***	-0.26 (-0.38, -0.14)***
Values	-0.27 (-0.43, -0.11)***	-0.25 (-0.39, -0.11)***
Acceptance	-0.24 (-0.35, -0.13)***	-0.23 (-0.33, -0.14)***
Present-moment focus	-0.22 (-0.31, -0.14)***	-0.23 (-0.29, -0.16)***
Committed action	-0.27 (-0.49, -0.05)*	-0.28 (-0.47, -0.08)**
Mindfulness	-0.23 (-0.30, -0.15)***	-0.23 (-0.29, -0.16)***
Gratitude	-0.23 (-0.38, -0.09)**	-0.16 (-0.29, -0.04)**
Savoring	-0.31 (-0.49, -0.14)***	-0.20 (-0.35, -0.05)**
Optimism	-0.24 (-0.43, -0.04)*	-0.14 (-0.31, 0.03)
Personal strengths	-0.20 (-0.38, -0.03)*	-0.10 (-0.25, 0.05)
Timepoints (level 1)	609	591-595 <sup>a</sup>
Groups (level 2)	300	296-297 <sup>a</sup>
Trials (level 3)	149	148

$\beta$  Standardized estimate, CI Confidence interval.

<sup>a</sup>Varying number of timepoints and groups, as the number of outliers excluded differed between active elements.

\*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$ .

interventions and might eventually explain more of the effect than specific active elements<sup>31</sup>. It is also worth noting that there was substantial overlap in included studies with the meta-analysis by Linardon and colleagues, as 136 of the trials in our meta-analysis were also included in their study. This overlap might also partially explain the similarity in some of the observed effect sizes. Nevertheless, our fine-grained analysis complements and extends their work by identifying the specific intervention ingredients most strongly associated with symptom improvement. To better disentangle the efficacy of active elements, future studies might want to deploy factorial designs or micro-randomized trials by repeatedly assigning people to

specific active elements and assessing their impact on mental health<sup>32,33</sup>. Future meta-analyses may also focus on identifying the best subset of active elements to find the most effective combination of elements.

Most of the third-wave active elements, including mindfulness and ACT elements, as well as positive psychology elements, were also significantly associated with improvements in symptoms of depression. This is generally in line with the efficacy of mindfulness-based interventions<sup>34,35</sup>, ACT<sup>36</sup>, and positive psychology<sup>37</sup> in improving symptoms of depression. One finding that should be highlighted is that the effect size for mindfulness was comparable to the overall effect of mindfulness-based interventions

**Table 3 | Results of bivariate mixed-effect meta-regression models for the association between individual active elements and improvements in symptoms of anxiety**

Active element	Associations (B, 95% CI)	
	Complete data	Outliers excluded
Functional analysis	-0.17 (-0.41, 0.07)	-0.18 (-0.36, -0.01)*
Exposure in vivo	-0.32 (-0.54, -0.10)**	-0.32 (-0.49, -0.15)***
Imagery-based exposure	-0.85 (-1.22, -0.48)***	-0.74 (-1.04, -0.43)***
Interoceptive exposure	-0.33 (-0.63, -0.02)*	-0.36 (-0.59, -0.13)**
Desensitization	-0.55 (-0.94, -0.15)**	-0.57 (-0.86, -0.28)***
Problem-solving	-0.25 (-0.38, -0.12)***	-0.28 (-0.38, -0.19)***
Goal-setting	-0.13 (-0.27, 0.01)	-0.16 (-0.27, -0.06)**
Self-monitoring	-0.18 (-0.29, -0.07)**	-0.19 (-0.27, -0.11)***
Journaling	-0.16 (-0.35, 0.04)	-0.20 (-0.34, -0.06)**
Cognitive restructuring	-0.20 (-0.33, -0.08)**	-0.20 (-0.29, -0.11)***
Thought record	-0.12 (-0.33, 0.09)**	-0.15 (-0.30, -0.00)*
Self-reinforcement	-0.18 (-0.46, 0.10)	-0.21 (-0.41, -0.01)*
Social skills	-0.14 (-0.31, 0.02)	-0.18 (-0.29, -0.06)**
Activity scheduling	-0.20 (-0.39, -0.02)*	-0.24 (-0.37, -0.11)***
Behavior experiment	-0.23 (-0.45, -0.01)*	-0.27 (-0.43, -0.10)**
Self-compassion	-0.19 (-0.33, -0.05)**	-0.23 (-0.33, -0.12)***
Mental imagery	-0.13 (-0.31, 0.04)	-0.17 (-0.30, -0.05)**
Worry exposure	-0.15 (-0.60, 0.30)	-0.11 (-0.45, 0.22)
Relaxation	-0.29 (-0.38, -0.20)***	-0.27 (-0.33, -0.21)***
Behavior activation	-0.35 (-0.49, -0.21)***	-0.25 (-0.35, -0.14)***
Psychoeducation	-0.30 (-0.39, -0.21)***	-0.25 (-0.31, -0.18)***
Graded tasks	-0.09 (-0.36, 0.19)	-0.13 (-0.32, 0.07)
Stimulus control	-0.33 (-0.58, -0.09)*	-0.30 (-0.49, -0.11)**
Externally-focused attention	-0.24 (-0.48, 0.00)	-0.28 (-0.45, -0.10)**
Cognitive defusion	-0.21 (-0.37, -0.06)*	-0.26 (-0.38, -0.15)***
Values	-0.13 (-0.32, -0.06)	-0.18 (-0.32, -0.04)*
Acceptance	-0.18 (-0.31, -0.04)**	-0.22 (-0.32, -0.13)***
Present-moment focus	-0.22 (-0.32, -0.12)***	-0.24 (-0.30, -0.17)***
Committed action	-0.19 (-0.48, -0.10)	-0.26 (-0.47, -0.04)*
Mindfulness	-0.21 (-0.31, -0.11)***	-0.23 (-0.29, -0.16)***
Gratitude	-0.08 (-0.24, 0.08)	-0.13 (-0.24, -0.01)*
Savoring	-0.11 (-0.29, 0.06)	-0.16 (-0.29, -0.03)*
Optimism	0.00 (-0.28, 0.29)	-0.04 (-0.24, 0.17)
Personal strengths	-0.05 (-0.28, 0.17)	-0.09 (-0.26, 0.07)
Timepoints (level 1)	526	508–511 <sup>a</sup>
Groups (level 2)	263	258–260 <sup>a</sup>
Trials (level 3)	130	130

β Standardized estimate, CI Confidence interval.

<sup>a</sup>Varying number of timepoints and groups, as the number of outliers excluded differed between active elements.

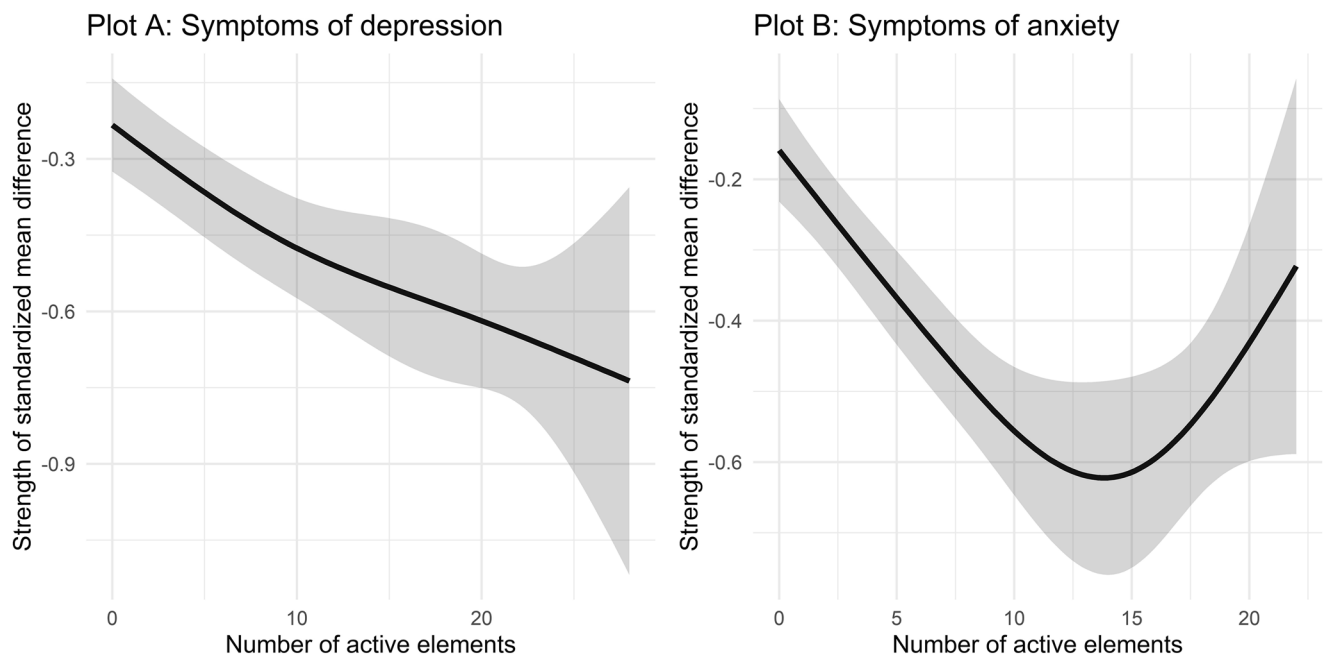
\*\*\*p < 0.001, \*\*p < 0.01, \*p < 0.05.

reported in the previous meta-analysis by Linardon and colleagues<sup>12</sup>, who found an effect size of 0.26 across all mindfulness-based apps for improving depressive symptoms. Our approach still provides a more fine-grained estimate of the effect of mindfulness, since it also includes groups delivering mindfulness content that are not explicitly labeled as mindfulness interventions in the trial. This could be the case if mindfulness is integrated into a multidisciplinary intervention involving multiple different active elements. Nevertheless, this also highlights an important shortcoming of our coding scheme that deserves attention. Eventually, we were unable to code at least some of the active elements, such as mindfulness or gratitude, in a more fine-

grained manner. For gratitude, for example, one could theoretically implement different types of gratitude exercises within one mental health app<sup>38</sup>. This limitation stems from the fact that there is a limited number of trials deploying these third-wave or positive psychology elements<sup>28</sup>, which restricts our ability to further differentiate between specific types or intensities of these active elements. This issue is further complicated by the lack of an established framework for active elements, not only for digital mental health interventions but also for mental health interventions in general. In health psychology, the taxonomy of behavior change techniques provides a reliable method for coding intervention content<sup>39,40</sup>, but this taxonomy was

**Table 4 | Overview of covariates included in bivariate mixed-effect meta-regression models**

Variable	Description	Coding
Pre-defined covariates		
Age	Mean age of participants in the trial	Continuous
Gender	Percentage of female participants in the trial	Continuous
Comorbidity	If the population has a comorbidity with a physical or mental disorder (e.g., depressed patients being treated for cancer)	Categorical (0 = no, 1 = yes)
Clinical severity	Whether the sample was nonclinical, subclinical, or clinical. Subclinical: participants were screened for elevated levels of symptoms; Clinical: participants were selected after diagnosis was made by a healthcare professional	Categorical (0 = nonclinical, 1 = subclinical, 2 = clinical)
Length to follow-up	The time from baseline to a particular timepoint of the trial in weeks	Continuous
Post-hoc covariate		
Adjunctive intervention	Whether the smartphone-based mental health app was an adjunctive treatment	Categorical (0 = no, 1 = yes)



**Fig. 3 | Cubic splines for the association between total number of active elements and improvements in symptoms of depression (Plot A) and anxiety (Plot B).** This plot shows improvement in symptoms of depression or anxiety as a non-linear

function of number of elements used in mental health apps. Improvement in symptoms is expressed as standardized within-group change from baseline to follow-up assessments.

primarily developed in the context of health-related interventions, such as treatments targeting smoking cessation. In the context of mental health interventions, a comparable comprehensive taxonomy is currently lacking. To better understand what works in mental health interventions, future research could focus on the development of such a taxonomy tailored to mental health interventions. Furthermore, future meta-analyses or experimental studies could benefit from incorporating outcomes related to positive functioning and well-being<sup>41</sup>, rather than focusing solely on symptom reduction. It is possible that certain active elements primarily influence symptom-related outcomes, while others may have a greater impact on positive functioning and well-being.

Several exposure-based elements, such as imagery-based exposure and interoceptive exposure (but not exposure *in vivo*), were also associated with improvements in depressive symptoms. In some cases, such as for interoceptive exposure, the association is even considerably stronger than for other elements. Additionally, desensitization and stimulus control emerged as the most significant elements after excluding outliers, suggesting that they may be particularly effective in alleviating depressive symptoms. Traditionally, such exposure-based elements are effective in treating anxiety

symptoms<sup>42</sup>, making this finding somewhat surprising. However, it is also important to note that imagery-based exposure, interoceptive exposure, and desensitization were only investigated in a small number of trials. This raises the possibility that their apparent efficacy may be driven by a few highly effective studies, rather than by the specific impact of the individual active elements themselves. We recommend interpreting the results of these less prevalent elements with caution, particularly when they no longer reach significance after the exclusion of outliers. Future research should prioritize (micro-)randomized trials that specifically isolate these promising but understudied techniques to confirm their efficacy.

It is interesting to note that most of the exposure-based elements were significantly associated with improvements in symptoms of anxiety, even after excluding outliers. These included exposure *in vivo*, imagery-based exposure, interoceptive exposure, and desensitization. This is relevant for two reasons. First, it suggests that exposure exercises delivered via apps are effective and could therefore be integrated into digital interventions. This has the potential to significantly reduce the burden on therapists if individuals can engage in these exercises independently<sup>43</sup>. The combination of exposure exercises with virtual reality could be especially promising in this

context<sup>44</sup>. Second, it indicates that there may be a symptom-specific effect of exposure elements on anxiety, as we did not observe similarly robust effects for depressive symptoms after excluding outliers. This finding is important, as it supports the potential for greater personalization in mental health apps, depending on which symptoms the intervention aims to target.

We also found that many other active elements were associated with improvements in symptoms of anxiety, such as problem-solving, self-monitoring, and cognitive restructuring. Overall, these findings suggest that many elements that are effective in reducing depressive symptoms are also effective for anxiety. This is also in line with the meta-analysis by Linardon and colleagues, who found similar effects for anxiety as for symptoms of depression<sup>12</sup>, and it also aligns with the ongoing development of transdiagnostic treatment approaches, such as the Unified Protocol, which proposes that similar compositions of treatment elements may be equally effective across a range of mental disorders<sup>45</sup>. Clinically, these similar effects for depression and anxiety may also be explained by the strong symptomatic overlap between the two conditions, as well as their high degree of comorbidity<sup>46</sup>. While many of the active elements were significantly associated with improvements in symptoms of depression and anxiety, it is important to note that there are substantial differences in the magnitude of these associations. This contributes to our understanding by identifying which elements are more effective, thereby informing future research and guiding the development of apps to maximize effect sizes. Notably, as for symptoms of anxiety, more active elements became statistically significant after removing outliers, while the magnitude of the associations remained relatively unchanged. This is likely due to reduced heterogeneity in the models, which decreases the uncertainty of the estimates.

Another relevant finding is that app-based interventions incorporating a greater number of active elements tended to be more effective. This suggests that multicomponent apps may be particularly beneficial, likely because they offer a more comprehensive intervention package that addresses a broader range of problems. Additionally, including a wider variety of elements provides people with more options, which could enhance adherence and engagement, which remain challenging to enhance in contemporary digital interventions<sup>47–49</sup>. For symptoms of depression, this relationship appeared to be linear, whereas for anxiety, a U-shaped relationship was observed. This is interesting, as it might suggest a “sweet spot” in the number of active elements. Due to the relatively low number of trials that included a high number of active elements, the evidence in this meta-analysis remains limited and should be interpreted with caution. Nevertheless, it could be valuable for future research to compare different interventions with a varying number of active elements to examine whether such a sweet spot can be identified empirically.

Several limitations need to be considered when interpreting the findings. First, the coding of active elements was based on a framework that focused on a selected subset of elements. Some therapeutic models, such as DBT or schema therapy, are likely to be rare in app-based mental health interventions, since they often require intensive, individualized, and therapist-guided approaches. Therefore, the current framework does not represent an exhaustive list of active therapeutic elements. In this context, some elements could theoretically also be coded in a more fine-grained manner. For example, cognitive restructuring includes many different techniques. However, distinguishing different techniques remains challenging due to the limited number of trials and a lack of detailed reporting. Second, the coding of elements was partially conducted manually by a data extractor and was based on published intervention materials. This likely led to an underestimation of active elements, as underreporting is a common issue in behavioral intervention studies<sup>29,30</sup>. Third, the dosage of each active element remains unknown. It is reasonable to assume that the same element could be delivered with varying intensity across different apps, which is an important factor to consider in future analyses.

In sum, this meta-analysis is the first to examine the fine-grained impact of specific active elements on symptoms of depression and anxiety in

app-based mental health interventions. We found substantial variation in the frequency with which different active elements were used, with the most common being psychoeducation, relaxation, mindfulness, and self-monitoring. Many active elements were significantly associated with reductions in depressive and anxiety symptoms. However, some elements appeared to primarily affect either depression or anxiety, and there were notable differences in the magnitude of these associations. In addition, we found that interventions incorporating a greater number of active elements tend to be more effective. Notably, we observed considerable variation in the quality of intervention reporting. We therefore urgently encourage authors of future trials to report active elements more consistently and thoroughly. Overall, this meta-analysis provides a foundation for optimizing digital mental health interventions by identifying which active elements are most effective. By pinpointing the components that may drive change, these findings offer concrete guidance for designing more targeted and impactful mental health apps.

## Methods

A systematic literature review and meta-analysis were conducted, which were pre-registered at PROSPERO (ID: CRD42025630092, registered January 17, 2025). The study protocol and extensive supporting information are available on OSF (<https://osf.io/wc59a/>). This article follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)<sup>50</sup>. A completed PRISMA reporting checklist can be found in the Supplementary Information (Table S5).

## Eligibility criteria

Papers were included if (1) the study design was a randomized controlled trial, (2) the intervention was a smartphone-based mental health app, and (3) the trial included a validated instrument assessing symptoms of depression or anxiety. We defined smartphone-based mental health apps as mobile applications offered on smartphones designed to improve psychological distress or well-being, offering therapeutic techniques based on evidence-based treatment models, and potentially including features like self-management tools, data tracking, or educational resources. Smartphone-based mental health interventions were allowed to involve human guidance, could be combined with other technologies such as wearables, and may be offered in addition to face-to-face therapy. Trials that held an information session, an intake meeting, or a single session of psychoeducation before delivering the app program were included. Secondary analyses were included only if they provided additional findings not reported in the original paper. Moreover, smartphone-based interventions administered as adjunctive treatments (i.e., in combination with treatment as usual) were included, as long as the adjunctive smartphone-based intervention was primarily targeting mental health outcomes. Trials were excluded that evaluated interventions focusing on lifestyle behaviors (e.g., weight loss, diet), text-message interventions, or cognitive training (e.g., cognitive bias modification). Papers that were not peer-reviewed were excluded, as well as conference proceedings, abstracts, dissertations, and study protocols.

## Information sources

Relevant papers were identified through searches in Web of Science, PsycINFO (via EBSCO), and PubMed. PsycINFO and PubMed were chosen to cover both the fields of psychology and biomedical sciences, while Web of Science was included due to its interdisciplinary reach.

## Search strategy

The search string was adapted from the one deployed by Linardon and colleagues<sup>12</sup> to ensure coverage of relevant terms like “smartphone”, “mental health” and “randomized controlled trials” and their variants. Search strings were adjusted to match the formatting requirements of each database (see review protocol for details, <https://osf.io/vwgtv/>). The search was conducted on April 30, 2024. Backward citation tracking was performed by examining the reference list of Linardon and colleagues<sup>12</sup>.

## Selection process

Two master's students trained by the first and last author screened titles and abstracts of potentially eligible studies. A pilot screening of 750 records yielded a Cohen's kappa of 0.72, indicating sufficient interrater reliability to proceed. When eligibility was uncertain, studies were included for full-text screening. Full-text articles were then assessed against predefined inclusion and exclusion criteria by three other trained master's students. Initially, J.K. and F.F. independently screened 10% of the full texts to validate decisions and resolve discrepancies. Reviewers followed a standardized, hierarchical checklist, stopping the process when a study failed to meet any criterion. Remaining uncertainties during the screening of full texts were resolved through discussion.

## Data items

Data items that were extracted included study characteristics (e.g., first author's last name, study year, country), sample characteristics (e.g., mean age, percentage of female participants), and treatment characteristics (e.g., length, presence of personalization or gamification features). For an overview of all data items, see the review protocol (<https://osf.io/vwgty>).

**Active elements.** The presence of 34 pre-registered active elements for each group of every trial was extracted. The active elements were derived from established frameworks, reviews, and trial protocols addressing CBT for depression<sup>51–53</sup>, anxiety<sup>54–57</sup>, and CBT in general<sup>58,59</sup>. In addition, active elements were integrated from “third-wave” CBT interventions<sup>22</sup>, including Acceptance and Commitment Therapy (ACT)<sup>23</sup> and Mindfulness-Based Cognitive Therapy (MBCT)<sup>24</sup>. For ACT, these were cognitive defusion, values, acceptance, present-moment focus, and committed action<sup>23,60</sup>. Consistent with mindfulness-based protocols, “mindfulness” was also added as a distinct active element<sup>61,62</sup>. Finally, four active elements from the field of positive psychology were included that were aimed at enhancing positive behaviors and cognitions, namely gratitude, savoring, optimism, and personal strengths<sup>63–65</sup>. While these might not strictly fall under the umbrella of CBT, we still expected them to be relatively prevalent in mental health apps<sup>28</sup>. We also included them because we see them as complementary to CBT elements, whose focus is rather on dysfunctional cognitions or behaviors. In the Supplementary Information (Table S1), an overview and description of the 34 active elements can be found.

## Data extraction

Data extraction was conducted by four master's students who received a training session from JK and JPS. During the training session, data from one example trial were extracted collaboratively. Afterwards, three papers were extracted independently by the four extractors and J.K. and J.P.S. and discussed inconsistencies and uncertainties. The four extractors then continued with data extraction, while J.K. provided feedback in cases of uncertainty.

Since many specific active elements were expected to be unreported in the published papers, a Qualtrics survey was created to gather additional information on whether specific active elements were used in each group of every trial. The corresponding author of each paper included in the review was contacted via email to indicate, for each of the 34 active elements, whether it was used in each group of the respective trial (coded as yes or no). In case information was missing for data extraction, this survey was sent together with the email in which the missing information was requested. After the initial request, two reminders were sent after two and four weeks. Deviating from the review protocol, a final step was added to the data extraction process in which one trained research assistant manually coded the presence of the 34 active elements for each group of every trial based on published materials. This was necessary due to the relatively low response rate to the Qualtrics survey (36.4%,  $n = 75$  of contacted trial authors responded to the survey). In cases where trial groups were not adequately described to reliably extract active elements, the group was coded as not well-described. In addition, 70 of all trial authors were contacted whose

papers lacked reporting of information relevant for data extraction and analyses (e.g., means, standard deviations, or sample characteristics), of which 39 (56%) responded with the requested information. Only those 70 authors whose papers lacked this information were contacted with the request to provide that additional information. In contrast, authors from all trials were invited to complete the Qualtrics survey to provide details on the active elements included in their interventions.

## Risk of bias

The risk of bias for all included RCTs was assessed using the Cochrane Collaboration Risk of Bias tool, with each criterion rated as “high”, “low”, or “insufficient information”<sup>66</sup>. One trained rater conducted the assessment after receiving training. A senior rater (J.K.) independently evaluated 10% of the articles to compare results, and once a high level of agreement was confirmed, the trained rater proceeded to assess the remaining articles. Based on previous meta-analyses<sup>12,67,68</sup>, it was evaluated if (1) the inclusion and exclusion criteria were adequately described, (2) the method for generating the allocation sequence was truly random, and (3) the allocation sequence was concealed from those assigning participants to intervention groups. It was also examined if (4) participants and study personnel were sufficiently blinded to the assigned intervention, (5) outcome assessors were unaware of the intervention assignments (if only self-report measures were used, risk of bias was coded as “low”). Furthermore, (6) the completeness of outcome data was examined by verifying if an intention-to-treat analysis was performed (if there were no drop-outs, risk of bias was coded as “low”), and (7) risk of bias due to dropouts was evaluated, either through a description of the numbers and reasons, via dropout analysis, or by a flow chart that detailed the dropouts and reasons for dropout.

## Data analyses

Experimental and control groups that were not well-described were excluded from the analyses, as interventions that were not well-described are likely to lead to an underestimation of reported active elements. Sensitivity analyses were conducted with all studies, including trials that were not well-described, and we found that the findings of the primary meta-analyses did not substantially differ. Standardized mean differences were calculated by subtracting the average baseline score from the average post-measurement score, divided by the standard deviation of the difference<sup>69</sup>. This was done for each group within each trial and for each timepoint, resulting in time-varying within-group effect sizes nested within groups and trials. For the calculation of effect sizes, a correlation of 0.5 was assumed<sup>70</sup>, with sensitivity analyses suggesting no substantial differences when correlations of 0.3 or 0.7 were used. Hedge's  $g$  correction was applied to account for bias due to small sample sizes.

For all analyses, a bivariate mixed-effects meta-regression model was used, which is implemented in R (version 4.3.2)<sup>71</sup> in the *metafor* package<sup>72</sup>. The bivariate mixed-effects meta-regression model is a meta-analysis framework that allows for efficient modeling of active elements varying between trials and groups<sup>26,27</sup>. The nested structure of the data was accounted for by including random effects for timepoints nested within groups, and groups nested within trials. This approach utilizes effect sizes from all timepoints and groups, including comparator groups, thereby increasing statistical power and precision. Separate univariable models for each of the 34 active elements were run, in which the respective active element was entered as a fixed effect (coded as 0 or 1). Active elements were coded as varying between groups within and across trials. Since our approach relies on within-group changes over time instead of between-group effect sizes, the calculated standardized mean differences are not based on random assignment to groups. Therefore, as predefined in the study protocol, the following covariates were included in all models to control for potential confounding: age, gender, comorbidity, clinical severity, and length to follow-up. In addition, the variable adjunctive treatment was added as a post-hoc covariate. Furthermore, we conducted exploratory moderation analyses examining whether associations between active elements and clinical outcomes differed by clinical status. To this end,

an interaction term between each active element and clinical status was included in all models. Moderation analyses were restricted to active elements that were present in at least 5% of intervention groups, as elements with lower prevalence were expected to yield poorly identified interaction effects. Clinical status was coded as non-clinical/subclinical (0) versus clinical (1). Descriptions of all covariates are provided in Table 4. Each model was run separately for symptoms of depression and anxiety.

To examine whether results differed depending on the source of element coding, we conducted additional sensitivity analyses. Specifically, we reran all models, again restricted to active elements present in at least 5% of groups, adding an interaction term between the active element and coding source (0 = extractor-coded, 1 = author-reported via survey). In addition, outliers were identified by examining standardized residuals from the meta-regression model, by flagging timepoints with standardized residuals exceeding  $\pm 2.576$  as potential outliers. These studies were visualized and removed in a sensitivity analysis to assess their impact on the results.

In an additional step, we examined whether the number of active elements used in each group is associated with changes in symptoms of depression and anxiety. For this purpose, the sum score was calculated reflecting the number of active elements in each group. We first examined this in a linear meta-regression model. However, because it could also be that the association between the number of elements and improvements in symptoms is nonlinear, we additionally fitted cubic spline models and plotted the association between the number of active elements and changes in symptoms as a nonlinear relationship. Syntax and outputs, including those of all sensitivity analyses, are available on OSF (<https://osf.io/vwgtj>).

### Data availability

The datasets generated and/or analyzed during the current study are not publicly available due to future publications but are available from the corresponding author upon reasonable request.

### Code availability

The underlying code for this study is available on OSF and can be accessed via this link: <https://osf.io/wc59a/>.

Received: 24 July 2025; Accepted: 12 February 2026;

Published online: 11 March 2026

### References

1. Torous, J. et al. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. *World Psychiatry* **20**, 318–335 (2021).
2. Torous, J. et al. Towards a consensus around standards for smartphone apps and digital mental health. *World Psychiatry* **18**, 97 (2019).
3. Chan, S., Torous, J. B., Hinton, L. & Yellowlees, P. M. In *E-mental health*, 217–229 (Springer, 2016).
4. Naslund, J. A. et al. Digital technology for treating and preventing mental disorders in low-income and middle-income countries: a narrative review of the literature. *Lancet Psychiatry* **4**, 486–500 (2017).
5. Naslund, J. A. & Deng, D. Addressing mental health stigma in low-income and middle-income countries: a new frontier for digital mental health. *Ethics Med. Public Health* **19**, 100719 (2021).
6. Chan, S., Godwin, H., Gonzalez, A., Yellowlees, P. M. & Hilty, D. M. Review of use and integration of mobile apps into psychiatric treatments. *Curr. psychiatry Rep.* **19**, 1–9 (2017).
7. Liu, Q. et al. Changes in the global burden of depression from 1990 to 2017: findings from the Global Burden of Disease study. *J. Psychiatr. Res.* **126**, 134–140 (2020).
8. Santomauro, D. F. et al. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. *Lancet* **398**, 1700–1712 (2021).
9. Keynejad, R. C., Dua, T., Barbui, C. & Thornicroft, G. WHO Mental Health Gap Action Programme (mhGAP) Intervention Guide: a systematic review of evidence from low and middle-income countries. *BMJ Ment. Health* **21**, 30–34 (2018).
10. Alonso, J. et al. Treatment gap for anxiety disorders is global: results of the World Mental Health Surveys in 21 countries. *Depress. anxiety* **35**, 195–208 (2018).
11. Goodwin, R. D. et al. Trends in US depression prevalence from 2015 to 2020: the widening treatment gap. *Am. J. Prev. Med.* **63**, 726–733 (2022).
12. Linardon, J. et al. Current evidence on the efficacy of mental health smartphone apps for symptoms of depression and anxiety. A meta-analysis of 176 randomized controlled trials. *World Psychiatry* **23**, 139–149 (2024).
13. Krebs, P. & Duncan, D. T. Health app use among US mobile phone owners: a national survey. *JMIR mHealth uHealth* **3**, e4924 (2015).
14. Firth, J. et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. *World Psychiatry* **16**, 287–298 (2017).
15. Firth, J. et al. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. *J. Affect. Disord.* **218**, 15–22 (2017).
16. Weisel, K. K. et al. Standalone smartphone apps for mental health—a systematic review and meta-analysis. *NPJ Digit. Med.* **2**, 118 (2019).
17. Lecomte, T. et al. Mobile apps for mental health issues: meta-review of meta-analyses. *JMIR mHealth uHealth* **8**, e17458 (2020).
18. Linardon, J., Messer, M., Goldberg, S. B. & Fuller-Tyszkiewicz, M. The efficacy of mindfulness apps on symptoms of depression and anxiety: an updated meta-analysis of randomized controlled trials. *Clin. Psychol. Rev.* **107**, 102370 (2024).
19. Goldberg, S. B., Lam, S. U., Simonsson, O., Torous, J. & Sun, S. Mobile phone-based interventions for mental health: a systematic meta-review of 14 meta-analyses of randomized controlled trials. *PLOS Digit. Health* **1**, e0000002 (2022).
20. Linardon, J. Can acceptance, mindfulness, and self-compassion be learned by smartphone apps? A systematic and meta-analytic review of randomized controlled trials. *Behav. Ther.* **51**, 646–658 (2020).
21. Cohen, Z. D., Barnes-Horowitz, N. M., Forbes, C. N. & Craske, M. G. Measuring the active elements of cognitive-behavioral therapies. *Behav. Res. Ther.* **167**, 104364 (2023).
22. Hayes, S. C. & Hofmann, S. G. Third-wave” cognitive and behavioral therapies and the emergence of a process-based approach to intervention in psychiatry. *World Psychiatry* **20**, 363–375 (2021).
23. Hayes, S. C., Luoma, J. B., Bond, F. W., Masuda, A. & Lillis, J. Acceptance and commitment therapy: model, processes and outcomes. *Behav. Res. Ther.* **44**, 1–25 (2006).
24. Kuyken, W. et al. How does mindfulness-based cognitive therapy work? *Behav. Res. Ther.* **48**, 1105–1112 (2010).
25. Öst, L.-G. Efficacy of the third wave of behavioral therapies: a systematic review and meta-analysis. *Behav. Res. Ther.* **46**, 296–321 (2008).
26. Kraiss, J. et al. Estimating the true effectiveness of smoking cessation interventions under variable comparator conditions: a systematic review and meta-regression. *Addiction* **118**, 1835–1850 (2023).
27. Black, N. et al. Behaviour change techniques associated with smoking cessation in intervention and comparator groups of randomized controlled trials: a systematic review and meta-regression. *Addiction* **115**, 2008–2020 (2020).
28. Marshall, J. M., Dunstan, D. A. & Bartik, W. Apps with maps—anxiety and depression mobile apps with evidence-based frameworks: systematic search of major app stores. *JMIR Ment. Health* **7**, e16525 (2020).
29. de Bruin, M. et al. Underreporting of the active content of behavioural interventions: a systematic review and meta-analysis of randomised trials of smoking cessation interventions. *Health Psychol. Rev.* **15**, 195–213 (2021).
30. Franco, A., Malhotra, N. & Simonovits, G. Underreporting in psychology experiments: evidence from a study registry. *Soc. Psychol. Personal. Sci.* **7**, 8–12 (2016).

31. Gan, D. Z., McGillivray, L., Han, J., Christensen, H. & Torok, M. Effect of engagement with digital interventions on mental health outcomes: a systematic review and meta-analysis. *Front. Digit. Health* **3**, 764079 (2021).
32. Bidargaddi, N., Schrader, G., Klasnja, P., Licinio, J. & Murphy, S. Designing m-Health interventions for precision mental health support. *Transl. Psychiatry* **10**, 222 (2020).
33. Klasnja, P. et al. Microrandomized trials: an experimental design for developing just-in-time adaptive interventions. *Health Psychol.* **34**, 1220 (2015).
34. Hofmann, S. G., Sawyer, A. T., Witt, A. A. & Oh, D. The effect of mindfulness-based therapy on anxiety and depression: A meta-analytic review. *J. Consult. Clin. Psychol.* **78**, 169 (2010).
35. Blanck, P. et al. Effects of mindfulness exercises as stand-alone intervention on symptoms of anxiety and depression: systematic review and meta-analysis. *Behav. Res. Ther.* **102**, 25–35 (2018).
36. Bai, Z., Luo, S., Zhang, L., Wu, S. & Chi, I. Acceptance and commitment therapy (ACT) to reduce depression: a systematic review and meta-analysis. *J. Affect. Disord.* **260**, 728–737 (2020).
37. Chakhssi, F., Kraiss, J. T., Sommers-Spijkerman, M. & Bohlmeijer, E. T. The effect of positive psychology interventions on well-being and distress in clinical samples with psychiatric or somatic disorders: a systematic review and meta-analysis. *BMC Psychiatry* **18**, 211 (2018).
38. Kloos, N., Austin, J., van 't Klooster, J.-W., Drossaert, C. & Bohlmeijer, E. Appreciating the good things in life during the covid-19 pandemic: a randomized controlled trial and evaluation of a gratitude app. *J. Happiness Stud.* **23**, 4001–4025 (2022).
39. Kok, G. et al. A taxonomy of behaviour change methods: an intervention mapping approach. *Health Psychol. Rev.* **10**, 297–312 (2016).
40. Michie, S. et al. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. *Psychol. Health* **26**, 1479–1498 (2011).
41. Iasiello, M. et al. Measuring well-being: creating a taxonomy and item bank for positive mental health. <https://doi.org/10.21203/rs.3.rs-6862291/v1> (2025).
42. Abramowitz, J. S., Deacon, B. J. & Whiteside, S. P. *Exposure Therapy for Anxiety: Principles and Practice* (Guilford Publications, 2019).
43. Schwob, J. T. & Newman, M. G. Brief imaginal exposure exercises for social anxiety disorder: a randomized controlled trial of a self-help momentary intervention app. *J. Anxiety Disord.* **98**, 102749 (2023).
44. Powers, M. B. & Emmelkamp, P. M. Virtual reality exposure therapy for anxiety disorders: a meta-analysis. *J. Anxiety Disord.* **22**, 561–569 (2008).
45. Barlow, D. H., Allen, L. B. & Choate, M. L. In *The Neurotic Paradox*, Vol. 1, 141–166 (Routledge, 2020).
46. Lamers, F. et al. Comorbidity patterns of anxiety and depressive disorders in a large cohort study: the Netherlands Study of Depression and Anxiety (NESDA). *J. Clin. Psychiatry* **72**, 341 (2011).
47. Linardon, J. & Fuller-Tyszkiewicz, M. Attrition and adherence in smartphone-delivered interventions for mental health problems: a systematic and meta-analytic review. *J. Consult. Clin. Psychol.* **88**, 1 (2020).
48. Koelen, J. et al. Dropout and completion in iCBT for university students: insights from a thematic analysis. *Internet Interv.* **40**, 100831 (2025).
49. Lipschitz, J. M., Pike, C. K., Hogan, T. P., Murphy, S. A. & Burdick, K. E. The engagement problem: a review of engagement with digital mental health interventions and recommendations for a path forward. *Curr. Treat. Options Psychiatry* **10**, 119–135 (2023).
50. Page, M. J. et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* **372**, n71 (2021).
51. Cuijpers, P. & Kleiboer, A. M. 40 Self-directed approaches to the treatment of depression. in *The Oxford Handbook of Mood Disorders*, 469–477 (Oxford University Press, 2017).
52. López-López, J. A. et al. The process and delivery of CBT for depression in adults: a systematic review and network meta-analysis. *Psychol. Med.* **49**, 1937–1947 (2019).
53. Watkins, E., Newbold, A., Tester-Jones, M., Collins, L. & Mostazir, M. Investigation of active ingredients within internet cognitive-behavioral therapy for depression: a randomized optimization trial. *JAMA Psychiatry*, **80**, 942–951 (2023).
54. Behar, E., DiMarco, I. D., Hekler, E. B., Mohlman, J. & Staples, A. M. Current theoretical models of generalized anxiety disorder (GAD): conceptual review and treatment implications. *J. Anxiety Disord.* **23**, 1011–1023 (2009).
55. Clark, D. M. et al. Cognitive therapy versus fluoxetine in generalized social phobia: a randomized placebo-controlled trial. *J. Consult. Clin. Psychol.* **71**, 1058 (2003).
56. Kaczurkin, A. N. & Foa, E. B. Cognitive-behavioral therapy for anxiety disorders: an update on the empirical evidence. *Dialogues Clin. Neurosci.* **17**, 337–346 (2015).
57. Newman, M. G., Zainal, N. H. & Hoyer, J. Cognitive-behavioral therapy (CBT) for generalized anxiety disorder (GAD). in *Generalized Anxiety Disorder and Worrying: a Comprehensive Handbook for Clinicians and Researchers*, 203–230 (2020).
58. Bennett-Levy, J. Mechanisms of change in cognitive therapy: the case of automatic thought records and behavioural experiments. *Behav. Cogn. Psychother.* **31**, 261–277 (2003).
59. Rice, R. H. Cognitive-behavioral therapy. *Sage Encycl. Theory Couns. Psychother.* **1**, 194 (2015).
60. van Aubel, E. et al. Blended care in the treatment of subthreshold symptoms of depression and psychosis in emerging adults: a randomised controlled trial of Acceptance and Commitment Therapy in Daily-Life (ACT-DL). *Behav. Res. Ther.* **128**, 103592 (2020).
61. van Emmerik, A. A., Berings, F. & Lancee, J. Efficacy of a mindfulness-based mobile application: a randomized waiting-list controlled trial. *Mindfulness* **9**, 187–198 (2018).
62. Kabat-Zinn, J. An outpatient program in behavioral medicine for chronic pain patients based on the practice of mindfulness meditation: theoretical considerations and preliminary results. *Gen. Hosp. Psychiatry* **4**, 33–47 (1982).
63. Carr, A. et al. Effectiveness of positive psychology interventions: a systematic review and meta-analysis. *J. Posit. Psychol.* **16**, 749–769 (2021).
64. Geschwind, N., Arntz, A., Bannink, F. & Peeters, F. Positive cognitive behavior therapy in the treatment of depression: a randomized order within-subject comparison with traditional cognitive behavior therapy. *Behav. Res. Ther.* **116**, 119–130 (2019).
65. Van Agteren, J. et al. A systematic review and meta-analysis of psychological interventions to improve mental wellbeing. *Nat. Hum. Behav.* **5**, 631–652 (2021).
66. Higgins, J. P. et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* **343**, d5928 (2011).
67. Kraiss, J. T., Redelinghuys, K. & Weiss, L. A. The effects of psychological interventions on well-being measured with the Mental Health Continuum: a meta-analysis. *J. Happiness Stud.* **23**, 3655–3689 (2022).
68. Tönis, K. J., Kraiss, J. T., Linssen, G. C. & Bohlmeijer, E. T. The effects of positive psychology interventions on well-being and distress in patients with cardiovascular diseases: a systematic review and Meta-analysis. *J. Psychosom. Res.* **170**, 111328 (2023).
69. Harrer, M., Cuijpers, P., Furukawa, T. A. & Ebert, D. D. *Doing Meta-Analysis With R: a Hands-On Guide*. 1st edn (Chapman & Hall/CRC Press, 2021).
70. Harrer, M. et al. Standardized effect sizes are far from “Standardized”: a primer and empirical illustration in depression psychotherapy meta-analyses. *PLoS Ment. Health* **2**, e0000347 (2025).
71. R Core Team: A language and environment for statistical computing (R Foundation for Statistical Computing, Vienna, Austria, 2025).
72. Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* **36**, 1–48 (2010).

## Acknowledgements

We acknowledge all master-level graduate students involved in study selection and data extraction, especially Angelika Freier, Rosalia Mardjo, Insa Holtkamp, and Nefise Aydin. We also want to thank all trial authors who responded to the Qualtrics survey and our requests for additional data. No funding was received for this study.

## Author contributions

J.K. and J.P.S. drafted the original idea for this study, supervised data collection, prepared the data, and wrote the initial manuscript. J.K. drafted the study protocol and performed the data analysis and visualization. FF contributed to data extraction and validation. F.C. and J.A.K., drawing on their clinical expertise, supported the development of the study protocol and the framework of active elements. F.B.A. conducted the data extraction of active elements and co-developed the coding scheme for the coding of active elements. All authors reviewed and provided feedback on the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41746-026-02466-z>.

**Correspondence** and requests for materials should be addressed to Jannis Kraiss.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2026