

<https://doi.org/10.1038/s41746-025-01567-5>

# A meta-analysis of persuasive design, engagement, and efficacy in 92 RCTs of mental health apps



Lee Valentine<sup>1,2</sup>✉, Jordan D. X. Hinton<sup>3</sup>, Kriti Bajaj<sup>1</sup>, Larissa Boyd<sup>1</sup>, Shaunagh O'Sullivan<sup>1</sup>, Rory P. Sorenson<sup>1,2</sup>, Imogen H. Bell<sup>1,2</sup>, Miguel Sobredo Vega<sup>4</sup>, Ping Liu<sup>1,2</sup>, Wilma Peters<sup>1,2</sup>, Shaminka N. Mangelsdorf<sup>1,2</sup>, Thomas W. Wren<sup>1,2</sup>, Carl Moller<sup>5</sup>, Shane Cross<sup>1,2</sup>, Carla McEnery<sup>1,2</sup>, Sarah Bendall<sup>1,2</sup>, Jennifer Nicholas<sup>1,2,6</sup> & Mario Alvarez-Jimenez<sup>1,2,6</sup>

This systematic review and meta-analysis examined the efficacy of digital mental health apps and the impact of persuasive design principles on intervention engagement and outcomes. Ninety-two RCTs and 16,728 participants were included in the meta-analyses. Findings indicate that apps significantly improved clinical outcomes compared to controls ( $g = 0.43$ ). Persuasive design principles ranged from 1 to 12 per app (mode = 5). Engagement data were reported in 76% of studies, with 25 distinct engagement metrics identified, the most common being the percentage of users who completed the intervention and the average percentage of modules completed. No significant association was found between persuasive principles and either efficacy or engagement. With 25 distinct engagement metrics and 24% of studies not reporting engagement data, establishing overall engagement with mental health apps remains unfeasible. Standardising the definition of engagement and implementing a structured framework for reporting engagement metrics and persuasive design elements are essential steps toward advancing effective, engaging interventions in real-world settings.

Mental health care systems face a multitude of serious challenges, including an ever-increasing demand for mental health support, insufficient funding, employment insecurity, high staff turnover, understaffing, and clinician burnout<sup>1–4</sup>. Concurrently, people seeking mental health support frequently encounter barriers to care such as long waitlists, reduced appointment availability, prohibitive financial costs, and limited access to ongoing sessions<sup>5–7</sup>.

The growing need for mental health care, alongside the proliferation of smartphone ownership, has catalysed the development of innovative digital interventions designed to address unmet mental health needs<sup>2,8</sup>. These technologies represent a paradigm shift in mental health care delivery<sup>9</sup> by improving treatment accessibility and offering the potential to mitigate pressure on overburdened in-person services. By providing scalable, cost-effective, and evidence-based treatment options, digital interventions present a promising approach to meeting the increasing demand for mental health support<sup>10–12</sup>.

Digital mental health apps can be self-guided and fully automated, allowing users to access support independent of a clinician<sup>13</sup>. These apps

offer the potential to increase access to support, reduce costs<sup>11,14,15</sup>, and help manage wait times for clinical services<sup>16</sup>. Digital interventions can also complement traditional face-to-face care through 'blended care' models, in which clinicians integrate technology into their practice to enhance both the online and offline components of care<sup>17–20</sup>. For example, clinicians may incorporate digital mood tracking between sessions, provide access to complementary online psychosocial content to reinforce therapeutic concepts, and use synchronous (e.g., real-time chat or video calls) or asynchronous (e.g., secure messaging) online communication to maintain engagement and support outside of scheduled appointments. Moreover, emerging digital interventions have the potential to address critical gaps across various phases of care—such as waitlists, discharge, and relapse prevention—thereby enhancing their value throughout the treatment journey<sup>21</sup>. The scalability of digital interventions holds promise for delivering personalised treatments to a much larger population than traditional face-to-face services can accommodate<sup>2,22</sup>. Consequently, evidence-based digital technology has emerged as a valuable resource for addressing the

<sup>1</sup>Orygen, Melbourne, VIC, Australia. <sup>2</sup>Centre for Youth Mental Health, The University of Melbourne, Melbourne, VIC, Australia. <sup>3</sup>Australian Research Centre in Sex, Health, and Society, Latrobe University, Melbourne, VIC, Australia. <sup>4</sup>Hospital Universitario Infanta Leonor, Madrid, Spain. <sup>5</sup>The Institute for Mental and Physical Health and Clinical Translation (IMPACT), Deakin University, Melbourne, VIC, Australia. <sup>6</sup>These authors contributed equally: Jennifer Nicholas, Mario Alvarez-Jimenez. ✉e-mail: [lee.valentine@unimelb.edu.au](mailto:lee.valentine@unimelb.edu.au)

disparity between the demand for mental health care and the available supply<sup>1</sup>.

Digital mental health interventions are generally well-regarded by clinicians and young people<sup>23</sup>. They have demonstrated efficacy in both high-prevalence disorders such as depression and anxiety<sup>12,24,25</sup>, as well as complex and severe disorders such as psychosis<sup>26</sup>. Additionally, there are indications that digital interventions can be cost-effective compared with alternate treatments<sup>11,27</sup>. However, low user engagement remains a significant and long-standing issue<sup>3,28–30</sup>. In a review of real-world user engagement with popular mental health apps, Baume et al.<sup>31</sup> reported an engagement rate of 3.9% by day 15, which declined to 3.3% by day 30. In a systematic review conducted by Fleming et al.<sup>29</sup> examining user engagement in samples of people with depression and anxiety, sustained usage varied widely, ranging from 0.5% to 28.6%. These low and inconsistent rates of engagement, despite demonstrating positive outcomes in controlled environments<sup>12</sup>, raise concerns about the clinical utility and effectiveness of digital mental health interventions in real-world scenarios<sup>3</sup>.

Overcoming the engagement challenge is an essential component in bridging the gap between potential effectiveness and practical impact<sup>2</sup>. To be efficacious, digital interventions must enable offline therapeutic action—desired health behaviours initiated within a digital mental health app that are subsequently adopted into real-world settings<sup>32</sup>. This aligns with Cole-Lewis et al.'s<sup>33</sup> concept of “Big E” (health behaviour engagement) proposed in the framework for digital behaviour change interventions, which emphasises the importance of engagement in leading to actual health behaviour changes. According to this model, achieving Big E—the real-world adoption of health behaviours—is contingent on “Little e” (user engagement with the app's features), which includes interactions with both app elements (e.g., games) and embedded behaviour change techniques designed to influence health outcomes (e.g., providing choice to support autonomy, as informed by Self-Determination Theory)<sup>47</sup>. The framework underscores that meaningful interaction with a digital intervention is a necessary precursor to real-world change, reinforcing that without engagement, even the most well-designed interventions are unlikely to achieve their intended impact.

In digital health literature, there is no universally accepted definition of user engagement<sup>3</sup>. This review adopts Borghouts et al.<sup>34</sup> comprehensive definition of engagement, which includes both initial adoption and continued interaction with the digital intervention, evidenced by behaviours such as signing up, using its features, and sustained use over time.

While the causal link between engagement and intervention efficacy remains unclear, it is broadly acknowledged that some degree of engagement is necessary for users to benefit from an intervention. Further, a lack of engagement complicates attributing positive outcomes to the intervention<sup>3</sup>. The relationship between user engagement and intervention effectiveness is complex and presumed to be influenced by various factors, such as type of engagement, the intervention itself, and individual user characteristics<sup>34</sup>. A nuanced understanding of the diverse and potentially interconnected factors influencing engagement is essential for shaping user behaviour and translating clinical efficacy into real-world outcomes.

The exploration of persuasive design holds promise for enhancing user engagement and supporting improved intervention outcomes<sup>30,35,36</sup>. The concept of persuasive design was developed to leverage technology to positively influence behaviour change and discourage harmful behaviour at an individual level<sup>37</sup>. This method is based on the concept that technology can serve more than just a functional purpose; it can also act as a catalyst to promote and support targeted behaviours, emotions and mental states<sup>35,37</sup>. Through the strategic application of persuasive design principles and strategies, these systems aim to motivate and support users in fostering positive shifts in attitudes and behaviours<sup>35–38</sup>.

The persuasive systems design (PSD) framework, proposed by Oinas-Kukkonen and Harjuma<sup>38</sup>, consists of 28 principles categorised into four domains: (1) primary task support, which facilitates the primary goal of the intervention, (2) dialogue support, which enables communication between the intervention and the user, (3) system credibility support, which enhances the trustworthiness and credibility of the intervention; and (4) social support, which leverages a social experience within the intervention. Each domain is further broken down into 7 distinct persuasive principles. These principles serve as a roadmap for developers to craft persuasive systems that produce more compelling products to engage users and foster positive behaviour change over time<sup>39</sup>. See Table 1 for more information.

**Table 1 | Persuasive systems design framework<sup>38</sup>**

<b>Primary task support:</b> facilitating the completion of the primary tasks or goals that the user aims to achieve through the system	<b>Dialogue support:</b> enabling effective communication and interaction between the system and the user	<b>Credibility support:</b> enhancing the perceived trustworthiness and credibility of the system	<b>Social support:</b> incorporating social influence and social interaction features within the system
Reduction: reducing complex behaviours into manageable tasks	Praise: providing praise or positive feedback for using the system	Trustworthiness: information provided should be fair and objective	Social learning: providing a way for users to observe the outcomes of others performing target behaviours
Tunnelling: guiding the user through a process or experience in a manageable way	Rewards: providing rewards to users for using the system/ achieving target behaviours	Expertise: information provided should demonstrate knowledge, experience, and competence	Social comparison: providing means for users to compare their performance with other users
Tailoring: customising information to reflect the needs, interests, personality and usage context of a user group	Reminders: remind users of their target behaviour	Surface credibility: the system conveys a competent appearance	Normative influence: facilitate the gathering of users with similar goals, fostering a sense of community and shared norms
Personalisation: providing individualised content and experiences	Suggestion: offer relevant suggestions to the user	Real-world feel: highlight the people or organisation behind the system	Social facilitation: users can perceive that others on the system are performing the target behaviour
Self-monitoring: keeping track of a user's performance or status	Similarity: the system reflects the user or user group	Authority: the system should reference people in roles of authority	Cooperation: users can cooperate to meet a common goal
Simulation: providing opportunities to engage in an experience allows users to directly observe the immediate cause-and-effect relationships of the target behaviour	Liking: a visually attractive system	Third-party endorsements: the system should reference endorsements from respected sources	Competition: users can compete against other users
Rehearsal: providing opportunities to practice behaviour to help users prepare for real-world situations	Social role: the system assumes a social, personable, or human-like manner	Verifiability: providing links to third-party sources to verify the accuracy of the system's content	Recognition: users who perform target behaviours should be recognised publicly on the system

Although Limited research has examined the role of PSD principles in digital interventions for mental health, Kelders et al.<sup>30</sup> identified a significant relationship between user adherence to web-based health interventions and the application of persuasive design elements. While the study did not focus solely on mental health, it found that design principles within the dialogue support domain—which aim to facilitate effective communication between the system and the user—were associated with increased adherence. As Kelders et al.<sup>30</sup> note, the study was constrained by a lack of usable reported usage data, as well as by the coding process, which relied solely on published descriptions. The authors recommend that future research further investigate the relationship between persuasive technology—particularly primary task support—and clinical outcomes in digital interventions.

Orji and Moffatt<sup>36</sup> reviewed 16 years of literature on persuasive technology in health and wellness. The authors reported that although 92% of the studies reported positive results, the review did not identify a statistically significant correlation between the use of persuasive principles and the intervention outcomes. In 2021, McCall et al.<sup>25</sup> conducted a systematic review, meta-analysis, and meta-regression to examine the impact of persuasive design principles in self-directed eHealth interventions. Their findings provided modest preliminary support that interventions utilising more persuasive elements in the primary task domain were more effective for treating depression, but not anxiety. In contrast, Wu et al.<sup>40</sup>, in a separate meta-analysis of persuasive design in smartphone apps for anxiety and depression, reported a negative association between persuasive design principles and engagement, as measured by completion rates, further highlighting the inconsistent evidence surrounding the role of persuasive design in digital interventions.

Given the rapid proliferation of mental health apps and the persistent challenge of sustaining user engagement, reviewing the application of persuasive design principles, and their impact on user engagement rates and overall effectiveness, is both relevant and timely. Previous research has often focused on broader eHealth interventions beyond mental health, restricted the scope to web-based interventions, or concentrated on specific conditions such as depression and anxiety. In contrast, this review takes a platform-specific approach by examining smartphone apps designed to address a broad range of mental health conditions. This focus helps address existing gaps and provides a more comprehensive understanding of how persuasive design principles influence both engagement and intervention efficacy across diverse mental health domains. Specifically, this review aims to (1) conduct a systematic review and meta-analysis of randomised controlled trials of smartphone mental health apps, (2) systematically assess the prevalence and types of persuasive design principles used in these interventions, and (3), via meta-regression, examine the relationships between persuasive design principles and the efficacy and engagement levels of digital mental health apps. Comprehensively exploring these factors will provide valuable insights to guide the development and refinement of mental health apps, ultimately contributing to enhanced mental health outcomes for users.

## Results

The search identified 5030 records, with 4028 remaining after duplicate removal. Following title and abstract screening, 390 articles proceeded to full text review. Of these, 119 met the inclusion criteria for the systematic review, with 92 providing sufficient data for meta-analysis. The remaining 27 studies were included in the narrative synthesis but excluded from meta-analysis due to insufficient pre/post-outcome data or limited intervention details. The PRISMA 2020 flow diagram outlining the selection process is presented in Fig. 1.

Title and abstract screening was conducted independently by two authors (L.V. and J.D.X.H.), with proportional agreement rates of 89.81% (3615/4025) for title and abstract screening and 77.24% (275/356) for full-text screening.

## Study characteristics

A total of 119 studies, comprising 30,251 participants, were included in the systematic review. Studies were published between 2016 and 2022 and spanned 27 countries, with the largest proportion of participants from the

United States (31.1%), followed by the United Kingdom and Germany (8.4% each) and Australia (7.6%). Sample types varied, with 36 studies (30.3%) including non-clinical populations, 48 (40.3%) targeting subclinical populations, and 35 (29.4%) focusing on clinical samples. Mean participant age ranged from 14 years to 60 years (median = 34 years). Most interventions ( $n = 80$ , 67%) were fully self-guided, while 39 (33%) incorporated human support. Apps were more likely to include human support in clinical (60%, 21/35) and subclinical (29%, 14/48) populations than in non-clinical samples (11%, 4/36). Intervention duration ranged from 10 days to 18 months.

Depression was the most commonly targeted mental health condition, accounting for 25.64% of the studies. A total of 11% of the studies focused on a combination of depression, anxiety, and stress, whereas 9.4% of the studies focused on stress alone. Other mental health conditions targeted included eating disorders and body dissatisfaction/dysmorphic disorder (11%), generalised anxiety (7.69%), post-traumatic stress disorder (PTSD; 5.98%), general mental health and wellness (4.27%), psychotic spectrum disorders (3.4%), postnatal depression (3.4%), psychological distress (3.4%), a combination of psychotic spectrum and bipolar disorder (2.56%), non-suicidal and suicidal injury (1.7%), obsessive-compulsive disorders (OCD; 1.7%), agoraphobia and panic disorder (1.7%), suicidal ideation (0.85%), sleep disturbance (0.85%), resilience (0.85%), loneliness (0.85%), burnout (0.85%), and bipolar disorder (0.85%).

Most interventions were grounded in one or more psychological frameworks. Cognitive-behavioural therapy (CBT) was the most commonly used (60.5%,  $n = 72$ ), followed by mindfulness-based approaches (23.5%,  $n = 28$ ), Acceptance and Commitment Therapy (ACT) (4.2%,  $n = 5$ ), and Dialectical Behaviour Therapy (DBT) (2.5%,  $n = 3$ ). A small proportion of studies (5%,  $n = 6$ ) did not specify a therapeutic framework. A detailed summary of study characteristics is presented in the supplementary information.

## Risk of bias

The risk of bias was assessed using the Cochrane Risk of Bias Tool version 2<sup>41</sup> across five domains. Of the 119 studies included, 7% were classified as having a high risk of bias in at least one domain, while 13% were rated as having a low risk of bias across all domains. The majority (80%) were categorised as having either low or unclear risk in at least one domain. A comprehensive assessment of the risk of bias across the five categories for each study is presented in Table 1 of the supplementary information.

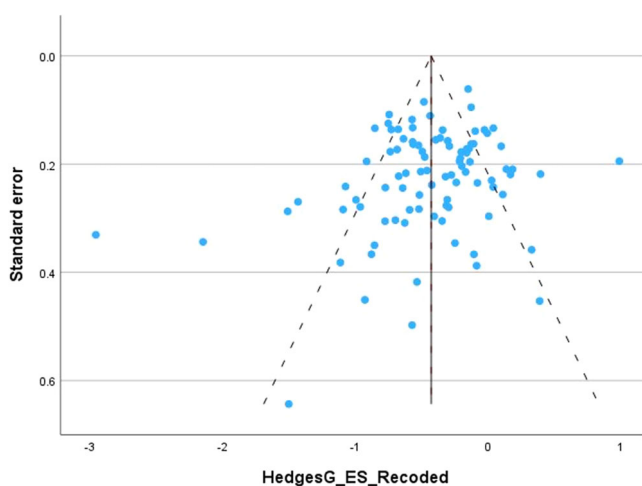
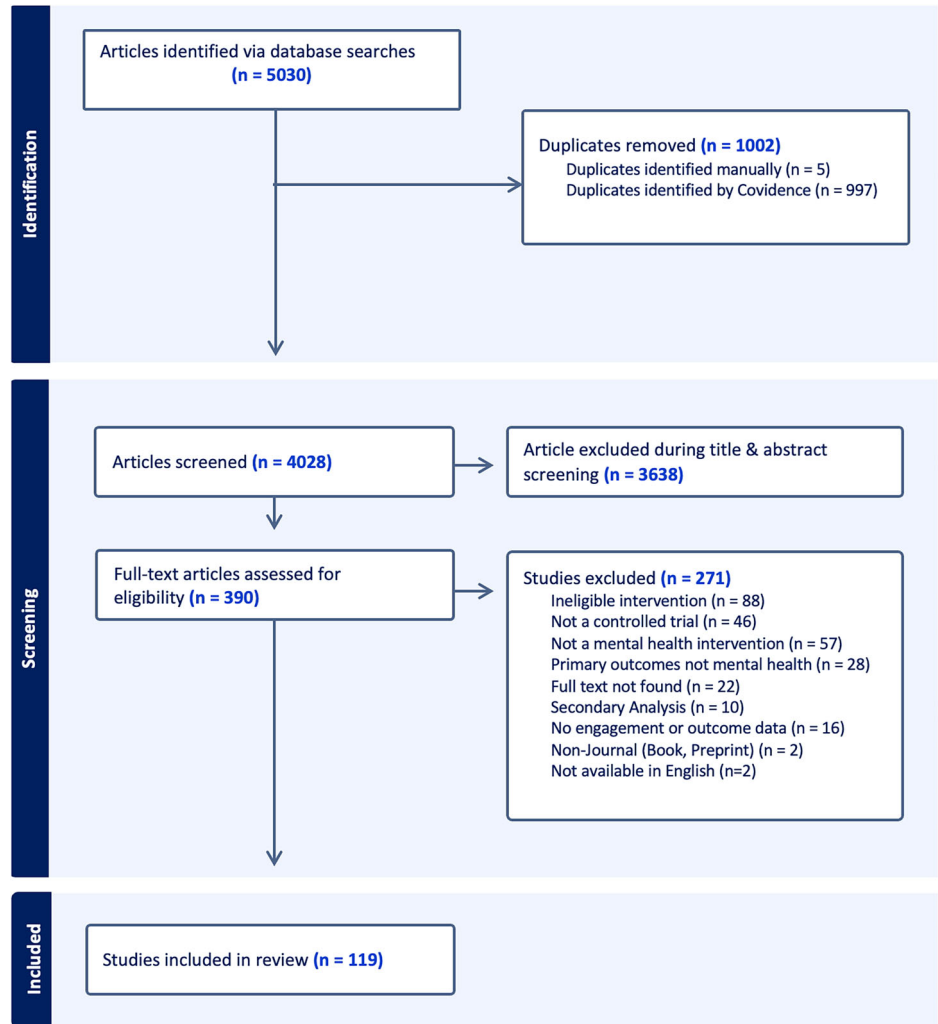
## Publication bias

Publication bias for the overall effects meta-analysis was explored via visual inspection of the funnel plot of standard errors and trim and fill estimates. As shown in Fig. 2, the funnel plot appears symmetrical across both sides of the final estimate. Furthermore, trim and fill analyses revealed no imputed studies on the left- or right-hand side of this estimate. Hence, we noted limited evidence of publication bias.

## Efficacy: overall and sub-group effects of intervention outcomes

Of all studies included in this review, 92 ( $N = 16,782$  total participants) provided sufficient information for effect sizes to be calculated for intervention efficacy and were therefore included in the meta-analysis. Across these studies, sample sizes ranged from 16 to 2271 participants ( $M = 182.41$ ,  $SD = 266.53$ ) at study intake. There was an overall significant effect of intervention efficacy, with a medium size (Hedges  $g = 0.43$ , 95% CI  $[-0.53, -0.34]$ ,  $p < 0.001$ ) indicating that at the post-intervention time point across studies, those in the intervention groups showed significant improvements in mental health outcomes compared with those in the control groups. Notably, however, there was substantial heterogeneity present within this effect ( $I^2 = 83.4$ , range 46.3–98.7). Thus, we examined potential moderators to try and explain some of this variance. Table 2. displays the meta-analytic effects and heterogeneity estimates for both the total effect and sub-group moderation analyses, and Fig. 3. presents the forest plot of effect estimates per study.

**Fig. 1 | A total of 5030 records were identified via database searches.** After removing 1002 duplicates, 4028 articles underwent title and abstract screening, with 3638 excluded. Full-text eligibility assessment was conducted for 390 articles, of which 271 were excluded. The final review included 119 studies.



**Fig. 2 | The funnel plot visualises the distribution of standard errors for primary studies included in the overall effects meta-analysis.** Blue dots represent individual studies, while dashed lines indicate the 95% pseudo-confidence intervals. The estimated overall effect size (observed studies only) is shown as a dashed line, and the estimated effect size with imputed studies (if any) is shown as a solid black line. The plot appears symmetrical, and trim-and-fill analyses identified no imputed studies, suggesting limited evidence of publication bias in the meta-analysis.

The sub-group moderation analysis revealed that the overall effect of intervention efficacy did not significantly differ according to sample type ( $Q(2) = 3.86, p = 0.145$ ), outcome type ( $Q(9) = 14.48, p = 0.106$ ), the presence of human support in interventions ( $Q(1) = 0.07, p = 0.789$ ), or intervention type ( $Q(4) = 3.40, p = 0.493$ ) (Table 2).

The analysis demonstrated that digital interventions were significantly more effective than control conditions in reducing symptoms of depression, anxiety, stress, PTSD, and body image/eating disorders. No significant effects were observed for social anxiety, psychosis, suicide/self-harm, broad mental health outcomes, or postnatal depression. However, it is important to note that these outcomes were each represented by only 2 to 4 studies, limiting the strength of the pooled estimates. In contrast, depression, anxiety, and stress were represented by 8 to 29 studies each, allowing for more robust and reliable effect size estimates.

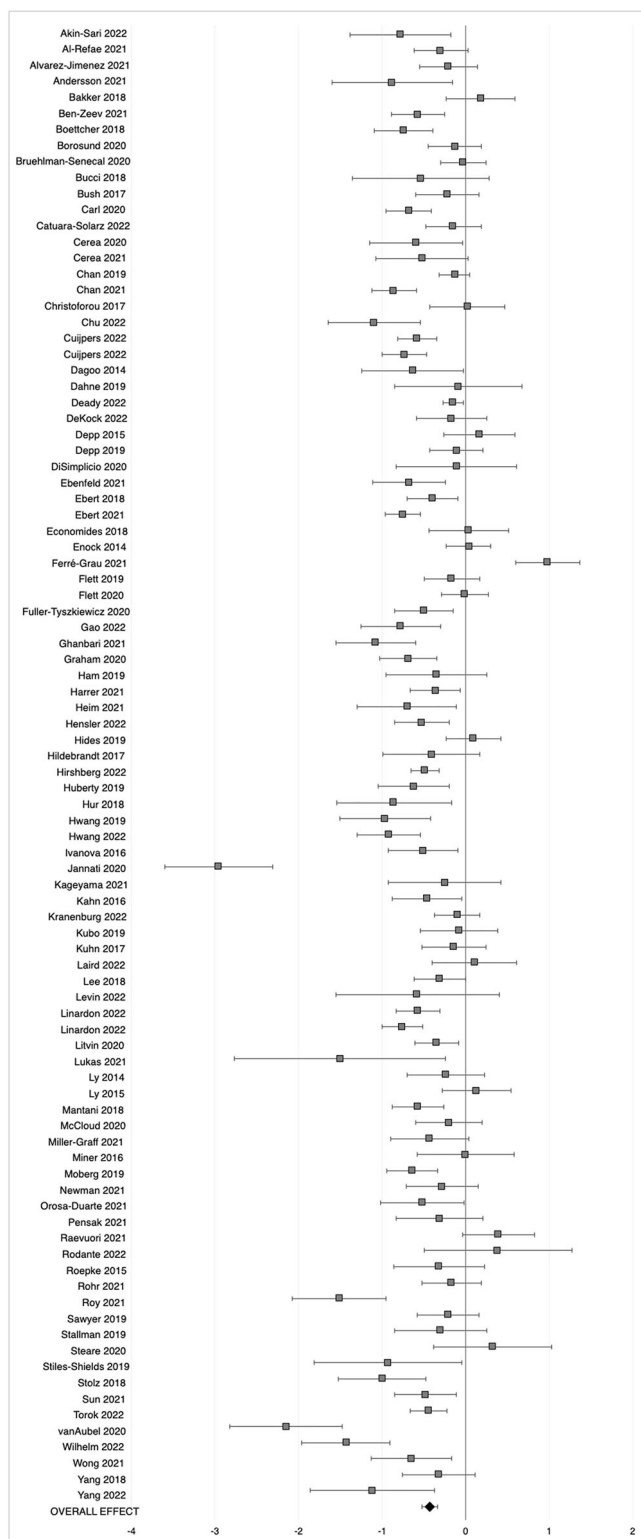
### Engagement: overview and key metrics across studies

Seventy-six percent ( $n = 90$ ) of studies included in the systematic review provided data on engagement, while the remaining 25% ( $n = 29$ ) did not. Among the papers reporting engagement data, a total of 25 distinct engagement indicators were identified. These indicators were grouped into ten overall engagement metrics: (1) rate of uptake, (2) time (min/h) spent on the app, (3) days of active use, (4) logins, (5) modules completed, (6) study metrics, (7) messages sent and received, (8) posts and comments made, (9) participant self-reports, and (10) miscellaneous; see Table 3 for details. The most commonly reported user engagement metrics were: the percentage of participants who completed

Table 2 | Meta-analytic and heterogeneity estimates for the overall effect, and sub-group moderation effects, of intervention efficacy

	Effect size estimates			Heterogeneity estimates	
	k	Hedge's g	SE	[95% CI]	p
Overall effect	92	-0.43	0.05	[-0.53, -0.34]	<001
Sample type					
Non-clinical	28	-0.30	0.08	[-0.45, -0.15]	<001
Sub-clinical	37	-0.49	0.07	[-0.62, -0.36]	<001
Clinical	27	-0.49	0.12	[-0.72, -0.26]	<001
Study outcome					
Depression	29	-0.42	0.08	[-0.57, -0.27]	<001
Anxiety	8	-0.71	0.15	[-1.01, -0.41]	<001
Stress	21	-0.38	0.07	[-0.52, -0.24]	<001
Social anxiety	4	-0.55	0.24	[-1.02, -0.08]	0.021
PTSD	5	-0.28	0.10	[-0.48, -0.09]	0.005
Body image/ eating disorder	5	-0.73	0.15	[-1.03, 0.43]	<001
Psychosis	2	-0.08	0.43	[-0.93, 76]	0.851
Suicide/ self-harm	3	-0.21	0.23	[-0.66, 24]	0.358
Mental health (broadly)	2	0.43	0.55	[-0.64, 1.51]	0.433
Postnatal depression	3	-1.08	0.92	[-2.88, 72]	0.240
Human support within intervention					
No	66	-0.44	0.06	[-0.55, -0.33]	<001
Yes	26	-0.41	0.09	[-0.59, -0.23]	<001
Intervention type					
CBT (only)	34	-0.54	0.09	[-0.71, -0.37]	<001
CBT (+ others)	11	-0.37	0.12	[-0.62, -0.13]	0.003
Mindfulness-based	18	-0.37	0.09	[-0.53, -0.20]	<001
Psychoeducation	6	-0.53	0.14	[-0.81, -0.24]	<001
Others	20	-0.33	0.13	[-0.58, -0.07]	0.012

k number of studies included.



**Fig. 3 |** The figure displays individual study effect size estimates (Hedges'  $g$ ) with corresponding 95% confidence intervals. Squares represent the point estimates of intervention efficacy. Horizontal lines indicate confidence intervals, and the vertical line at zero represents no effect. Studies to the left of zero reflect a reduction in symptoms and thus favour the intervention, while those to the right suggest a negative or null effect.

the entire programme or completed the programme per protocol, ranging from 8%<sup>42</sup> to 100%<sup>43</sup>; the average percentage of modules completed, ranging from 30.95%<sup>44</sup> to 98%<sup>45</sup>; and the mean number of logins or visits to the app during the intervention period, ranging from 11.52 logins<sup>46</sup> to 106.84 logins.<sup>26</sup>

A correlation analysis was conducted to examine the relationship between engagement and efficacy. Due to substantial variation in engagement measurement across studies, the analysis focused on the percentage of programme or per-protocol completion ( $n = 17$ ), the most commonly reported metric, alongside pre- and post-intervention effectiveness data. No significant relationship was identified (Table 4).

### Persuasive design principles: identification, frequency, and additional design features

None of the studies included in this review explicitly identified the use of specific persuasive design principles. However, 88% ( $n = 81$ ) of the studies included in the meta-analysis provided sufficient descriptions of the app intervention, or referenced protocol papers, development papers, websites, or app store information, which allowed for deductive coding based on the PSD framework<sup>38</sup>.

A total of 22 out of the 28 persuasive principles in the PSD framework were identified across the reviewed apps. The number of principles per app ranged from 1 to 12, with a mode of 5. Table 5 presents the persuasive principles identified in the review, the percentage of apps that incorporated each principle, and updated examples of how these principles were operationalised in the reviewed apps. Among the four domains, principles from the primary task support domain, which facilitates the completion of primary tasks or goals that the user seeks to achieve through the system, were the most frequently coded, accounting for 50% of the persuasive principles identified. This was followed by dialogue support (23%) and credibility support (22%). Only 5% of the persuasive principles were derived from the social support domain, which includes principles such as social learning and cooperation. Tunnelling, which is intended to guide the user through a process or experience in a manageable way, was the most frequently identified principle overall, and was present in 88% of the apps. Rehearsal, which provides opportunities to practice behaviour to help users prepare for real-world situations, was identified in 84% of the apps. 5.

In addition to the persuasive principles outlined in the PSD framework, this study identified additional recurring design features in the evaluated apps that may influence user behaviour and motivation. These features, summarised in Table 6, include: (1) staged disclosure, (2) goal setting, (3) explicit self-pacing, and (4) limited use.

The identified features are grounded in motivational theories, including self-determination theory<sup>47</sup> and goal setting theory<sup>48</sup>. These features promote progression, autonomy, and competence, enhancing motivation and engagement by balancing user control with elements of novelty, competition, and exclusivity.

Figure 4 uses a heatmap to illustrate the percentage of digital mental health interventions incorporating PSD features across specific mental health conditions. To reduce heterogeneity, only interventions targeting a single mental health condition with more than one intervention per condition were included, resulting in the analysis of 61 apps. Each cell represents the percentage of interventions using a specific PSD principle, with red tones indicating higher prevalence and green tones representing lower prevalence.

Notable patterns include a higher prevalence of primary task support and dialogue support principles, with limited use of social support features. *Rehearsal*, *Trustworthiness*, and *Tunnelling* were the most commonly applied principles. In contrast, social features such as *Social Comparison*, *Cooperation* and *Social Learning* were rarely used, highlighting their limited current use in digital mental health apps.

### Meta-regression: persuasive design principles and intervention efficacy

To explore the relationship between overall intervention efficacy and persuasive design principles, we conducted a meta-regression in which the number of persuasive design principles used across studies was regressed onto effect size estimates (Hedges'  $g$ ). Figure 5 displays a scatterplot illustrating the relationship between these variables. Overall, the analysis revealed no relationship between the number of persuasive design principles and intervention efficacy across studies,  $b = 0.01$ ,  $SE = 0.02$ , 95% CI =

**Table 3 | Range of user engagement metrics across digital mental health interventions**

User engagement metric	Definition	Number of studies	Range (lowest and highest reported measurement in this category)
Rate of uptake	The number of participants that downloaded the intervention/created an account/commenced the intervention/activated the app/ used the app/ logged on at least once	11	67.5% ( $n = 27/40$ ) <sup>55</sup> to 99% ( $n = 127/128$ ) <sup>56</sup>
Time (min/h) spent on app	Mean time spent on app over trial (min)	6	41.14 <sup>57</sup> –530.4 min <sup>58</sup>
	Median time spent on app over trial (min)	2	26.8 <sup>59</sup> –318 min <sup>58</sup>
	Mean daily use (min)	2	10.8 <sup>60</sup> –15.25 min <sup>61</sup>
	Median daily use (min)	1	13.93 min <sup>61</sup>
	Mean hours spent on app per week (h)	1	3.10 <sup>62</sup>
	Mean weekly use (min)	2	37.9 <sup>62</sup> –103 min <sup>63</sup>
Days of active use	Mean number of days of active use overall	5	6.38 <sup>55</sup> –23.07 days <sup>57</sup>
	Median number of days of active use overall	1	17 days <sup>61</sup>
Logins	Percentage of users logged on at least once a day	3	94% <sup>64</sup> –96% <sup>65</sup>
	Average percentage of logins daily	1	75% <sup>66</sup>
	Mean number of logins/visits over the intervention period	8	11.52 <sup>46</sup> –106.84 logins <sup>26</sup>
	Median logins to app overall	3	9.5 <sup>67</sup> –19 logins <sup>68</sup>
Modules completed	Percentage of users that completed the first/introductory module	6	60.1% <sup>56</sup> – 95% <sup>56,69</sup>
	Percentage of users that completed the whole programme OR percentage of users that completed the programme per protocol	17	7/84 (8%) <sup>12</sup> –25/25 (100%) <sup>43</sup>
	Mean percentage of modules completed	13	12.38/40 (30.95%) <sup>44</sup> –6.84/7 (98%) <sup>45</sup>
	Median completion of modules	1	5 modules (5/8 = 63%) <sup>58</sup>
Study metrics	Number of trial surveys completed	2	48% <sup>70,71</sup> – 68.7% <sup>70</sup>
	Number of participants still engaged at the completion of the study	2	8.1% at 18 months <sup>26</sup> –26% at 12 weeks <sup>72</sup>
Messages sent and received	Mean number of messages sent from user to coach/therapist/clinician	3	10.35 <sup>44</sup> –22 <sup>73</sup>
	Mean number of messages sent from coach/therapist/clinician to user	3	25.08 <sup>44</sup> –54 <sup>73</sup>
Posts and comments made	Mean number of newsfeed posts and/or comments	1	21.49 <sup>26</sup>
	Median number of newsfeed posts and/or comments	1	7 <sup>26</sup>
Participant self-report	User self-reported engagement	15	N/A
Miscellaneous	Non-comparable engagement metrics	11	N/A

**Table 4 | Correlation between engagement and intervention efficacy**

	Engagement	PD_TOTAL	PD_NEW_TOTAL	PD_TOTALExcTunneling	PD_PTS	PD_DS	PD_SCS
Engagement	1	–0.119	–0.267	–0.119	0.011	–0.191	–0.127
<i>p</i> -value		0.651	0.300	0.651	0.966	0.462	0.627
<i>N</i>	17	17	17	17	17	17	17

[–0.04, 05],  $p = 0.804$ . Similarly, we found no significant association between intervention efficacy and any of the four persuasive design domains (primary task:  $estimate = -0.08$ , 95% CI [–0.17, 01],  $p = 0.078$ ), dialogue:  $estimate = 0.06$ , 95% CI [–0.03, 14],  $p = 0.184$ ), system credibility:  $estimate = 0.001$ , 95% CI [–0.11, 11],  $p = -0.981$ ), social support:  $estimate = 0.05$ , 95% CI [–0.06, 16],  $p = 0.375$ )).

#### Relationship between persuasive design principles and engagement: correlation and meta-analysis findings

The intended meta-regression to explore the relationship between PSD principles and engagement could not be conducted due to substantial variation in how engagement was conceptualised, measured and reported across studies. Instead, a two-tailed Pearson's correlation analysis was performed to examine the relationship between the number of persuasive design principles used and app engagement. As there was substantial variation in measurements of engagement across studies, this correlation analysis focused on the most consistently reported engagement metric: percentage of users that completed the whole programme or percentage of

users that completed the programme per protocol ( $n = 17$ ). No significant relationship was found between the number of persuasive design principles used and programme completion,  $r(17) = 0.21$ ,  $p = 0.43$ .

A separate random-effects meta-analysis was conducted to test between-group differences (intervention vs control) at post-intervention for each individual PSD principle. The analysis revealed the significant effects in favour of digital interventions compared with control conditions for the majority of PSD principles (see Table 7). However, no significant effects were observed for the use of social role, surface credibility, real-world feel, third-party endorsement, normative influence and cooperation. Importantly, these latter outcomes were represented by only 3–9 studies each, whereas the others represented up to 81 studies, providing a more robust pooled effect size in most cases.

#### Discussion

This study found that mental health apps significantly improved clinical outcomes compared to control groups, with a medium effect size ( $N = 16,728$ ,  $g = 0.43$ ). Significantly positive effects were identified for

**Table 5 | Prevalence and examples of persuasive design principles used in digital mental health apps: percentage of studies reporting each principle**

Persuasive design principle	Studies % (n)	Example of use in app
Tunnelling	88% (81)	"This intervention has modules that cover a variety of topics and can be accessed in any order, including coping effectively with depression and anxiety, enhancing happiness, improving sleep quality, and practising mindfulness." <sup>74</sup>
Rehearsal	84% (77)	"In each web session, participants were encouraged to practice the prescribed strategies via several homework exercises. These homework exercises were presented in the app component of the intervention, which allowed users to practice these skills digitally and in their daily life." <sup>75</sup>
Trustworthiness	80% (74)	"The healthy minds programme (HMP) smartphone-based meditation intervention was constructed around the awareness, connection, insight, and purpose (ACIP) model of well-being proposed by Dahl and colleagues (2020)." <sup>76</sup>
Reminders	55% (51)	"As in the treatment condition, the app emits an alarm prompting participants to access the app at 3 pseudo-randomized time points per day, 6 days a week between 10.00 and 22.00 for 12 weeks alongside usual treatment." <sup>56</sup>
Personalisation	50% (46)	"Guided by the individual formulation, clinicians sent each client weekly tailored content suggestions." <sup>26</sup>
Self-monitoring	49% (45)	"Participants rated their anxiety before and after relaxation sessions using a visual analogue scale." <sup>77</sup>
Suggestion	48% (44)	"The main component of the app is a 30-day intervention involving one 5–10 min 'challenge' per day." <sup>46</sup>
Expertise	43% (40)	"Coaches were two bachelor's degree-level individuals who received training in the coaching manual and weekly supervision with a clinical psychologist." <sup>78</sup>
Reduction	32% (29)	"Besides written materials, all contents have also been audio-recorded to facilitate users to listen to the content if they are unable to read the materials on the go." <sup>79</sup>
Rewards	25% (23)	"The app includes gamification entailing an ad hoc character named TIVA—from the Spanish word "positIVA," meaning "positive." This character grows up and changes every time the caregiver performs the daily activity." <sup>50</sup>
Simulation	11% (10)	"In the "Lesson" phase, the core skills of the day are introduced via case examples in which the main character experiences the most commonly reported depressive and/or anxiety symptoms." <sup>81</sup>
Praise	10% (9)	"Following the completion of six levels pertaining to two beliefs, users may see an encouraging statement such as "Excellent! Now you've learned how to better deal with your thoughts and to better recognize the way you overestimate threat." <sup>43</sup>
Normative influence	10% (9)	"A peer-support community that is not moderated by professionals, but is rather a place where users can post their thoughts, challenges, and questions and receive support from others using the app." <sup>68</sup>
Social learning	9% (8)	"After challenge completion, users may choose to anonymously share a short written reflection about the challenge with another user. The app randomly selects another user and sends the written reflection." <sup>82</sup>
Social role	7% (6)	"Each skill is taught in a tutorial format by an avatar named "Dr. Joy", who introduces the player to the game, and explains the processes and skills involved." <sup>83</sup>
Third-party endorsements	7% (6)	"Calm is an award-winning mindfulness app that's available for iOS, Android, and desktop. Developed in 2012, Calm has over 100 million downloads and more than 700,000 five-star reviews." <sup>63</sup>
Social facilitation	5% (5)	"Individuals are also able to choose positive messages ("message in a bottle") to send out to other (anonymous) users." <sup>84</sup>
Authority	4% (4)	"The mobile programme complied with the United States Health Insurance Portability and Accountability Act (HIPAA)." <sup>44</sup>
Tailoring	3% (3)	Users input their current emotional distress and MoodMission provides a tailored list of five CBT-based activities, called "Missions." <sup>85</sup>
Cooperation	3% (3)	"Mothers can also reply and answer each other's questions." <sup>86</sup>
Surface-credibility	2% (2)	"Sleep stories narrated by celebrities like Matthew McConaughey and LeBron James." <sup>82</sup>
Social comparison	2% (2)	"Immediately after each response, a bar graph of the response rate of others is displayed, allowing the participant to see the answers most people chose (i.e., the correct answers)." <sup>87</sup>

**Table 6 | Prevalence and application of novel persuasive design principles in digital mental health apps**

Novel persuasive design principle	Studies % (n)	Definition	Example of use in app
Staged disclosure	25% (23)	Staged disclosure presents information gradually by restricting access to later content until prior tasks are completed, often incorporating gamification elements such as <i>unlocking levels</i> through mastery. Unlike tunnelling, which allows users to navigate sequential information freely, staged disclosure enforces a controlled sequence, reducing cognitive overload and enhancing engagement by encouraging task completion and sustained interaction.	"During the course of the programme, participants were prevented from changing the sequence order of the programme and were unable to proceed to the next session without completing the previous session." <sup>81</sup>
Goal setting	15% (13)	Encourages users to establish specific, measurable, and achievable objectives. This approach motivates users by providing clear targets to strive for, which can enhance their sense of purpose, commitment, and progress tracking.	"They then choose from a list of 29 skills any number they want to work on, define their current level of ability and set a goal for how well they want to be able to master a skill." <sup>82</sup>
Self-paced	9% (8)	Users proceed through content or tasks at their own pace and according to their preferences. This approach respects users' autonomy and control over their interaction with the system, promoting a more personalised and empowering experience.	"The app was designed to be self-guided. Participants were instructed to go at a pace that suited them." <sup>75</sup>
Limited use	7% (6)	Limiting the availability or access to certain features, content, or functions within an application or system. This approach is employed to create a sense of exclusivity, thereby increasing the perceived value and desirability of those elements.	"Following the completion of three levels each day, a screen prompting users to stop using the app for that day appears." <sup>43</sup>

**Fig. 4 |** This figure presents a heatmap illustrating the percentage of digital mental health apps within each mental health condition category that incorporate specific PSD principles. To reduce heterogeneity, only interventions targeting a single mental health condition with more than one intervention per condition were included, resulting in the analysis of 61 apps. Each cell shows the percentage of apps within the specific mental health category that incorporated a given PSD principle. Colour intensity reflects prevalence, with red tones indicating higher usage and green tones indicating lower usage.

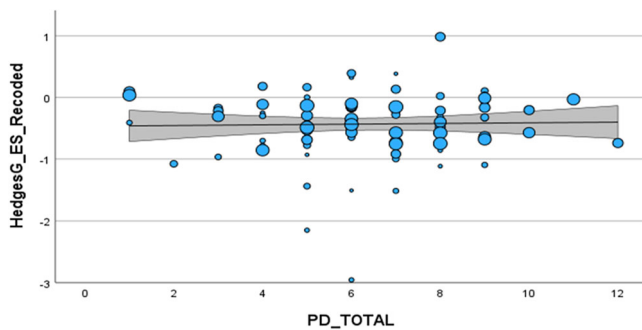
	Depression (n=20)	Disordered Eating (n=4)	Distress (n=4)	Stress (n=13)	Anxiety (n=9)	Postnatal Depression (n=4)	Psychotic Spectrum Disorders (n=3)	PTSD (n=2)	OCD (n=2)
Reduction	10	25	50	54	22	0	0	0	0
Tunnelling	75	75	75	92	89	100	100	100	50
Tailoring	5	0	0	8	0	0	67	0	0
Personalisation	70	0	50	50	44	25	67	100	0
Self-Monitoring	35	75	25	14	67	75	67	100	0
Simulation	10	0	0	0	22	25	33	0	0
Rehearsal	75	75	50	100	89	75	67	100	100
Praise	25	25	0	8	11	0	0	0	50
Rewards	30	25	25	23	22	25	0	0	50
Reminders	45	50	25	54	67	50	100	0	100
Suggestions	45	5	50	69	67	25	0	50	100
Similarity	0	0	0	0	0	0	0	0	0
Liking	0	0	0	0	0	0	0	0	0
Social Role	5	0	0	0	22	0	0	0	0
Trustworthiness	85	75	75	77	67	75	100	100	50
Expertise	40	25	25	38	56	100	67	0	50
Surface Credibility	0	0	0	0	0	0	0	0	0
Real-World Feel	0	0	25	23	0	0	0	0	0
Authority	5	0	0	23	11	0	0	0	0
Third-Party Endorsement	0	0	25	0	0	0	0	0	0
Verifiability	0	0	25	0	0	0	0	0	0
Social Learning	15	0	0	0	0	25	33	0	0
Social Comparison	5	0	0	0	0	0	0	0	0
Normative Influence	20	0	0	0	0	25	33	0	0
Social Facilitation	10	0	0	0	0	25	33	0	0
Cooperation	5	0	0	0	0	25	0	0	0
Competition	5	0	0	0	0	0	0	0	0
Recognition	0	0	0	0	0	0	0	0	0

interventions targeting depression, anxiety, stress, and body image/eating disorders, whereas no significant effects were observed for interventions addressing psychosis, suicide/self-harm, postnatal depression, or overall mental health. These results may reflect the much larger body of research focused on depression, anxiety, and stress, underscoring the need for further investigation into digital interventions for less-studied conditions. Notably, no significant differences in outcomes were identified on the basis of sample type (clinical, subthreshold, or non-clinical), the presence of human support, or the clinical approach (e.g., CBT, mindfulness, psychoeducation).

Notably, none of the studies included in the meta-analysis explicitly reported the use of persuasive design principles in their app descriptions. Through deductive coding, it was determined that 79% of the principles from the PSD framework were present across the apps included in the meta-analysis. We found that apps used between 1 and 12 persuasive design

principles, with a mode of 5. This aligns closely with the findings of McCall et al.’s<sup>25</sup> systematic review and meta-analysis, which identified between 1 and 13 principles per intervention, with a mean of 4.95. The most commonly implemented principles were tunnelling (88%), rehearsal (84%), trustworthiness (80%), reminders (55%), and personalisation (50%). Principles from the primary task support domain were the most frequently employed, a finding consistent with previous research<sup>25,30</sup>.

In contrast to previous research, however, this study did not find any significant association between the number or domain of persuasive design principles (e.g., primary task, dialogue) and app efficacy. Similarly, no significant relationship was observed between engagement and either individual persuasive principles or domains. These findings differ from earlier studies, which have reported both positive<sup>25,30,36,40</sup> and negative<sup>40</sup> associations between PSD principles and engagement or efficacy outcomes.



**Fig. 5 |** This figure illustrates the relationship between the total number of persuasive design principles implemented in digital mental health interventions (PD\_TOTAL) and intervention efficacy, measured by Hedges' *g* effect sizes. Each point represents an individual study, with point size reflecting study weight in the meta-regression. The shaded region represents the 95% confidence interval around the regression line. The analysis found no significant relationship between the number of persuasive design principles and intervention efficacy,  $b = 0.01$ ,  $SE = 0.02$ , 95% CI  $[-0.04, 0.05]$ ,  $p = 0.804$ .

These discrepancies may stem from several factors, including differences in study scope, coding practices, methodological approaches, and analytical techniques. For instance, some previous studies focused on digital interventions within the broader digital health field rather than specifically on mental health<sup>30,36</sup>, while others examined web-based interventions rather than smartphone apps<sup>30</sup> or exclusively investigated interventions targeting depression and anxiety<sup>25,40</sup>. Variations in scope, mental health focus, and intervention modalities may partly explain the divergent findings across studies.

Another factor could relate to the variability in coding practices across research teams. As noted, none of the included studies explicitly reported their use of persuasive design principles, necessitating reliance on subjective coding decisions made by individual research teams. This process introduces variability, as coding decisions are subject to subjective interpretations that may differ across research groups. These challenges highlight the need for more transparent and standardised reporting of persuasive principles to minimise bias and improve consistency and comparability across studies.

Methodological differences may also account for the observed discrepancies across reviews. Previous reviews typically coded interventions for PSD principles based solely on the descriptions provided in study outcome publications<sup>30,40</sup>. In contrast, our approach sought to enhance accuracy of PSD coding by incorporating supplementary materials, such as study protocols, development papers, and publicly available websites. This broader range of source material was intended to enable a more comprehensive assessment of the PSD principles operating within each intervention. However, this more expansive approach to coding may have contributed to the different findings relative to earlier reviews.

Finally, differences in study design and analytical approaches may have also contributed to the difference in findings. Our study utilised a methodology that focused exclusively on RCT-tested apps across all mental health conditions and a total participant sample size of 16,728, providing greater statistical power and the opportunity for subgroup analysis to enhance the reliability and precision of effect estimates.

Inconsistent reporting of PSD principles poses a significant challenge in evaluating their presence, implementation, and impact in digital mental health interventions. Without explicit disclosure from authors, it remains difficult to determine whether these principles have been incorporated, to what extent, or with what level of fidelity. This lack of transparency reflects a broader issue: the absence or inconsistent application of standardised frameworks for detailing persuasive features in digital interventions. Such gaps result in key design elements being either unreported or varying widely in their quality and integration. Addressing these challenges requires the adoption of standardised frameworks, such as the PSD framework,

alongside consensus on best practices for documenting and evaluating engagement-enhancing features. Clear and consistent reporting throughout the app development and evaluation process is essential to improve transparency, reduce variability, and ensure that the potential impacts of persuasive design in digital interventions can be rigorously assessed.

Building on these findings, we now turn to potential reasons why no significant relationship between PSD principles and engagement was identified in the current study. While coding practices and inconsistent reporting may have contributed to this outcome, several other factors may also explain the limited evidence for PSD principles enhancing engagement. These include variability in the application of PSD principles, challenges in reliably measuring engagement, and the influence of external variables that extend beyond the scope of the PSD framework.

As mentioned, inadequate reporting of how persuasive design principles are incorporated into interventions may have led to missed or misattribution of PSD principles to apps. This could also be a threshold issue, where a principle is minimally applied and marked as present, whereas another app may implement the same principle more thoroughly or effectively, yet both receive the same score. Second, substantial variability in how engagement is measurement across studies limits our ability to comprehensively explore the relationship between PSD principles and engagement. Third, using more PSD principles does not necessarily lead to proportionally higher engagement or efficacy. This non-linear relationship can make it harder to detect associations. For example, certain combinations of PSD principles may be more effective in fostering engagement and efficacy for specific subgroups of users. Fourth, other variables influencing engagement (e.g., overall app aesthetics and design, human support, user characteristics, social reciprocity, social ranking, variable unpredictable rewards, etc) may mask the relationship between the PSD principles researched in this study and engagement.

Finally, while it is possible that PSD principles do not directly influence engagement or efficacy, this explanation appears less plausible. Given the well-established role of user experience and design in shaping interactions with technology, it is improbable that persuasive design principles have no effect. Therefore, it is more reasonable to consider that other factors, such as variability in application or measurement, may account for the lack of a clear relationship in this study.

To shift focus from the relationship between PSD principles and engagement specifically, we now address the significant variability in how engagement was defined and reported across studies. This variability hindered the ability to test associations through meta-analysis and identify possible patterns of engagement. Notably, 24% of the included studies did not report engagement metrics at all, a considerable proportion given the importance of the engagement challenge in the digital intervention field. This aligns with previous findings by Lipschitz et al.<sup>2</sup>, who highlighted the low rates of engagement reporting in digital mental health interventions for depression, where only 64% of studies reported daily usage and just 23% provided retention rates for the final treatment week. The present study extends these concerns to mental health apps more broadly, suggesting that inconsistent and selective reporting of engagement metrics may skew perceived levels of user interaction. In some cases, this may reflect a form of selective outcome reporting, whereby studies present only the most favourable engagement metrics—those that cast the intervention in a positive light—while omitting others that may be less compelling. Such selective reporting can artificially inflate engagement data, potentially overestimating the use, relevance, and even efficacy of interventions. These discrepancies underscore the need for standardised engagement metrics to ensure accurate, comprehensive reporting and a clearer understanding of user interaction with digital mental health interventions.

We found significant variability in the engagement metrics reported across studies, identifying 25 distinct metrics, which we grouped into 10 categories. These categories included: (1) rate of uptake, (2) time spent on the app (min/h), (3) days of active use, (4) logins, (5) modules completed, (6) study metrics, (7) messages sent and received, (8) posts and comments made, (9) participant self-reports, and (10) miscellaneous. Among all the

**Table 7 | Random-effects meta-analysis of between-group differences (intervention vs control) at post-intervention, moderated by the use of individual persuasive design principles**

Effect size estimates								
Persuasive design principal	Moderator (yes/no)	k	Hedges G	SE	Z	p	95% CIs	
							Lower	Upper
Reduction	No (0)	63	−0.432	0.066	−6.512	<0.001	−0.562	−0.302
	Yes (1)	29	−0.444	0.065	−6.802	<0.001	−0.572	−0.316
Tunnelling	No (0)	11	−0.465	0.206	−2.264	0.024	−0.868	−0.062
	Yes (1)	81	−0.433	0.050	−8.738	<0.001	−0.530	−0.336
Tailoring	No (0)	89	−0.429	0.050	−8.589	<0.001	−0.526	−0.331
	Yes (1)	3	−0.551	0.248	−2.220	0.026	−1.038	−0.065
Personalisation	No (0)	46	−0.461	0.076	−6.043	<0.001	−0.610	−0.311
	Yes (1)	46	−0.407	0.061	−6.639	<0.001	−0.527	−0.287
Self-monitoring	No (0)	47	−0.362	0.065	−5.574	<0.001	−0.490	−0.235
	Yes (1)	45	−0.504	0.072	−6.980	<0.001	−0.645	−0.362
Simulation	No (0)	82	−0.398	0.046	−8.574	<0.001	−0.489	−0.307
	Yes (1)	10	−0.760	0.255	−2.984	0.003	−1.260	−0.261
Rehearsal	No (0)	15	−0.243	0.083	−2.943	0.003	−0.406	−0.081
	Yes (1)	77	−0.467	0.056	−8.406	<0.001	−0.576	−0.358
Praise	No (0)	83	−0.438	0.053	−8.236	<0.001	−0.542	−0.334
	Yes (1)	9	−0.351	0.095	−3.707	<0.001	−0.536	−0.165
Rewards	No (0)	69	−0.471	0.056	−8.350	<0.001	−0.582	−0.360
	Yes (1)	23	−0.303	0.094	−3.241	0.001	−0.486	−0.120
Reminders	No (0)	41	−0.435	0.078	−5.580	<0.001	−0.588	−0.282
	Yes (1)	51	−0.432	0.063	−6.866	<0.001	−0.556	−0.309
Suggestion	No (0)	48	−0.483	0.081	−5.998	<0.001	−0.640	−0.325
	Yes (1)	44	−0.383	0.058	−6.655	<0.001	−0.496	−0.270
Similarity	No (0)	92	N/A	N/A	N/A	N/A	N/A	N/A
	Yes (1)	0	N/A	N/A	N/A	N/A	N/A	N/A
Liking	No (0)	92	N/A	N/A	N/A	N/A	N/A	N/A
	Yes (1)	0	N/A	N/A	N/A	N/A	N/A	N/A
Social role	No (0)	86	−0.452	0.048	−9.334	<0.001	−0.547	−0.357
	Yes (1)	6	−0.143	0.245	−0.582	0.561	−0.624	0.338
Trustworthiness	No (0)	18	−0.263	0.115	−2.285	0.022	−0.489	−0.037
	Yes (1)	74	−0.473	0.052	−9.029	<0.001	−0.576	−0.370
Expertise	No (0)	52	−0.435	0.050	−8.702	<0.001	−0.532	−0.337
	Yes (1)	40	−0.428	0.094	−4.537	<0.001	−0.613	−0.243
Surface credibility	No (0)	90	−0.436	0.050	−8.807	<0.001	−0.533	−0.339
	Yes (1)	2	−0.271	0.365	−0.742	0.458	−0.985	0.444
Real-world feel	No (0)	90	−0.436	0.050	−8.807	<0.001	−0.533	−0.339
	Yes (1)	2	−0.271	0.365	−0.742	0.458	−0.985	0.444
Authority	No (0)	88	−0.444	0.051	−8.700	<0.001	−0.544	−0.344
	Yes (1)	4	−0.209	0.100	−2.086	0.037	−0.405	−0.013
Third-party endorsement	No (0)	86	−0.446	0.051	−8.731	<0.001	−0.546	−0.346
	Yes (1)	6	−0.237	0.145	−1.633	0.102	−0.522	0.047
Verifiability	No (0)	92	N/A	N/A	N/A	N/A	N/A	N/A
	Yes (1)	0	N/A	N/A	N/A	N/A	N/A	N/A
Social learning	No (0)	84	−0.437	0.053	−8.222	<0.001	−0.541	−0.333
	Yes (1)	8	−0.381	0.102	−3.731	<0.001	−0.581	−0.181
Social comparison	No (0)	90	−0.425	0.050	−8.589	<0.001	−0.522	−0.328
	Yes (1)	2	−0.764	0.158	−4.840	<0.001	−1.074	−0.455
Normative Influence	No (0)	83	−0.462	0.048	−9.562	<0.001	−0.557	−0.367
	Yes (1)	9	−0.159	0.190	−0.836	0.403	−0.532	0.214

**Table 7 (continued) | Random-effects meta-analysis of between-group differences (intervention vs control) at post-intervention, moderated by the use of individual persuasive design principles**

Effect size estimates								
Persuasive design principal	Moderator (yes/no)	k	Hedges G	SE	Z	p	95% CIs	
Social facilitation	No (0)	87	−0.439	0.052	−8.510	<0.001	−0.540	−0.338
	Yes (1)	5	−0.344	0.117	−2.940	0.003	−0.574	−0.115
Cooperation	No (0)	89	−0.440	0.049	−8.921	<0.001	−0.537	−0.344
	Yes (1)	3	−0.197	0.325	−0.607	0.544	−0.834	0.440
Competition	No (0)	92	N/A	N/A	N/A	N/A	N/A	N/A
	Yes (1)	0	N/A	N/A	N/A	N/A	N/A	N/A
Recognition	No (0)	92	N/A	N/A	N/A	N/A	N/A	N/A
	Yes (1)	0	N/A	N/A	N/A	N/A	N/A	N/A

engagement metrics, the most commonly reported were: (1) the percentage of users who completed the entire intervention or who completed the intervention per protocol (14% of studies), (2) self-reported engagement (12% of studies), and (3) the mean percentage of modules completed (11% of studies). To support comparability across studies, we recommend using either *the percentage of users who completed the entire intervention or completed the intervention per protocol*, and/or *the mean percentage of modules completed*. Given its susceptibility to recall bias and other potential sources of error, we advise interpreting self-report data with caution and avoided when objective metrics are available.

The focus on engagement metrics in digital interventions stems from the understanding that engagement is necessary for the intervention to be effective. Engagement metrics are, therefore, crucial for ensuring that users are meaningfully exposed to the intervention content and essential for attributing positive outcomes to the intervention in question<sup>3</sup>. However, to obtain an accurate understanding of the relationship between engagement and effectiveness, we must also understand the behaviour change mechanisms of the intervention, that is, how the intervention drives real-world change. The factors influencing engagement with the digital intervention, such as persuasive design, do not fully align with those driving real-world behaviour change, which are often rooted in deeper motivations, self-efficacy, and the perceived relevance and feasibility of recommended behaviours<sup>49</sup>. Despite the aim of promoting behaviour change, Orji and Moffat<sup>36</sup> reported that 55% of studies did not link mental health interventions to any behaviour change theory. This absence of a theoretical underpinning was also observed in the present study, where the extraction of behaviour change theories underlying the reviewed studies was halted due to insufficient reporting.

Therefore, to advance the field, we propose several recommendations aimed at addressing gaps in the development, reporting, and evaluation of digital mental health interventions. Establishing standardised engagement metrics is crucial for enabling reliable comparisons across studies and facilitating future meta-analyses. We recommend adopting guidelines such as CONSORT-EHEALTH<sup>50</sup> to improve reporting consistency and to provide clearer insights into how persuasive design principles influence engagement and outcomes. We also endorse, and propose extending, the five-point checklist developed by Lipschitz et al.<sup>2</sup>, which includes metrics such as adherence, rate of uptake, level-of-use, duration-of-use, and number of completers. We suggest incorporating two additional measures aimed at providing a more comprehensive understanding of intervention impact. First, the description of interventions should include detailed information on the design principles and features of the platform that promote engagement; authors should explicitly state the frameworks (e.g., PSD) and specific principles employed. Second, authors should explicitly link the intervention to an overarching model of behaviour change. By systematically reporting these principles alongside consistent engagement metrics, researchers can better establish possible links between design elements, user engagement outcomes, real-world behaviour change, theoretical models,

and clinical outcomes. This approach will facilitate a deeper understanding of user engagement and what drives therapeutic change.

To build on these recommendations, it is also important to acknowledge that while persuasive design principles aim to enhance engagement, therapeutic change requires a nuanced understanding of how engagement interacts with psychological mechanisms specific to mental health conditions. While persuasive principles aim to enhance user engagement, it is important to recognise that engagement alone may not directly lead to behaviour change unless paired with mechanisms targeting the underlying psychological processes relevant to a mental health condition. For instance, the negative association Wu et al.<sup>40</sup> found between PSD features and engagement (as measured by completion rate), points to the complexity of this relationship and suggests that other factors may moderate the impact of persuasive techniques on outcomes. For instance, Wu et al.<sup>40</sup> speculate that engagement may not be the primary factor driving the association with effectiveness. We recognise that behaviour change in digital mental health intervention is likely influenced by a complex interplay of factors, including both engagement-driven processes and condition-specific mechanisms of action.

Despite its strengths, this study had several notable limitations. First, the use of persuasive principles was not explicitly reported in any studies included in the analysis. Although coding was guided by a well-defined framework and conducted by an expert team of user experience designers, many of the reviewed apps were not commercially available. As a result, limiting their comprehensive description and accessibility for analysis. As a result, we were unable to directly access or interact with these interventions and had to rely solely on the authors' descriptions in published papers to code for persuasive system design features. This method raises concerns that some apps may have employed persuasive principles not described in their app descriptions, potentially affecting assessments of effectiveness and outcomes. Additionally, there is uncertainty about the extent to which persuasive principles were actually incorporated, even among those apps that were coded for it. This limitation highlights the urgent need for greater transparency in intervention reporting, especially for apps that are not commercially available, to facilitate more detailed evaluations of how persuasive design principles impact mental health app outcomes.

Additionally, we could only conduct a single correlation analysis between PSD principles and one engagement metric. With only 76% of studies reporting on engagement or adherence, and those metrics being highly heterogeneous, a meta-analysis was not feasible. This inconsistency in reporting and operationalising engagement metrics significantly hinders progress in understanding app and digital intervention engagement. Consequently, these limitations constrained our ability to comprehensively evaluate the relationships between persuasive design principles and app engagement and effectiveness. Despite these limitations, to our knowledge, this is the most extensive examination to date on the relationships between PSD principles, engagement and efficacy in digital mental health. We included 119 RCTs, coded all persuasive design principles with high inter-

rater reliability, reviewed all protocols and associated reports, including downloading the apps when available, and categorised all engagement data reported into meaningful categories.

In conclusion, this study provides the most extensive and systematic evaluation to date of the relationship between persuasive design principles, engagement, and efficacy in digital mental health apps. Results identified that mental health apps are moderately effective for depression, anxiety, stress, and body image/eating disorders but show limited efficacy for psychosis, suicide/self-harm, and postnatal depression, highlighting the need for further research in these areas. While persuasive design principles are intended to enhance user engagement, our study found limited evidence of their direct impact, likely due to inconsistent reporting, variability in their application, and a lack of standardised engagement metrics. Further, owing to the rapid development of technology and its capabilities, persuasive principle frameworks require revision and updating to include strategies such as staged disclosure as identified in this study. These findings underscore critical limitations in the field that require immediate attention. To address the ongoing issue of poor engagement with digital mental health interventions, we need a concerted effort to establish a uniform understanding of engagement as a foundation for harmonising the reporting of engagement metrics. Additionally, it is essential to align descriptions of the persuasive strategies and behaviour change frameworks used to underpin engagement. Addressing these methodological gaps is critical for enabling large-scale data pooling, which will facilitate a more nuanced understanding of the factors influencing engagement, tailored to specific populations and contexts. Without such foundational improvements, the key challenge of engagement in digital mental health will remain intractable, limiting its real-world impact.

## Methods

This study comprised a systematic review, meta-analyses, and meta-regression. Details on the conduct of the study in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020)<sup>51</sup> guidelines are provided in the supplementary information. Additionally, it adhered to the prospectively registered protocol in PROSPERO (CRD42022352123) in August 2022, which can be accessed at <https://www.crd.york.ac.uk/prospero/>.

## Eligibility criteria

To be included, studies had to meet the following criteria: (1) were randomised controlled trials; (2) delivered a digital mental health intervention; (3) were delivered via a smartphone app; (4) aimed at addressing a mental health condition with psychological/psychosocial approaches; (5) delivered the therapeutic intervention primarily via the app; (6) reported quantifiable measurements of user engagement, intervention adherence, or effectiveness data; (7) contained an app description that was adequate enough to allow coding of its persuasive design principles; (8) were published in peer-reviewed academic journals; and (9) were written in English.

The exclusion criteria were as follows: (1) case studies and feasibility studies without a control group, reviews, theses, or book chapters and (2) data reported in another study (i.e., conference abstracts where data are subsequently published elsewhere).

## Synthesis approach

To examine our primary aims, we conducted a systematic narrative synthesis and meta-analysis of the available literature. The available studies that were identified as meeting criteria via review were systematically synthesised and reported by examining (a) the types of samples, interventions, and mental health contexts in which smartphone-based therapeutic interventions have been applied; (b) which persuasive principles were present in each app description and the number of overall principles present in each app; (c) the type of metrics used to report on intervention engagement and level of engagement (if available); and (d) pre and post-intervention outcome data.

## Search strategy

A systematic search was conducted via the Ovid platform to identify relevant studies. The databases searched included PsycINFO, PubMed, and Embase, as they provide a well-established body of research on engagement and clinical outcomes in the clinical context, which was the focus of our review. A search term strategy was designed to target four key concepts: (1) *app* (app OR smartphone OR iphone OR android OR mobile OR “mobile phone” OR “mobile app” OR “mobile application” OR “mobile based” OR “smartphone app” OR “smartphone application”), (2) *digital intervention* (digital intervention OR “mobile intervention” OR “smartphone intervention” OR “digital mental health” OR “mobile health intervention” OR “digital technology” OR “mHealth” OR “eHealth” AND “mobile health”), (3) *mental ill health* (“mental ill health” OR disorder OR “psychiatric disorder” OR well-being OR wellbeing OR depress\* OR psycho\* OR bipolar OR anxiety OR schizophrenia OR affective OR self-harm OR self-injury OR distress OR mood OR body image OR eating disorder OR suicid\* OR “posttraumatic stress” OR PTSD OR agoraphobia OR phobia\* OR panic OR funct\* OR OCD OR stress OR symptom,\* and (4) *randomised controlled trial* (randomised OR randomized OR RCT OR waitlist OR allocate\*). NOT protocol OR “single arm” OR “systematic review” OR “scoping review.”

## Study selection and data collection

The searches were completed in August 2022, after which all retrieved articles identified by the search strategy were downloaded and then uploaded to Covidence systematic review software, Veritas Health Innovation, Melbourne, Australia. Available at [www.covidence.org](http://www.covidence.org). After automated identification and removal of duplicates, two reviewers (L.V., J.D.X.H.) independently screened the titles and abstracts of all articles according to the eligibility criteria. Once consensus on eligibility was reached in the screening phase, two researchers (L.V., J.D.X.H.) independently conducted full-text reviews. If conflicts arose in the screening or full-text review phases via the Covidence interrater feature, the reviewers (L.V., J.D.X.H.) resolved them through discussion.

## Data extraction

Four authors (L.V., M.S.V., S.O., and R.S.) independently extracted data related to the study, sample, intervention characteristics and engagement and efficacy rates (i-v. below). Fifty percent of the full dataset was double-extracted to ensure reliability (W.P., R.P.S., and L.V.). Any discrepancies were resolved by consensus with a third author (J.N.). The following data were extracted:

- (i) Study characteristics: authors, year of publication, country of participant recruitment,
- (ii) Sample characteristics: total sample size, number of participants in the control and intervention group(s), primary mental health condition, age, gender, population type (clinical/sub-clinical/non-clinical),
- (iii) Intervention characteristics: app name, description, theoretical treatment model (e.g., CBT), intervention length (in days), if there was human support, control conditions (e.g., waitlist, treatment as usual, active control), primary mental health target including discrete symptoms (i.e., stress), symptom clusters (i.e., a depression diagnoses) and measures of behaviour (i.e., social functioning);
- (iv) Engagement rates: whether engagement data were reported (yes/no), engagement/adherence metric (e.g., number of logins, modules completed), and engagement rates;
- (v) Efficacy data: means, standard deviations (SDs), and sample sizes (*N*) of the primary outcome measures for both the intervention and control group(s) at the pre and post-intervention time points. If certain metrics (e.g., SDs) were missing, they were either calculated using available data or the authors were contacted for additional information. When multiple primary outcome measures were presented, the first listed measure was considered the primary outcome for extraction. An exception to this procedure was made for subsequent outcome

measures pertaining to depression or anxiety, which were also included in the extraction process, and

(vi) Persuasive principles: the presence or absence of the 28 persuasive design principles outlined in the Persuasive Systems Design (PSD) framework for each app description. Each principle was coded dichotomously for each app, with a value of 1 indicating its presence and 0 indicating its absence, resulting in a total of 28. Data were extracted from descriptions of apps within the included studies, and, whenever possible, app information from supplementary materials such as protocols, development papers, app stores, or websites was gathered and used to provide more context for each app.

These data were extracted by three authors (K.B., L.B., and L.V.), who first collaborated to operationalise the PSD framework (Table 1.) and apply its principles to contemporary app design. KB, a senior UX designer, and L.B., a senior product designer, are both experts in product design and user needs. Initially, the three authors coded 20% of the apps collaboratively to ensure consistency in their approach before coding independently once they were confident that they were aligned in their deductive approach. This collaborative effort aimed to maintain accuracy and consistency in applying the PSD framework to the app descriptions. In cases of discrepancies requiring resolution, a fourth author (J.N.) was consulted. During this coding process, reoccurring principles that were not part of the PSD framework were identified and discussed among the reviewers. These principles were then added to the deductive coding framework for later analysis.

### Risk of bias assessment of individual studies

Six reviewers independently evaluated the risk of bias in each study via the Cochrane risk of bias (RoB2) tool. Each of the five risk domains of the RoB2 was scored against a three-point rating scale relating to a low, some concern, or high risk of bias.

### Meta-analysis, meta-regression, and correlation analysis

A meta-regression analysis was conducted to investigate the relationship between intervention efficacy and the implementation of persuasive design principles. Additionally, this analysis examined the associations between intervention efficacy and variables related to sample type (e.g., clinical, sub-threshold, non-clinical), outcome type (e.g., depression, anxiety etc.), the presence of human support (blended vs self-guided), and intervention type (e.g., CBT, mindfulness etc.). A two-tailed Pearson's correlation analysis was also performed to examine the relationship between the number of persuasive design principles and app engagement. All analyses were conducted in SPSS (v29).

The meta-analysis and meta-regression used random-effects models with maximum likelihood estimations. Initially, we computed Hedges'  $g$  effect sizes from available estimates (e.g.,  $M$ ,  $SD$ ) reported within studies to reflect the size of the between-group (control vs intervention) effects at the post-intervention time-point, and then meta-analysed them. The primary outcome variable (e.g., depression) as reported within each study was used across all analyses. To maintain coding consistency, estimates calculated on the basis of positive-valued outcomes (i.e., where higher scores indicate better mental health) were reverse-coded to align with the majority of outcomes. Therefore, negative Hedges'  $g$  values represent cases where the intervention group was more effective than the control group. To gain a deeper understanding of intervention efficacy, we conducted sub-group moderator analyses across various variables, including sample type (non-clinical, sub-clinical, clinical), outcome type (depression, anxiety, stress, social anxiety, PTSD, body image/eating disorder, psychosis, suicide/self-harm, general mental health, postnatal depression), presence of human support (blended vs self-guided), and type of intervention (CBT, CBT+, mindfulness-based, and psychoeducation). This approach allowed us to identify specific factors that may influence the overall effectiveness of digital mental health interventions.

Additionally, a meta-regression was performed, regressing the efficacy of the interventions (Hedges'  $g$ ) on the total number of persuasive design

principles identified in each study, to determine the potential relationship between persuasive design principles and intervention efficacy. Throughout all analyses, we assessed heterogeneity using common metrics (e.g., Cochran's  $Q$ ,  $I^2$ <sup>52,53</sup>); and examined potential publication bias by (a) visually inspecting the funnel plot of standard errors by Hedges'  $g$  to assess for asymmetry, and (b) examining trim and fill estimates to determine if the potential for missing estimates affected the overall effect size<sup>54</sup>.

A two-tailed Pearson's correlation analysis was run to examine the relationship between the number of persuasive design principles used and app engagement, as the literature presents conflicting findings regarding the relationship between persuasive design principles and user engagement levels.

### Data availability

The datasets analysed during this review are available from the corresponding author upon request.

### Code availability

The code used for the analyses in this review is available from the corresponding author upon request.

Received: 21 October 2024; Accepted: 14 March 2025;

Published online: 29 April 2025

### References

1. Lattie, E. G. et al. Digital mental health interventions for depression, anxiety and enhancement of psychological well-being among college students: systematic review. *J. Med. Internet Res.* **21**, e12869 (2019).
2. 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*. <https://doi.org/10.1007/s40501-023-00297-3> (2023).
3. Lipschitz, J. M. et al. Digital mental health interventions for depression: scoping review of user engagement. *J. Med. Internet Res.* <https://doi.org/10.2196/39204> (2022).
4. Welles, C. C. et al. Sources of clinician burnout in providing care for underserved patients in a safety-net healthcare system. *J. Gen. Intern. Med.* **38**, 1468–1475 (2023).
5. Corcadden, L., Callander, E. J. & Topp, S. M. Who experiences unmet need for mental health services and what other barriers to accessing health care do they face? Findings from Australia and Canada. *Int. J. Health Plan. Manag.* **34**, 761–772 (2019).
6. Mulraney, M et al. How long and how much? Wait times and costs for initial private child mental health appointments. *J. Paediatr. Child Health* **57**, 526–532 (2021).
7. Subotic-Kerry, M. et al. While they wait: a cross-sectional survey on wait times for mental health treatment for anxiety and depression for adolescents in Australia. *BMJ open* **15**, e087342 (2025).
8. Torous, J., Myrick, K.J., Rauseo-Ricupero, N. & Firth, J. Digital mental health and COVID-19: using technology today to accelerate the curve on access and quality tomorrow. *JMIR mental health* **7**, e18848 (2020).
9. Schueller, S. M., Muñoz, R. F. & Mohr, D. C. Realizing the potential of behavioral intervention technologies. *Curr. Dir. Psychol. Sci.* **22**, 478–483 (2013).
10. Bond, R. R. et al. Digital transformation of mental health services. *npj Ment. Health Res.* **2**, 13 (2023).
11. Engel, L. et al. The cost-effectiveness of a novel online social therapy to maintain treatment effects from first-episode psychosis services: results from the horizons randomized controlled trial. *Schizophr. Bull.* **50**, 427–436 (2024).
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. Saad, A., Bruno, D., Camara, B., D'Agostino, J. & Bolea-Alamanac, B. Self-directed technology-based therapeutic methods for adult patients receiving mental health services: systematic review. *JMIR Mental Health* **8**, e27404 (2021).
14. Gega, L. et al. Digital interventions in mental health: evidence syntheses and economic modelling. *Health Technol. Assess.* **26**, 1–182 (2022).
15. Gentili, A. et al. The cost-effectiveness of digital health interventions: a systematic review of the literature. *Front Public Health.* **10**, 787135 (2022).
16. Eichstedt, J. A. et al. Waitlist management in child and adolescent mental health care: a scoping review. *Child Youth Serv. Rev.* **160**, 107529 (2024).
17. Cross, S. et al. Developing a theory of change for a digital youth mental health service (moderated online social therapy): mixed methods knowledge synthesis study. *JMIR Form. Res.* **7**, e49846 (2023).
18. Erbe, D., Eichert, H. C., Riper, H. & Ebert, D. D. Blending face-to-face and internet-based interventions for the treatment of mental disorders in adults: systematic review. *J. Med. Internet Res.* **19**, e306 (2017).
19. O'Sullivan, S. et al. A novel blended transdiagnostic intervention (eOrygen) for youth psychosis and borderline personality disorder: uncontrolled single-group pilot study. *JMIR Ment. Health* **11**, e49217 (2024).
20. Valentine, L. et al. Blended digital and face-to-face care for first-episode psychosis treatment in young people: qualitative study. *JMIR Ment. Health* **7**, e18990, <https://doi.org/10.2196/18990> (2020).
21. Alvarez-Jimenez, M. et al. A national evaluation of a multi-modal, blended, digital intervention integrated within Australian youth mental health services. *Acta Psychiatr. Scand.* **151**, 317–331, <https://doi.org/10.1111/acps.13751> (2025).
22. Valentine, L., D'Alfonso, S. & Lederman, R. Recommender systems for mental health apps: advantages and ethical challenges. *AI Soc.* **38**, 1627–1638 (2023).
23. Bell, I. H. et al. The impact of COVID-19 on youth mental health: a mixed methods survey. *Psychiatry Res.* **321**, 115082 (2023).
24. Gan, D. Z. Q., 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.* <https://doi.org/10.3389/fdgth.2021.764079> (2021).
25. McCall, H. C., Hadjistavropoulos, H. D. & Sundström, C. R. F. Exploring the role of persuasive design in unguided internet-delivered cognitive behavioral therapy for depression and anxiety among adults: systematic review, meta-analysis, and meta-regression. *J. Med. Internet Res.* <https://doi.org/10.2196/26939> (2021).
26. Alvarez-Jimenez, M. et al. The Horyzons project: a randomized controlled trial of a novel online social therapy to maintain treatment effects from specialist first-episode psychosis services. *World Psychiatry* **20**, 233–243 (2021).
27. Jankovic, J., Richards, F. & Priebe, S. Advance statements in adult mental health. *Adv. Psych. Treat.* **16**, 448–55 (2010).
28. de Beurs, D., van Bruinessen, I., Noordman, J., Friele, R. & van Dulmen, S. Active involvement of end users when developing web-based mental health interventions. *Front. Psychiatry* **8**, 72 (2017).
29. Fleming, T. et al. Beyond the trial: systematic review of real-world uptake and engagement with digital self-help interventions for depression, low mood, or anxiety. *J. Med. Internet Res.* **20**, e199, <https://doi.org/10.2196/jmir.9275> (2018).
30. Kelders, S. M., Kok, R. N., Ossebaard, H. C. & Van Gemert-Pijnen, J. E. W. C. Persuasive system design does matter: A systematic review of adherence to web-based interventions. *J. Med. Internet Res.* <https://doi.org/10.2196/jmir.2104> (2012).
31. Baumel, A., Muench, F., Edan, S. & Kane, J. M. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. *J. Med. Internet Res.* **21**, e14567 (2019).
32. Baumel, A. Therapeutic activities as a link between program usage and clinical outcomes in digital mental health interventions: a proposed research framework. *J. Technol. Behav. Sci.* <https://doi.org/10.1007/s41347-022-00245-7> (2022).
33. Cole-Lewis, H., Ezeanochie, N. & Turgiss, J. Understanding health behavior technology engagement: pathway to measuring digital behavior change interventions. *JMIR Form. Res.* <https://doi.org/10.2196/14052> (2019).
34. Borghouts, J. et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. *J. Med. Internet Res.* <https://doi.org/10.2196/24387> (2021).
35. Fogg, B. J. Persuasive computers: perspectives and research directions. in *Conference on Human Factors in Computing Systems—Proceedings* 225–232 (Stanford University, 1998).
36. Orji, R. & Moffatt, K. Persuasive technology for health and wellness: state-of-the-art and emerging trends. *Health Inform. J.* **24**, 66–91 (2018).
37. Fogg, B. J., Cuellar, G. & Danielson, D. Motivating, influencing, and persuading users: an introduction to captology. in *Human–Computer Interaction Fundamentals* 133–146 (2009). <https://doi.org/10.1201/b10368-13>
38. Oinas-Kukkonen, H. & Harjumaa, M. Persuasive systems design: key issues, process model, and system features. *Commun. Assoc. Inf. Syst.* <https://doi.org/10.17705/1CAIS.02428> (2009).
39. Parada, F., Martínez, V., Espinosa, H. D., Bauer, S. & Moessner, M. Using persuasive systems design model to evaluate 'cuida tu Ánimo': an internet-based pilot program for prevention and early intervention of adolescent depression. *Telemed. e-Health* **26**, 251–254 (2020).
40. Wu, A. et al. Smartphone apps for depression and anxiety: a systematic review and meta-analysis of techniques to increase engagement. *NPJ Digit Med.* <https://doi.org/10.1038/s41746-021-00386-8> (2021).
41. Higgins, J. P. T., Savović, J., Page, M. J., Elbers, R. G. & Sterne, J. A. C. Assessing risk of bias in a randomized trial. Draft version (29 January 2019). *Cochrane Handbook for Systematic Reviews of Interventions* (Wiley, 2019).
42. Sun, Y. et al. Effectiveness of smartphone-based mindfulness training on maternal perinatal depression: randomized controlled trial. *J. Med. Internet Res.* **23**, e23410 (2021).
43. Cerea, S. et al. Cognitive behavioral training using a mobile application reduces body image-related symptoms in high-risk female university students: a randomized controlled study. *Behav. Ther.* **52**, 170–182 (2021).
44. Newman, M. G., Jacobson, N. C., Rackoff, G. N., Bell, M. J. & Taylor, C. B. A randomized controlled trial of a smartphone-based application for the treatment of anxiety. *Psychother. Res.* **31**, 443–454, <https://doi.org/10.1080/10503307.2020.1790688> (2021).
45. Torok, M. et al. The effect of a therapeutic smartphone application on suicidal ideation in young adults: findings from a randomized controlled trial in Australia. *PLoS Med.* **19**, e1003978 (2022).
46. Deady, M. et al. Preventing depression using a smartphone app: a randomized controlled trial. *Psychol. Med.* **52**, 457–466 (2022).
47. Deci, E. L. & Ryan, R. M. Self-determination theory. in *International Encyclopedia of the Social & Behavioral Sciences: Second Edition* (Scientific Research, 2015); <https://doi.org/10.1016/B978-0-08-097086-8.26036-4>.
48. Locke, E. A. & Latham, G. P. Building a practically useful theory of goal setting and task motivation. A 35-year odyssey. *Am. Psychol.* **57**, 705–717 (2002).
49. Michie, S., van Stralen, M. M. & West, R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement. Sci* **6**, 42 (2011).
50. Eysenbach, G. et al. ConSORT-eHealth: Improving and standardizing evaluation reports of web-based and mobile health interventions. *J. Med Internet Res.* **13**, e126 (2011).

51. Page, M. J. et al. Updating guidance for reporting systematic reviews: development of the PRISMA 2020 statement. *J. Clin. Epidemiol.* **134**, 103–112 (2021).
52. Cochran, W. G. The combination of estimates from different experiments. *Biometrics* **10**, 101 (1954).
53. Higgins, J. P. T., Thompson, S. G., Deeks, J. J. & Altman, D. G. Measuring inconsistency in meta-analyses. *BMJ*. <https://doi.org/10.1136/bmj.327.7414.557> (2003).
54. Duval, S. & Tweedie, R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* **56**, 455–463 (2000).
55. Lukas, C. A., Eskofier, B. & Berking, M. A gamified smartphone-based intervention for depression: randomized controlled pilot trial. *JMIR Ment. Health* **8**, e16643 (2021).
56. Carl, J. R. et al. Efficacy of digital cognitive behavioral therapy for moderate-to-severe symptoms of generalized anxiety disorder: a randomized controlled trial. *Depress Anxiety* **37**, 1168–1178 (2020).
57. Flett, J. A. M., Conner, T. S., Riordan, B. C., Patterson, T. & Hayne, H. App-based mindfulness meditation for psychological distress and adjustment to college in incoming university students: a pragmatic, randomised, waitlist-controlled trial. *Psychol. Health* **35**, 1049–1074 (2020).
58. Stolz, T. et al. A mobile App for social anxiety disorder: a three-arm randomized controlled trial comparing mobile and PC-based guided self-help interventions. *J. Consult Clin. Psychol.* **86**, 493–504 (2018).
59. Steare, T. et al. Smartphone-delivered self-management for first-episode psychosis: The ARIES feasibility randomised controlled trial. *BMJ Open* **10**, e034927 (2020).
60. Mantani, A. et al. Smartphone cognitive behavioral therapy as an adjunct to pharmacotherapy for refractory depression: randomized controlled trial. *J. Med. Internet Res.* **19**, e373 (2017).
61. Catuara-Solarz, S. et al. The efficacy of “foundations,” a digital mental health app to improve mental well-being during COVID-19: proof-of-principle randomized controlled trial. *JMIR Mhealth Uhealth* **10**, e30976 (2022).
62. Huberty, J. et al. Efficacy of the mindfulness meditation mobile app “calm” to reduce stress among college students: randomized controlled trial. *JMIR Mhealth Uhealth* **7**, e14273 (2019).
63. Laird, B., Puzia, M., Larkey, L., Ehlers, D. & Huberty, J. A mobile app for stress management in middle-aged men and women (Calm): feasibility randomized controlled trial. *JMIR Form. Res.* **6**, e30294 (2022).
64. Ebert, D. D. et al. Effectiveness of web- and mobile-based treatment of subthreshold depression with adherence-focused guidance: a single-blind randomized controlled trial. *Behav. Ther.* **49**, 71–83 (2018).
65. Heim, E. et al. Step-by-step: feasibility randomised controlled trial of a mobile-based intervention for depression among populations affected by adversity in Lebanon. *Internet Inter.* **24**, 100380 (2021).
66. Bucci, S. et al. Actissist: proof-of-concept trial of a theory-driven digital intervention for psychosis. *Schizophr. Bull.* **44**, 1070–1080, <https://doi.org/10.1093/schbul/sby032> (2018).
67. Roepke, AM et al. Randomized controlled trial of superbetter, a smartphone-based/internet-based self-help tool to reduce depressive symptoms. *Games Health J.* **4**, 235–246 (2015).
68. Moberg, C., Niles, A & Beermann, D Guided self-help works: randomized waitlist controlled trial of Pacifica, a mobile app integrating cognitive behavioral therapy and mindfulness for stress, anxiety, and depression. *J. Med. Internet Res.* **21**, e12556 (2019).
69. Ebert, DD et al. Effectiveness and moderators of an internet-based mobile-supported stress management intervention as a universal prevention approach: randomized controlled trial. *J. Med. Internet Res.* **23**, e22107 (2021).
70. Depp, CA, Perivoliotis, D, Holden, J, Dorr, J & Granholm, EL Single-session mobile-augmented intervention in serious mental illness: a three-arm randomized controlled trial. *Schizophr. Bull.* **45**, 752–762 (2019).
71. van Aubele, 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).
72. Ben-Zeev, D et al. Mobile health (mHealth) versus clinic-based group intervention for people with serious mental illness: a randomized controlled trial. *Psychiatr. Serv.* **69**, 978–985 (2018).
73. Raevuori, A et al. A therapist-guided smartphone app for major depression in young adults: a randomized clinical trial. *J. Affect Disord.* **286**, 228–238 (2021).
74. De Kock, JH et al. Brief digital interventions to support the psychological well-being of NHS staff during the COVID-19 pandemic: 3-arm pilot randomized controlled trial. *JMIR Ment. Health* **9**, e34002 (2022).
75. Linardon, J, Shatte, A, Rosato, J & Fuller-Tyszkiewicz, M Efficacy of a transdiagnostic cognitive-behavioral intervention for eating disorder psychopathology delivered through a smartphone app: a randomized controlled trial. *Psychol. Med.* **52**, 1679–1690 (2022).
76. Hirshberg, MJ et al. A randomized controlled trial of a smartphone-based well-being training in public school system employees during the COVID-19 pandemic. *J. Educ. Psychol.* **114**, 1895–1911 (2022).
77. Christoforou, M, Sáez Fonseca, JA & Tsakanikos, E Two novel cognitive behavioral therapy-based mobile apps for agoraphobia: randomized controlled trial. *J. Med. Internet Res.* **19**, e398 (2017).
78. Graham, AK et al. Coached mobile app platform for the treatment of depression and anxiety among primary care patients: a randomized clinical trial. *JAMA Psychiatry* **77**, 906 (2020).
79. Mak, WW et al. Efficacy and moderation of mobile app-based programs for mindfulness-based training, self-compassion training, and cognitive behavioral psychoeducation on mental health: Randomized controlled noninferiority trial. *JMIR Ment. Health* **5**, e60 (2018).
80. Ferré-Grau, C et al. A mobile app–Based intervention program for nonprofessional caregivers to promote positive mental health: randomized controlled trial. *JMIR Mhealth Uhealth* **9**, e21708 (2021).
81. Ham, K. et al. Preliminary results from a randomized controlled study for an app-based cognitive behavioral therapy program for depression and anxiety in cancer patients. *Front. Psychol.* **10**, 1592 (2019).
82. Boettcher, J et al. Adding a smartphone app to internet-based self-help for social anxiety: a randomized controlled trial. *Comput. Hum. Behav.* **87**, 98–108 (2018).
83. Litvin, S, Saunders, R, Maier, MA & Lüttke, S Gamification as an approach to improve resilience and reduce attrition in mobile mental health interventions: a randomized controlled trial. *PLoS One* **15**, e0237220 (2020).
84. McCloud, T, Jones, R, Lewis, G, Bell, V & Tsakanikos, E Effectiveness of a mobile app intervention for anxiety and depression symptoms in university students: randomized controlled trial. *JMIR Mhealth Uhealth* **8**, e15418 (2020).
85. Bakker, D, Kazantzis, N, Rickwood, D & Rickard, N A randomized controlled trial of three smartphone apps for enhancing public mental health. *Behav. Res. Ther.* **109**, 75–83 (2018).
86. Sawyer, A et al. The effectiveness of an app-based nurse-moderated program for new mothers with depression and parenting problems (EMUMS Plus): pragmatic randomized controlled trial. *J. Med. Internet Res.* **21**, e13689 (2019).
87. Hur, JW, Kim, B, Park, D & Choi, SW A scenario-based cognitive behavioral therapy mobile app to reduce dysfunctional beliefs in individuals with depression: a randomized controlled trial. *Telemed. e-Health* **24**, 710–716 (2018).

## Acknowledgements

S.B. reports financial support was provided by Australian National Health and Medical Research Council (2026634). M.A.J. was supported by an investigator grant (APP1177235) from the National Health and Medical Research Council and a Dame Kate Campbell Fellowship from the University of Melbourne. J. N. is supported by a National Health and Medical Research Council Emerging Leader Fellowship (ID 2009782).

## Author contributions

L.V., J.D.X.H., J.N., and M.A.J. designed the study. L.V., J.D.X.H., J.N., and M.A.J. developed the search terms. L.V. and J.D.X.H. conducted title/abstract and full-text screening. L.V., S.O., R.P.S., M.S.V., and W.P. conducted the data extraction. L.V., K.B., and L.B. conducted persuasive design principle coding. J.D.X.H., S.O., and P.L. conducted the analyses. T.W.W., C.M., S.N.M., I.H.B., and L.V., conducted the risk of bias assessment. L.V. wrote the first draft of the manuscript, with input from J.N., S.B., J.D.X.H., M.A.J., and S.C. All authors revised the manuscript and approved the final version of the submitted 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/s41746-025-01567-5>.

**Correspondence** and requests for materials should be addressed to Lee Valentine.

**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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

© The Author(s) 2025