Mediators and mechanisms of change in internet- and mobile-based interventions for depression: A systematic review

Matthias Domhardt a,*, Lena Steubl a, Johanna Boettcher b, Claudia Buntrock c, Eirini Karyotaki d,e, David D. Ebert d, Pim Cuijpers d,f, Harald Baumeister b

a Department of Clinical Psychology and Psychotherapy, Ulm University, Germany
b Department of Clinical Psychology and Psychotherapy, Psychologische Hochschule Berlin, Germany
c Department of Clinical Psychology and Psychotherapy, Friedrich-Alexander University Erlangen-Nuremberg, Germany
d Department of Clinical, Neuro and Developmental Psychology, Vrije Universiteit Amsterdam, the Netherlands
e Department of Global Health and Social Medicine, Harvard Medical School, United States
f Public Health Research Institute, Vrije Universiteit Amsterdam, the Netherlands

A R T I C L E   I N F O

Keywords:
Depression
e- and m-Health
Digital intervention
Psychotherapy
Process research
Mechanism of change

A B S T R A C T

The efficacy of Internet- and mobile-based interventions (IMIs) for depression in adults is well established. Yet, comprehensive knowledge on the mediators responsible for therapeutic change in these interventions is pending. Therefore, we conducted the first systematic review on mediators in IMIs for depression, investigating mechanisms of change in interventions with different theoretical backgrounds and delivery modes (PROSPERO CRD42019130301). Two independent reviewers screened references from five databases (i.e., Cochrane Library, Embase, MEDLINE/PubMed, PsycINFO and ICTRP), selected studies for inclusion and extracted data from eligible studies. We included 26 RCTs on mediators in IMIs for depression (6820 participants), rated their risk of bias and adherence to methodological quality criteria for psychotherapy process research. Primary studies examined 64 mediators, with cognitive variables (e.g., perceived control, rumination or interpretation bias) being the largest group of both examined (m = 28) and significant mediators (m = 22); followed by a range of other mediators, including mindfulness, acceptance and behavioral activation. Our findings might contribute to the empirically-informed advancement of interventions and mental health care practices, enabling optimized treatment outcomes for patients with depression. Furthermore, we discuss implications for future research and provide methodological recommendations for forthcoming mediation studies with more pertinent designs, allowing for inferences with higher causal specificity.

1. Introduction

Depression is highly prevalent (Kessler & Bromet, 2013) and associated with substantial personal and societal burden (Vos et al., 2016). Psychotherapeutic interventions with several theoretical backgrounds are effective in treating depressive symptoms, irrespective of whether they are implemented face-to-face (Barth et al., 2013; Cuijpers, Karyotaki, de Wit, & Ebert, 2019) or delivered remotely via the Internet (Cuijpers, Noma, Karyotaki, Cipriani, & Furukawa, 2019; Ebert et al., 2018; Karyotaki et al., 2018) or mobile-based communication technology (Firth et al., 2017; Linardon, Cuijpers, Carlbring, Messer, & Fuller-Tyszkiewicz, 2019). However, far less is known whereby and how these psychotherapeutic interventions “work” (Cuijpers, Cristea, Karyotaki, Reijnders, & Hollon, 2017; Cuijpers, Reijnders, & Huibers, 2019; Lemmens, Muller, Arntz, & Huibers, 2016), and in which aspects the change mechanisms in interventions with divergent theoretical orientations and delivery modes might differ. Evidence-based knowledge on these issues is important, as it can inform intervention development and mental health care practices (Cuijpers, 2016; Domhardt, Gelline, von Rezori, & Baumeister, 2019), highlighting ways to more powerful intervention packages with optimized treatment outcomes.

1.1. Mechanisms of change research – Concepts, definitions and methods

Knowledge on the mechanisms of change in psychotherapeutic interventions for depression can be obtained primarily by two research
approaches, component studies and mediation studies. Component studies aim to determine the influence of single intervention components on treatment efficacy, thereby contributing to the understanding whereby therapies work (Cuijpers, Reijnders, & Huibers, 2019; Domhardt et al., 2019). In these studies, multicomponent psychotherapeutic interventions are compared with disentangled variations of the same intervention (Borkovec & Costontguay, 1998) in which either a component is left out (i.e., dismantling studies) or added to the full treatment package (i.e., additive design studies). When these component studies are conducted within a framework of randomized controlled trials (RCTs), they generate strong causal evidence on the incremental (add on) effects of single intervention components (Cuijpers et al., 2017; Domhardt et al., 2019). These components can be consecutively considered as “active ingredients” of therapies and further differentiated into specific or common factors of psychotherapeutic interventions (Mulder, Murray, & Rucklidge, 2017). Specific factors are those therapeutic techniques and procedures that are inherent to a particular therapeutic approach, based on the respective model of psychotherapy and/or delineated from a disorder-specific etiological model (DeRubeis, Brotman, & Gibbons, 2008); examples for specific factors are cognitive restructuring and behavioral activation related to CBT, or defensive and reflective functioning related to psychodynamic therapies. In contrast, common factors are considered to be generally constituent in psychotherapeutic interventions, regardless of their theoretical orientation or specific therapeutic approach (Wampold, 2015); examples for common factors are the working alliance or the induction of expectations or hope. Meanwhile, the long-lasting debate about the relative importance and differential role of specific and common factors seems to shift to the joint notion that both factors – as well as their respective interactions – contribute to treatment effectiveness (Mulder et al., 2017; Tschacher, Junghan, & Pfammatter, 2014).

Nevertheless, insights that are generated with component studies do not necessarily reveal how psychotherapeutic interventions work, and which processes within patients lead to therapeutic change. These endeavors are the main purpose of mediation studies (Kazdin, 2007, 2009), which can be regarded as the essence and core approach of psychotherapy process research. Mediators, like self-efficacy (Bandura, 1978) or functional thoughts (Beck, Rush, Shaw, & Emery, 1979), are intervening variables that account statistically for the relationship between a psychotherapeutic intervention and an outcome variable (Kazdin, 2007), and thereby potentially point to mechanisms through which the intervention achieves its effects (Kraemer, Wilson, Fairburn, & Agras, 2002). Yet, it is important to bear in mind that mediators might be congruent and identical to a mechanism (i.e., the actual process responsible for change), but might also be a proxy for one or more other variables that do not explain the hypothesized mechanism (Kazdin, 2007). These proxy variables are either (a) unknown and/or unmeasured, or (b) both known and measured, with subsequent implications for the scope of causal inferences (Ohllson & Kendler, 2019). Accordingly, in the context of psychotherapy process research it is worthwhile to mediate are deduced theoretically (i.e., conceptually represent a plausible change mechanism that is induced by an intervention component), are defined and operationalized a priori, and are tested for other known potential intervening variables (Johansson & Hoglend, 2007; Kazdin, 2007; Kraemer et al., 2002; Laurenceau, Hayes, & Feldman, 2007).

In contrast, moderators are variables and demographic characteristics, like comorbidity or age, that influence the magnitude or direction of the relationship between an intervention and outcome, thereby indicating on whom and under which circumstances treatments have differential effects (Kazdin, 2007; Kraemer et al., 2002). Variables that qualify for moderators must precede an intervention and need to be independent to the intervention (Kraemer, Kiernan, Essex, & Kupfer, 2008); whereas variables that qualify for mediators always come after an intervention (during treatment) and are related to the intervention (Kraemer et al., 2008). As such, moderators suggest that different processes are involved (i.e., “moderated mediation”, James & Brett, 1984; Kazdin, 2007), but conceptually represent no mechanisms of therapeutic change themselves (Kraemer et al., 2002; Laurenceau et al., 2007).

To date, the most widely used method to statistically evaluate mediators was introduced by Reuben Baron and David Kenny with their seminal paper in 1986, also known as the causal steps approach. In this approach, four conditions are tested consecutively within a regression framework (Baron & Kenny, 1986), in order to demonstrate that a mediator \( M \) transmits the effect of an independent variable \( X \) (the intervention) on a dependent variable \( Y \) (the outcome). Although the method of Baron and Kenny to mediation analysis is still very prominent in different scientific disciplines, this approach has several limitations that are especially disadvantageous in psychotherapy process research on depression. For instance, as the statistical power associated with the causal steps method is very low (as are type I error rates), large sample sizes and large treatment effects are necessary to be able to detect indirect effects (MacKinnon et al., 2002; MacKinnon, Faiichild, & Fritz, 2007); yet, these are two conditions that are hardly present in psychotherapy process research on depression (Lemmens et al., 2016). Moreover, it is challenging to apply the causal steps method to models with multiple intervening variables and to assess each of the indirect effects independently (MacKinnon et al., 2007); however, it is most likely that several distinct processes are involved in psychotherapeutic interventions for common mental disorders (Cuijpers, Reijnders, & Huibers, 2019; Domhardt et al., 2019), and multi-mediator models should be evaluated accordingly (Kazdin, 2007, 2009). Most importantly, the causal steps method is only applicable when differential treatment effects are available. Considering the common finding that several psychotherapeutic interventions are equally effective in treating depression (e.g., Barth et al., 2013; Cuijpers, Karyotaki, et al., 2019), it is a significant downside that mediation analyses after Baron and Kenny are not practical under these circumstances – while knowledge on whether these equally effective treatments operate with shared or distinct working mechanisms would be of particular importance.

Consequently, recent advancements aimed to alleviate these limitations and strive to develop statistical methods more suitable for psychotherapy research. Most prominently, Helena Kraemer and the MacArthur group proposed a novel framework that builds upon the same logic of the causal steps method, but opened up opportunities to conduct mediation analyses even when no differential treatment effects are observed and/or the statistical power would be too low to detect intervening variable effects with the Baron and Kenny approach (Kraemer et al., 2002; Kraemer et al., 2008; MacKinnon et al., 2007). In the former case, an interaction between the intervention and the mediator is deemed to be sufficient (Kraemer et al., 2002), and in the latter a “joint significance” test is endorsed (MacKinnon et al., 2002); a procedure, in which it is simultaneously tested whether the independent variable is related to the intervening variable and whether the intervening variable is related to the dependent variable (Holllon & DeRubeis, 2009; MacKinnon et al., 2002). Further recent developments and approaches to mediation analyses are comprehensively summarized by Hayes (2018), Hofmann, Curtiss, and Hayes (2020) and MacKinnon (2017).

1.2. Criteria for the detection of causality and implications for study designs

Still, statistical mediation alone is not sufficient to detect mechanisms of change and corroborate causality (i.e., the relation between cause and effect, Gregor, 2006). Several other criteria strengthen the evidence for actual change processes and pave the way toward the determination of causality (Hernan & Robins, 2019; Kazdin, 2007, 2009) – a subject of philosophical debate for more than 2000 years (Kraemer et al., 2008; Ohllson & Kendler, 2019). Aside the aforementioned criteria of a theoretical foundation of plausible processes and the
necessity of a strong statistical association (between intervention, mediator and outcome), a basic prerequisite to assume causality is temporal precedence. The essential criterion of temporality is met, when a timeline can be established that the intervention leads to changes in the mediator, and the changes in the mediator precede changes in the outcome – and not vice versa (Kazdin, 2007; Kraemer et al., 2008). Hence, only longitudinal study designs with several concurrent assessments of the mediator and outcome variables with at least three measurement points are assumed to be suitable to evaluate this criterion and establish the actual directionality and shape of change (Laurenceau et al., 2007). The criteria of specificity (of the association between intervention, mediator and outcome), gradient (i.e., dose-response relationship between mediator and outcome) and consistency (of results on mediators across studies) can further back the case for causality (Kazdin, 2007, 2009). Moreover, findings are enormously strengthened when they result from an experimental design. This can be achieved either by randomly assigning adequate numbers of participants to different comparison conditions in an RCT (Laurenceau et al., 2007), or by directly manipulating the mediator variable itself (Kazdin, 2007). In this latter approach, it is safe to ascertain that the experimental variation of the mediator causally leads to the observed changes in the outcome, whereas in the former approach one can merely assume that the intervention effects are mediated by the purported construct and not by other conceptually differing proxy variables. This is also why the concurrent assessment of multiple mediators – as well as multiple outcomes (Laurenceau et al., 2007) – is highly advisable in order to falsify rivaling hypotheses on therapeutic processes (Kazdin, 2007, 2009). Altogether, multiple converging lines of research observant to these criteria and resorting to different methodologies, while addressing the same question (i.e., triangulation; Munafò & Davey Smith, 2018), are necessary to approximate specific mechanisms of change and prove causality (Kazdin, 2007, 2009; Ohlsson & Kendler, 2019).

However, it is also to acknowledge that the criteria of specificity, gradient and consistency are not absolute, and several important exceptions to these arguments need to be considered. For instance, it might be that different treatment modalities, such as CBT or pharmacotherapy, operate with shared and/or distinct processes, on the same and/or differing levels of action (e.g., behavioral activation vs. alterations in neurotransmission; Hollon, 2019). And these shared (e.g., modifying distorted cognitions) or distinct mechanisms might further happen simultaneously and/or consecutively during treatment, might be independent to another or influence each other dynamically (Hofmann et al., 2020). It is thus central to differentiate between non-specificity in causal processes and actual non-causality, as well as to consider possible differential associations between interventions and mediators on the one hand, and mediators and outcomes on the other hand (Hollon, DeRubeis, & Evans, 1987). Furthermore, the association of a mediator with an outcome might be linear and therefore imply a gradient, but might be also in another multifaceted relation (e.g., U- or S-shaped) that does not concur with a dose-response relationship (Hofmann et al., 2020; Laurenceau et al., 2007). These examples illustrate that the findings on mediators and mechanisms of change must not be necessarily consistent, but might even differ considerably between studies, owing to the complexities of human behavior change.

1.3. Evidence for mediators and mechanisms of change in psychotherapeutic interventions for depression

The evidence base on mediators in face-to-face psychotherapies for depression derived from systematic reviews indicates that intervention effects are mediated by a range of cognitive, behavioral and emotional processes, as well as by the acquisition and usage of different skills (Lemmens et al., 2016). The majority of studies investigated numerous cognitive mediators like dysfunctional thinking (Cristesu et al., 2015) or cognitive changes in general (Lorenzo-Luaces, German, & DeRubeis, 2015), and provide some support to the theorized cognitive processes of CBT and its third wave variant Acceptance and Commitment Therapy (ACT; Stockton et al., 2019) to this point. Furthermore, some specific cognitive processes of psychodynamic therapies like changes in insight (Johansson et al., 2010) and self-understanding (Leichsenring, Steinetz, & Crits-Christoph, 2018) are documented in samples with mixed diagnoses so far. Besides, behavioral activation has been shown to represent a relevant factor in the treatment of depression (Baumeister, 2017; Jacobson et al., 1996; Richards et al., 2016) and might unfold its effects by decreasing avoidance behavior, as well as increasing positive experiences and reward (Hopko, Ryba, McIndoo, & File, 2016). In addition, some emotional processes (Lemmens et al., 2016) are found to be therapeutic when addressed with different emotion-regulation strategies, like reappraisal, changes in rumination or problem-solving (Aldao, Nolen-Hoeksema, & Schweizer, 2010). In line with the training and growth aspect of evidence-based psychotherapies, the acquisition and usage of mindfulness skills (van der Velden et al., 2015), interpersonal skills (Lipsitz & Markowitz, 2013) and problem-solving skills (Malouff, Thorsteinsson, & Schutte, 2007) are all found to be significant mediators in psychotherapies for depression (Lemmens et al., 2016). Overall, only few mediation studies in these reviews met all quality criteria for process research at once, largely precluding causal inferences so far.

In contrast to the evidence base on face-to-face psychotherapies, the knowledge on the mediators in technology-delivered psychotherapeutic interventions for depression is even more fragmented and limited. To date, there is only one single narrative review on mediators in Internet-based and computerized CBT for depression (Mogoaşă, Cobei, David, Giosan, & Szentagotai, 2017), ignoring mobile-based interventions and neglecting therapeutic approaches beyond CBT. This gap in the literature is a significant drawback, as both Internet- and mobile-based interventions (IMIs) possess the potential to complement and augment existing psychotherapeutic services for depression (Firth et al., 2017; Karyotaki et al., 2017; Karyotaki, Ebert, et al., 2018), considering their particular features and presumed advantages, like flexibility and anonymity in conduct, possible cost-effectiveness and scalability on a population scale. In these ways, digital health interventions are thought to contribute to the mitigation of the disease burden associated with depression substantially (Buntrock et al., 2016; Domhardt, Steubl, & Baumeister, 2018; Ebert, Cuijpers, Muñoz, & Baumeister, 2017; Karyotaki et al., 2017; Karyotaki, Ebert, et al., 2018; Königsbauer, Letsch, Doebler, Ebert, & Baumeister, 2017).

1.4. The current study

Accordingly, specific knowledge on the mechanisms of change in guided (with accompanying human support) and unguided (without support) IMIs for depression is of high relevance, in order to build their distinct evidence base respective to their theoretical foundation, contributing to enhanced intervention development and informing the implementation of IMIs in routine clinical care. Furthermore, IMIs convey novel technological and methodological opportunities for psychotherapy research into change mechanisms that might not only disclose the active ingredients of IMIs themselves, but ancillary also contribute to unravel the processes in face-to-face interventions. This is because IMIs allow for a higher standardization of interventions, as prior confounds of psychotherapy research like treatment fidelity (i.e., the actual presentation of therapeutic content) or therapist factors (e.g., professional experience, personal views) can be hold constant, and intervention components can be clearly and reliably operationalized (Domhardt et al., 2019; Furukawa et al., 2019; Steubl, Sachser, Baumeister, & Domhardt, 2019). Hence, IMIs enable an unprecedented way to break down the utterly complex processes of psychotherapeutic interventions into paradigmatic fragments and allow for the evaluation of specific mechanisms of therapeutic change in original experimental designs (Domhardt et al., minor revision). In addition, studies with IMIs might facilitate the attainment of sample sizes sufficient to detect small, but clinically relevant intervening variable effects, as participants can be
recruited on large scale and interventions can be implemented irrespective of the constraints of space and time. Research exploiting these possibilities might thereby contribute to a better understanding to which degree technology-delivered and face-to-face psychotherapeutic interventions operate with shared and/or distinct processes; and disclose, if there are actually specific digital components, as well as unique mediators and mechanisms of change in IMIs for depression.

Considering these novel opportunities and acknowledging the fast growing field of digital health research (Andersson, Titov, Dear, Rozenthal, & Carbring, 2019), it seems both timely and worthwhile to provide the first comprehensive systematic overview on mechanisms of change in IMIs for depression. As such, with this study we aimed to systematically review and methodologically assess studies that reported on mediators in various forms of IMIs for depression in adults, informing psychotherapy process research, intervention development and the implementation of IMIs in routine care alike.

2. Method

This systematic review was conducted in line with the recommendations of the Cochrane Collaboration (Higgins & Green, 2011) and is reported in accordance to the PRISMA statement (Moher, Liberati, Tetzlaff, & Altman, 2009). Further detailed information on the design and methodology of this study can be found within the protocol at PROSPERO (CRD42019130301).

2.1. Eligibility criteria and search strategy

In this review we included (a) RCTs, in which (b) a guided or un-guided IMI for depression (of various theoretical backgrounds) was compared to (c) an active or non-active control group, (d) in adults (≥18 years) (e) with symptoms of depression based on self-report or diagnostic interview. Eligible studies had to report (f) depression symptom severity as outcome and (g) examine at least one mediator variable by means of statistical mediation analysis (all approaches that qualify for mediation analyses were eligible). Studies with (h) non-clinical samples (i.e., depression symptoms were assessed without applying a clinical cut-off score) were eligible as well, but are reported separately. Studies with (i) an experimental manipulation of mediator variables were also included. Studies had to be (j) written in English, had obtained ethical approval, and (k) were published in a peer-reviewed journal, or (l) indexed “with preliminary/unpublished results” in the International Clinical Trials Registry Platform (ICTRP). Overall, with the application of these inclusion criteria, we intended to enable a comprehensive overview of this branch of research, allowing for comparisons between studies with “clinical” and “non-clinical” samples, as well as between interventions with divergent theoretical underpinnings and delivery modes.

The search strategy was threefold. First, the electronic databases Cochrane Library (CENTRAL), Embase, MEDLINE/PubMed and PsycINFO were systematically searched with a predefined set of search strings in Ovid, from database inception to September 1, 2019. Second, the ICTRP was consulted to detect studies with preliminary or unpublished results. Third, the reference lists of all eligible studies were searched by hand, to identify other relevant studies.

The combinations of index, MESH and free terms were indicative for “Depression”, “Internet”, “Mobile”, “Psychotherapy”, “Intervention”, “Mediators” and “Mechanisms”. The full list of search strings, as well as the search strategies for each database can be found in the Appendix (Tables A.1, A.2).

2.2. Study selection and data extraction

Study selection and data extraction were conducted independently by two researchers (MD and LS). Disagreements were solved by discussion. Titles and abstracts were screened, and irrelevant records were discarded. The full texts of remaining references were retrieved, and assessed according to the aforementioned inclusion criteria.

The following data were extracted from eligible studies: (a) study identification items, (b) sample characteristics, (c) comparison groups, (d) number of participants, (e) mean age, (f) drop-out and intervention adherence, (g) number of modules and intervention duration, (h) post-treatment and follow-up, (i) measures of depression symptom severity, (j) potential mediators (with instruments and measurement points), and (k) statistical methods for mediation analyses, as well as (l) clinical outcomes.

2.3. Risk of bias assessment and evaluation of methodological quality for process research

To determine the risk of bias and methodological quality of included studies, two rating systems were applied independently by two reviewers (MD and LS), with disagreements solved by discussion. First, the Cochrane Collaboration’s tool for assessing risk of bias (Higgins et al., 2011) was deployed on all papers on all seven criteria (i.e., “random sequence generation”, “allocation concealment”, “blinding of participants and personnel”, “blinding of outcome assessment”, “incomplete outcome data”, “selective reporting”, and “other bias”), as having either “low”, “unclear” or “high risk of bias”.

Second, in order to evaluate the appropriateness of research designs of included studies to detect mechanisms of change and to approach causality, a rating system based on the criteria proposed by Alan Kazdin (2007, 2009) and adapted to psychotherapy research for depression by Lemmens et al. (2016) was applied. Each study was rated on the following criteria as either fulfilled (“+”) or not fulfilled (“−”): (a) RCT, (b) control group, (c) theoretical foundation, (d) minimal threshold for sample size per group (n > 40), (e) concurrent examination of multiple mediators, (f) evaluation of temporality (i.e., three or more longitudinal assessments of mediators and outcome variables), and (g) direct experimental manipulation of a mediator. In contrast to Lemmens et al. (2016), the criterion of theoretical foundation of mediators was deployed additionally in the rating system. In line with Lemmens et al. (2016), three aspects originally suggested by Kazdin (2007) were not incorporated in the rating system, as they are either not meaningfully applicable in single studies (consistency) or might be too exclusive for certain mechanisms (gradient and specificity). Nevertheless, these issues are addressed in the discussion.

Since our primary intention was to assess the methodological quality of included studies and to guide this novel branch of process research conceptually, we did not plan to meta-analytically evaluate mediators within this review (PROSPERO CRD42019130301). A theoretically-deduced framework on the mechanisms of change in IMIs for depression will be reported elsewhere (Domhardt et al., in preparation).

3. Results

3.1. Study characteristics

The systematic searches for mediation studies yielded in 3255 initial hits; after the removal of duplicates and title and abstract screening, 58 articles were assessed in full text review with reasons for exclusion outlined in a PRISMA flow chart (Fig. 1). Finally, a total of 26 RCTs met the eligibility criteria and were included in this review. Half of these studies examined samples with clinically depressed patients (n = 2769) (Choi et al., 2014; Clarke et al., 2014; Forand et al., 2018; Hirsch et al., 2018; Lamers et al., 2015; Newby, Williams, & Andrews, 2014; Pictet, Jernmann, & Ceschi, 2016; Pots, Trompeter, Schreurs, & Bohlmeijer, 2016; Seeley et al., 2019; Terides et al., 2017; van der Zanden, Galindo-Garre, Curie, Kramer, & Cuijpers, 2014; Warmerdam, van Straten, Jongsmra, Twisk, & Cuijpers, 2010; Williams, Blackwell, Mackenzie, Holmes, & Andrews, 2013); whereas the other 13 studies focused on samples in which depression symptoms were assessed without applying...
a clinical cut-off score (i.e., “non-clinical”), comprising mixed samples with other physical and mental impairments, or without pathology (n = 4051) (Auyeung & Mo, 2018; Bakker, Kazantzis, Rickwood, & Rickard, 2018; Heckendorf, Lehr, Ebert, & Freund, 2019; Kauer et al., 2012; Levin, Haeger, Pierce, & Twohig, 2017; Ma, She, Siu, Zeng, & Liu, 2018; Morgan, Mackinnon, & Jorm, 2013; Morris, Schueller, & Picard, 2015; Querstret, Cropley, & Fife-Schaw, 2018; van Luenen, Kraaij, Spinhoven, Wilderjans, & Garnefski, 2019; Viskovich & Pakenham, 2018; Viskovich & Pakenham, 2019; Willems et al., 2017). Table 1 provides a detailed overview on study characteristics and extracted data.

Most studies were conducted in western countries (92.3%; i.e., 34.6% Australia, 23.1% the Netherlands, 19.2% the USA, 7.7% the United Kingdom, 3.8% Germany, and 3.8% Switzerland; vs. 7.7% China) and were published within the past five years (2014–2019; 84.6%). The majority of studies (k = 18 studies; 69.2%) had recruited participants via (online) advertisements. Other recruitment settings included (online) clinics (k = 3; 11.5%), agencies or specialized centers (k = 2; 7.7%), and general practitioners (k = 2; 7.7%), as well as an established waitlist (k = 1; 3.8%). The mean ages of participants across studies ranged from 18.7 to 64.8 years (M = 36.6; SD = 12.3), and almost three-fourths of included participants were female (73.3%; range: 12% to 100%). Sample sizes ranged between 69 and 1162 participants, with a mean sample size of 262 participants (SD = 243). In total, 6820 participants were included in this review.

Internet-based interventions were evaluated in 21 studies (80.8%) and mobile-based interventions were examined in two studies (7.7%). Other studies investigated an intervention combining Internet- and mobile-components (k = 2; 7.7%) or psychotherapy via Internet-based video calls (k = 1; 3.8%). The most frequently researched theoretical background of IMIs was CBT (k = 8; 30.8%), followed by ACT (k = 4; 15.4%), Mindfulness Based Cognitive Therapy (MBCT; k = 2; 7.7%) and Positive Psychology (k = 2; 7.7%). Further underpinnings of IMIs
<table>
<thead>
<tr>
<th>Authors (Year); Country</th>
<th>Sample / recruitment strategy</th>
<th>Comparison groups</th>
<th>Number of participants per group (% females)</th>
<th>Mean age (SD)</th>
<th>Study dropout*</th>
<th>Mean number of modules completed time spent</th>
<th>Percentages of participants completing all modules</th>
<th>Modules (no.) / duration in weeks*</th>
<th>Post-treatment; follow-up (s) in weeks*</th>
<th>Depression symptom severity instruments</th>
<th>Potential mediators</th>
<th>Mediator instruments (measurement points in weeks*)</th>
<th>Statistical method for mediation analyses (authors)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choi et al. (2014); US</td>
<td>Homebound older adults with depressive symptoms / aging-service agencies</td>
<td>Tele-PST vs. in-person PST vs. telephone care calls</td>
<td>56 / 63 / 39 (79%)</td>
<td>64.8 (9.18)</td>
<td>12%</td>
<td>–</td>
<td>–</td>
<td>6/–</td>
<td>12; 24, 36</td>
<td>HAMD</td>
<td>↓ Disability status</td>
<td>WHODAS II (0, 12, 24, 36)</td>
<td>Latent growth curve mediation modelling within a framework of SEM (Cheong et al., 2003; Mackinnon, 2008) Mixed models repeated measures (Baron &amp; Kenny, 1986) and revised version of the Sobel Test (Preacher &amp; Hayes, 2004)</td>
</tr>
<tr>
<td>Clarke et al. (2014)*; AUS</td>
<td>Self-reported mild-to-moderate symptoms of depression, anxiety and stress / internet, radio, print</td>
<td>Mobile phone and web intervention vs. attention control condition vs. WLC</td>
<td>242 / 248 / 230 (70%)</td>
<td>38.9 (–)</td>
<td>28%</td>
<td>1.6 modules/ 14.7 log-ins / 49 self-monitor entries</td>
<td>–</td>
<td>12/7</td>
<td>8; –</td>
<td>DASS (depress. subscale)</td>
<td>↑ Self-efficacy (n.s.)</td>
<td>MHSES (0, 8)</td>
<td>PROCESS (Hayes, 2013)</td>
</tr>
<tr>
<td>Forand et al. (2018); US</td>
<td>Self-reported depressive symptoms / online, flyers, newspapers, referral</td>
<td>Guided iCBT vs. WLC</td>
<td>60 / 30 (75%)</td>
<td>33 (–)</td>
<td>24%</td>
<td>6.22 sessions /</td>
<td>56%</td>
<td>8/8</td>
<td>8; –</td>
<td>PHQ-9 (primary outcome); HAMD</td>
<td>↑ Cognitive skills ↑ Behavioral activation (n.s.)</td>
<td>CCTS-SR (0, 3, 8) BADS-SF (0, 3, 8)</td>
<td>PROCESS (Hayes, 2013)</td>
</tr>
<tr>
<td>Hirsch et al. (2018); UK</td>
<td>Diagnosed generalized anxiety disorder or depression (SCID) / online, newspapers, university e-mails</td>
<td>Standard CBM vs. CBM with prior RNT vs. active control</td>
<td>44** / 43** / 44** (82%)</td>
<td>29.9 (11.19)</td>
<td>16%</td>
<td>–</td>
<td>74%</td>
<td>10/4</td>
<td>4; 4</td>
<td>PHQ-9</td>
<td>↓ Negative interpretation bias (approached significance: p = .054) ↓ Negative interpretation bias (n.s.)</td>
<td>SST (0, 4, 8)</td>
<td>SEM (product of coefficients approach; Iacobucci et al., 2007)</td>
</tr>
<tr>
<td>Lamers et al. (2015); NL</td>
<td>Self-reported mild to moderate depressive symptoms / online, newspaper</td>
<td>Guided life-review vs. guided expressive writing vs. WLC</td>
<td>58 / 58 / 58 (77%)</td>
<td>56.9 (9.15)</td>
<td>10%</td>
<td>- / 5.25 h peer week / 2.29 h peer week</td>
<td>75.9%</td>
<td>7/10</td>
<td>13; 26, 52</td>
<td>CES-D</td>
<td>↑ Ego-integrity</td>
<td>Northwestern Ego-Integrity scale (0, 13, 26, 52) Stepwise regression analyses and bootstrapping (Preacher &amp; Hayes, 2004)</td>
<td></td>
</tr>
<tr>
<td>Newby et al. (2014); AUS</td>
<td>Self-reported and diagnosed (MINI) MDD, GAD or both / waitlist, online</td>
<td>Guided iCBT vs. WLC</td>
<td>46 / 53 (78%)</td>
<td>44 (12.2)</td>
<td>3%</td>
<td>–</td>
<td>89%</td>
<td>6/10</td>
<td>10; 13</td>
<td>PHQ-9</td>
<td>↓ Repetitive negative thinking (RNT) ↓ Positive metacognitive</td>
<td>RTQ (0, 4-10***, 10, 13) PBRS-A (0, 4-10***, 10, 13)</td>
<td>PROCESS (Hayes, 2013)</td>
</tr>
</tbody>
</table>

(continued on next page)
<table>
<thead>
<tr>
<th>Authors (Year); Country</th>
<th>Sample / recruitment strategy</th>
<th>Comparison groups</th>
<th>Number of participants per group (% females)</th>
<th>Mean age (SD)</th>
<th>Study dropout*</th>
<th>Mean number of modules completed / time spent</th>
<th>Percentages of participants completing all modules</th>
<th>Modules (no.) / duration in weeks*</th>
<th>Post-treatment; follow-up (s) in weeks*</th>
<th>Depression symptom severity instruments</th>
<th>Potential mediators</th>
<th>Mediator instruments (measurement points in weeks*)</th>
<th>Statistical method for mediation analyses (authors)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pictet et al. (2016); CH</td>
<td>Self-reported depressive symptoms / flyers, online</td>
<td>Imagery CBM vs. control CBM vs. WLC</td>
<td>34 (34)</td>
<td>26.7 (9.14)</td>
<td>4%</td>
<td>97%</td>
<td>4/1</td>
<td>1; 3</td>
<td>BDI-II</td>
<td>† Depression-related interpretation bias</td>
<td>AST-D-II (0,1,3)</td>
<td>PROCESS (Hayes, 2012)</td>
<td></td>
</tr>
<tr>
<td>Pots et al. (2016); NL</td>
<td>Self-reported mild to moderate depressive symptoms (&amp; MINI) / newspapers, online</td>
<td>Web-based ACT vs. online expressive writing vs. WLC</td>
<td>87 (87)</td>
<td>46.9 (12.06)</td>
<td>16%</td>
<td>73%</td>
<td>9/9-12</td>
<td>13; 26, 52</td>
<td>CES-D</td>
<td>† Psycholog. flexibility</td>
<td>AAQ-II (0, 13, 26, 52)</td>
<td>PROCESS (Hayes, 2013)</td>
<td></td>
</tr>
<tr>
<td>Seeley et al. (2019); US</td>
<td>Mothers with elevated depressive symptoms (CES-D; PHQ-8) / Head Start centers</td>
<td>Trial 1: Guided iCBT (+ MI) vs. delayed intervention / facilitated TAU</td>
<td>70</td>
<td>31.0 (6.4)</td>
<td>1%</td>
<td>6.4 sessions / 15.1 h</td>
<td>-</td>
<td>8/8</td>
<td>14</td>
<td>BDI-II (Trial 1)</td>
<td>† Behavioral activation†</td>
<td>BADS (0, 14/26)</td>
<td>Mplus (Muthén &amp; Muthén, 1998-2012); bootstrapping (Mackinnon et al., 2004)</td>
</tr>
<tr>
<td>Terides et al. (2017); AUS</td>
<td>Self-reported clinically mild symptoms of anxiety (GAD-7), depression (PHQ-9), or both (MINI) / seeking online treatment</td>
<td>iCBT vs. WLC</td>
<td>68 (68%); 80 (81%)</td>
<td>44.7 (12.75)</td>
<td>14%</td>
<td>63%</td>
<td>5/8</td>
<td>8; 13</td>
<td>PHQ-9</td>
<td>† CBT skills usage (cognitive restructuring, rewarding behaviors*, and activity scheduling)</td>
<td>FATS (0, 8, 13; both groups)</td>
<td>PROCESS (Hayes, 2013)</td>
<td></td>
</tr>
<tr>
<td>van der Zanden et al. (2014); NL</td>
<td>Young adults with self-reported depressive symptoms / adds in GP offices, institutions, online</td>
<td>iCBT group course vs. WLC</td>
<td>121 (121)</td>
<td>20.9 (2.2)</td>
<td>20%</td>
<td>3.2 sessions / -</td>
<td>20%</td>
<td>6/6</td>
<td>12; 24</td>
<td>CES-D</td>
<td>† Mastery</td>
<td>Mastery Scale (0, 12, 24)</td>
<td>Path analysis mediation models in Mplus (Baron &amp; Kenny, 1986; Mackinnon, 2008; Muthén &amp; Muthén, 1998-2012)</td>
</tr>
</tbody>
</table>

(continued on next page)
<table>
<thead>
<tr>
<th>Authors (Year); Country</th>
<th>Sample / recruitment strategy</th>
<th>Comparison groups</th>
<th>Number of participants per group (% females)</th>
<th>Mean age (SD)</th>
<th>Study dropout</th>
<th>Mean number of modules completed / time spent</th>
<th>Percentages of participants completing all modules</th>
<th>Modules (no.) / duration in weeks*</th>
<th>Post-treatment; follow-up (s) in weeks*</th>
<th>Depression symptom severity instruments</th>
<th>Potential mediators</th>
<th>Mediator instruments (measurement points in weeks*)</th>
<th>Statistical method for mediation analyses (authors)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warmerdam et al. (2010); NL</td>
<td>Self-reported depressive symptoms / adds in newspapers, online</td>
<td>Guided iCBT vs. guided online PST vs. WLC</td>
<td>88 (45 (12.1) 34% --</td>
<td>38.6%</td>
<td>8/8</td>
<td>5, 8; --</td>
<td>CES-D</td>
<td>↓ Dysfunction attitudes ↓ Worrying</td>
<td>PSWQ (0, 5, 8)</td>
<td>Mastery Scale (0, 5, 8)</td>
<td>SPSS-R (0, 5, 8)</td>
<td>Linear mixed modelling (Baron &amp; Kenny, 1986)</td>
<td></td>
</tr>
<tr>
<td>Warmerdam et al. (2010); NL</td>
<td>Self-reported depressive symptoms / adds in newspapers, online</td>
<td>Guided iCBT vs. guided online PST vs. WLC</td>
<td>88 (45 (12.1) 34% --</td>
<td>37.5%</td>
<td>5/5</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Williams et al. (2013); AUS</td>
<td>Diagnosed major depressive episode (MINI) / research unit</td>
<td>CBM-I (7-day) + iCBT (10-week) vs. WLC</td>
<td>38 (44.8 (12.05) 39% --</td>
<td>50%</td>
<td>13/11</td>
<td>11; --</td>
<td>BDI-II</td>
<td>↑ Change in interpretation bias</td>
<td>AST-D (0,1)</td>
<td>↑ Sense of autonomy ↑ Positive affects (n.s.) ↑ Coping self-efficacy</td>
<td>Subscale of BPSNPS (0, 1) Subscale of PANAS (0, 1) GSES (0, 4)</td>
<td>PROCESS (Hayes, 2012)</td>
<td></td>
</tr>
<tr>
<td>Williams et al. (2013); AUS</td>
<td>Diagnosed major depressive episode (MINI) / research unit</td>
<td>CBM-I (7-day) + iCBT (10-week) vs. WLC</td>
<td>31 (44.8 (12.05) 39% --</td>
<td>50%</td>
<td>13/11</td>
<td>11; --</td>
<td>BDI-II</td>
<td>↑ Change in interpretation bias</td>
<td>AST-D (0,1)</td>
<td>↑ Sense of autonomy ↑ Positive affects (n.s.) ↑ Coping self-efficacy</td>
<td>Subscale of BPSNPS (0, 1) Subscale of PANAS (0, 1) GSES (0, 4)</td>
<td>PROCESS (Hayes, 2012)</td>
<td></td>
</tr>
<tr>
<td>Non-clinical samples</td>
<td>Non-clinical samples</td>
<td>Non-clinical samples</td>
<td>Non-clinical samples</td>
<td>Non-clinical samples</td>
<td>Non-clinical samples</td>
<td>Non-clinical samples</td>
<td>Non-clinical samples</td>
<td>Non-clinical samples</td>
<td>Non-clinical samples</td>
<td>Non-clinical samples</td>
<td>Non-clinical samples</td>
<td>Non-clinical samples</td>
<td>Non-clinical samples</td>
</tr>
<tr>
<td>Auyeung and Mo (2018); CN</td>
<td>University students / mail, social media</td>
<td>Online PPI vs. writing group</td>
<td>70 (22.8 (3.78) 28% --</td>
<td>57% (≥ 5 sessions)</td>
<td>6/</td>
<td>1; --</td>
<td>CESD-10</td>
<td>↑ Sense of autonomy ↑ Positive affects (n.s.) ↑ Coping self-efficacy</td>
<td>↑ Emotionally focused self-care (n.s.) ↑ Emotional self-awareness (n.s.) ↑ Mental health literacy (n.s.)</td>
<td>ESAS-R (0, 4)</td>
<td>↑ Repetitive negative thinking</td>
<td>PTQ (0, 6, 13, 26)</td>
<td>Parallel multiple mediation analyses (Hayes and Rockwood, 2017)</td>
</tr>
<tr>
<td>Bakker et al. (2018); AUS</td>
<td>Community sample / social media</td>
<td>CBT toolkit app vs. mood tracking app vs. CBT strategy app vs. WLC</td>
<td>78 (34.2 (12.1) 55% --</td>
<td>36% (imputed 30-day)</td>
<td>--/4</td>
<td>4; --</td>
<td>PHQ-9</td>
<td>↑ Sense of autonomy ↑ Positive affects (n.s.) ↑ Coping self-efficacy</td>
<td>↑ Emotionally focused self-care (n.s.) ↑ Emotional self-awareness (n.s.) ↑ Mental health literacy (n.s.)</td>
<td>ESAS-R (0, 4)</td>
<td>↑ Repetitive negative thinking</td>
<td>PTQ (0, 6, 13, 26)</td>
<td>Parallel multiple mediation analyses (Hayes and Rockwood, 2017)</td>
</tr>
<tr>
<td>Heckendorf et al. (2019); GER</td>
<td>Community sample / online news magazine</td>
<td>Internet- and app-based gratitude intervention vs. WLC</td>
<td>132 (42.4 (10.9) 20% --</td>
<td>53.8%</td>
<td>5 / 5</td>
<td>6; 13, 26</td>
<td>CES-D</td>
<td>↓ Repetitive negative thinking</td>
<td>↓ Emotional self-awareness (n.s.) ↓ Mental health literacy (n.s.)</td>
<td>PTQ (0, 6, 13, 26)</td>
<td>Parallel multiple mediation analyses (Hayes and Rockwood, 2017)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heckendorf et al. (2019); GER</td>
<td>Community sample / online news magazine</td>
<td>Internet- and app-based gratitude intervention vs. WLC</td>
<td>130 (42.4 (10.9) 20% --</td>
<td>53.8%</td>
<td>5 / 5</td>
<td>6; 13, 26</td>
<td>CES-D</td>
<td>↓ Repetitive negative thinking</td>
<td>↓ Emotional self-awareness (n.s.) ↓ Mental health literacy (n.s.)</td>
<td>PTQ (0, 6, 13, 26)</td>
<td>Parallel multiple mediation analyses (Hayes and Rockwood, 2017)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued on next page)
<table>
<thead>
<tr>
<th>Authors</th>
<th>Sample / recruitment strategy</th>
<th>Comparison groups</th>
<th>Number of participants per group (% females)</th>
<th>Mean age (SD)</th>
<th>Study dropout</th>
<th>Mean number of modules completed / time spent</th>
<th>Percentages of participants completing all modules</th>
<th>Modules (no.) / duration in weeks*</th>
<th>Post-treatment; follow-up (s) in weeks*</th>
<th>Depression symptom severity instruments</th>
<th>Potential mediators</th>
<th>Mediator instruments (measurement points in weeks*)</th>
<th>Statistical method for mediation analyses (authors)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kauer et al.</td>
<td>Young people with mild or severe emotional or mental health concerns / GPs</td>
<td>Monitoring of 8 areas of functioning vs. attention control group</td>
<td>69 (3.2)</td>
<td>26%</td>
<td>3.3 entries per day / 17.7 days completed</td>
<td>56% (minimum dose)</td>
<td>−/2–4</td>
<td>2–4, 6</td>
<td>DASS</td>
<td>↑ Emotional self-awareness</td>
<td>ESA Scale (0, 2–4, 6)</td>
<td>Parallel process latent growth curve modelling in Mplus (Cheong et al., 2003; Muthén &amp; Muthén, 1998-2012)</td>
<td></td>
</tr>
<tr>
<td>Levin et al.</td>
<td>College students in distress / online, institutions, flyers</td>
<td>Web-based ACT vs. WLC</td>
<td>40 (2.73)</td>
<td>22%</td>
<td>–</td>
<td>55%</td>
<td>6/4</td>
<td>4; –</td>
<td>CCAPS-34</td>
<td>↑ Mindful acceptance ↓ Obstruction to valued living</td>
<td>VQ (0, 4)</td>
<td>Cross product of coefficient test in Mplus (Muthén &amp; Muthén, 1998-2012)</td>
<td></td>
</tr>
<tr>
<td>Ma et al.</td>
<td>Adults aiming to reduce stress / web-based adds</td>
<td>Group mindfulness-based intervention vs. self-directed mindfulness-based intervention vs. discussion group vs. WLC</td>
<td>48 (7.98)</td>
<td>60%</td>
<td>–</td>
<td>–</td>
<td>8/8</td>
<td>8; –</td>
<td>SDS</td>
<td>↑ Emotion regulation</td>
<td>DERS (0, 8)</td>
<td>PROCESS (Preacher &amp; Hayes, 2008)</td>
<td></td>
</tr>
<tr>
<td>Morgan et al.</td>
<td>Individuals with subthreshold depression / web-based sources (adds etc.)</td>
<td>E-mails promoting self-help strategies vs. non-directive depression information vs. discussion group vs. WLC</td>
<td>568 (13.8)</td>
<td>36.3</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>12/6</td>
<td>6; –</td>
<td>PHQ-9</td>
<td>↑ Use of self-help strategies</td>
<td>Frequency of strategy use (5-category scale; 0, 6)</td>
<td>MEDIATE (Preacher &amp; Hayes, 2008)</td>
</tr>
<tr>
<td>Morris et al.</td>
<td>Young adults / universities, online, social media</td>
<td>Web-based cognitive reappraisal program vs. online expressive writing</td>
<td>135 (5.3)</td>
<td>37%</td>
<td>21 log-ins / 9 min 18 s per session (1013.28 words)</td>
<td>–</td>
<td>–/3</td>
<td>3; –</td>
<td>CES-D</td>
<td>↑ Reappraisal</td>
<td>ERQ (0,3)</td>
<td>SPSS macro (Preacher &amp; Hayes, 2008)</td>
<td></td>
</tr>
<tr>
<td>Querstret et al.</td>
<td>Community / institutions, online, social media</td>
<td>Online mindfulness intervention vs. WLC</td>
<td>60 (10.45)</td>
<td>31%</td>
<td>–</td>
<td>75% (completed intervention in &lt;12 weeks)</td>
<td>–/4</td>
<td>4; 13, 26</td>
<td>PHQ-9</td>
<td>↑ Non-judging ↑ Describing (n.s.) ↑ Acting with awareness (n.s.)</td>
<td>FFMQ-SF (0, 4)</td>
<td>PROCESS (Hayes, 2013)</td>
<td></td>
</tr>
</tbody>
</table>
### Table 1 (continued)

<table>
<thead>
<tr>
<th>Authors (Year); Country</th>
<th>Sample / recruitment strategy</th>
<th>Comparison groups</th>
<th>Number of participants per group (% females)</th>
<th>Mean age (SD)</th>
<th>Study dropout*</th>
<th>Mean number of modules completed / time spent</th>
<th>Percentages of participants completing all modules</th>
<th>Modules (no.) / duration in weeks*</th>
<th>Post-treatment; follow-up (s) in weeks*</th>
<th>Depression symptom severity instruments</th>
<th>Potential mediators</th>
<th>Mediator instruments (measurement points in weeks*)</th>
<th>Statistical method for mediation analyses (authors)</th>
</tr>
</thead>
<tbody>
<tr>
<td>van Luenen et al. (2019); NL</td>
<td>People with HIV and depressive symptoms / HIV treatment centers</td>
<td>iCBT intervention vs. attention-only WLC</td>
<td>97 (46.3) (10.63)</td>
<td>19%</td>
<td>–</td>
<td>–</td>
<td>8 / 8</td>
<td>8; –</td>
<td>PHQ-2</td>
<td>↑ Activation</td>
<td>BADS (0, 1, 3, 5, 8)</td>
<td>Self-designed item</td>
<td>Multilevel SEM, exploratory analyses, and autoregressive latent trajectory analysis (Bollen &amp; Curran, 2004; Garland et al., 2015; Preacher et al., 2010)</td>
</tr>
<tr>
<td>Viskovich and Pakenham (2018); AUS</td>
<td>University students / institutions, social media, flyers</td>
<td>ACT with recommendation vs. ACT with no recommendation vs. ACT with stepped approach and enforced gaps</td>
<td>40 26.3 (7.96)</td>
<td>61%</td>
<td>–</td>
<td>27%</td>
<td>4 / 4</td>
<td>4; –</td>
<td>DASS-SF</td>
<td>↑ Acceptance</td>
<td>AAQ-II (0, 4)</td>
<td>MEMORE (Montoya &amp; Hayes, 2016)</td>
<td></td>
</tr>
<tr>
<td>Viskovich and Pakenham (2019); AUS</td>
<td>University students / social media, institutions, e-mail</td>
<td>Web-based ACT vs. WLC</td>
<td>596 26.9 (8.77)</td>
<td>42%</td>
<td>–</td>
<td>34% (started or completed module 4)</td>
<td>4 / 4</td>
<td>4; 12</td>
<td>DASS-SF</td>
<td>↑ Acceptance</td>
<td>AAQ-II (0, 4, 12)</td>
<td>PROCESS and MEMORE (Montoya &amp; Hayes, 2017; Preacher &amp; Hayes, 2004)</td>
<td></td>
</tr>
</tbody>
</table>

(continued on next page)
Table 1 (continued)

<table>
<thead>
<tr>
<th>Authors (Year); Country</th>
<th>Sample / recruitment strategy</th>
<th>Comparison groups</th>
<th>Number of participants per group (% females)</th>
<th>Mean age (SD)</th>
<th>Study dropout*</th>
<th>Mean number of modules completed / time spent</th>
<th>Percentages of participants completing all modules</th>
<th>Modules (no.) / duration in weeks*</th>
<th>Post-treatment; follow-up (s) in weeks*</th>
<th>Depression symptom severity instruments</th>
<th>Potential mediators</th>
<th>Mediator instruments (measurement points in weeks*)</th>
<th>Statistical method for mediation analysis (authors)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Willems et al. (2017); NL</td>
<td>Cancer survivors / hospitals</td>
<td>Stand-alone online intervention vs WLC</td>
<td>265</td>
<td>56.3 (11.17)</td>
<td>18%</td>
<td>2.22 modules / -</td>
<td>8.5% ≥ 5 modules</td>
<td>8 / -</td>
<td>&lt; 26</td>
<td>HADS-D</td>
<td>↑ Perceived personal control over cancer</td>
<td>↓ Problem-solving skills (n. s.)</td>
<td>PROCESS (Mackinnon, 2008; Hayes, 2013; Preacher &amp; Hayes, 2008)</td>
</tr>
</tbody>
</table>

Note. AAQ-II = Acceptance and Action Questionnaire-II (Bond et al., 2011); ACT = Acceptance and Commitment Therapy; AST-D = Ambiguous Scenarios Test – Depression (Berna, Lang, Goodwin, & Holmes, 2011); AST-D-II = Ambiguous Scenarios Test for depression-related bias-II (Rohrbacher & Reinecke, 2014); ATQ = Automatic Thoughts Questionnaire (Hollon & Kendall, 1980); AUS = Australia; BADS = Behavioral Activation for Depression Scale (Kanter, Mulick, Busch, Berlin, & Martell, 2007); BADS-SF = Behavioral Activation for Depression Scale – Short Form (Manos, Kanter, & Luo, 2011); BDI-II = Becks Depression Inventory, revised (Rebeck et al., 1996); BNPSF = Psychological Need Satisfaction and Frustration Scale (Chen et al., 2014); CBM = Cognitive Bias Modification; CBM-I = imagery-based Cognitive-Bias Modification; CBAPS-34 = Counseling Center Assessment of Psychological Symptoms-34 (Locke et al., 2012); CCTS-SR = Competencies of Cognitive Therapy Scale – Self-Report (Strunk, Hollars, Adler, Goldstein, & Braun, 2014); CERQ-short = Cognitive Emotion Regulation Questionnaire short version (Garnefski & Kraaij, 2006); CES-D = Center of Epidemiological Studies Depression scale (Radloff, 1977); CED-10 = Centre for Epidemiological Studies Depression Scale – Shortened Version (Andresen et al., 1994); Cognitive Fusion Questionnaire (Gillanders et al., 2014); CH = Confederalia Helvetica/Switzerland; CN = People’s Republic of China; CSES = Coping Self-Efficacy Scale (Chesney et al., 2006); DAS = Dysfunctional Attitudes Scale (Weismann, 1979); DASS-SF = Depression Anxiety Stress Scale (Short Form; Lovibond & Lovibond, 1995); DESR = Difficulties in Emotion Regulation Scale (Gratz & Roemer, 2004); Engaged Living Scale (Trompetter et al., 2013); EQ-R = Emotion Regulation Questionnaire (Gross & John, 2003); ESA Scale = adapted from 20-item Self-reflection and Insight Scale (Grant, Franklin, & Langford, 2002); 10-item Ruminative Response Scale (Treynor, Gonzalez, & Nolen-Hoeksema, 2003) and 12-item Meta-Evaluation Scale (Mayer & Stevens, 1994); ESAS-R = Emotional Self-Awareness Scale – Revised (Bakker & Rickard, 2018); FFMQ-SF = Five Facet Mindfulness Questionnaire – Short Form (Baer et al., 2006); FATS = Frequency of Actions and Thoughts Scale (Terides et al., 2016); GAD = Generalized Anxiety Disorder; GAD-7 = generalized anxiety disorder screener (Lowe et al., 2008); GP = General Practitioner; (HADS-A = Hospital Anxiety and Depression scale-Agency (Zigmond & Snaith, 1983); GDGRS = Goal Disengagement and Goal Re-engagement Scale (Wrooch et al., 2003); HADS-D = Hospital Anxiety and Depression scale-Depression (Zigmond & Snaith, 1983); HAMD = Hamilton Rating Scale for Depression (Hamilton, 1960); iCBT = internet-based Cognitive-Behavioral Therapy; IPQ-R = Illness perception questionnaire - revised (Moss-Morris et al., 2002); Mastery Scale (Pearlin & Schooler, 1978); MDD = Major Depressive Disorder; Mental Health Literacy Questionnaire (Bakker & Rickard, 2017; 2018); MHSES = Mental Health Self-efficacy Scale (Clarke et al., 2014); MI = motivational interviewing; Mindful Attention Awareness Scale (Brown & Ryan, 2003); MINI = Mini International Neuropsychiatric Interview (Sheehan et al., 1998); NL = the Netherlands; Northwestern Ego-Integrity scale (Janis et al., unpublished); PANAS = Positive and Negative Affect Schedule (Watson et al., 1988); PASPO = Perceived Ability to Bavor Positive Outcomes (Bryant, 1989); PBRS = Positive Beliefs about Rumination Scale-Adapted Version (Watkins & Moulds, 2005); PHLMS = Philadelphia Mindfulness Scale (Cardaciotto et al., 2008); PHQ-9 = Patient Health Questionnaire depression scale (Kroenke et al., 2001); PHQ-9 = first two questions of PHQ-9; Personal Values Questionnaire II (Blackledge, Ciarrochi, & Bailey, n.d.); PPI = Positive Psychology Intervention; PST = Problem-solving therapy; PWSQ = Penn State Worry Questionnaire (Meyer et al., 1990); PTQ = Perseverative Thinking Questionnaire (Ehring et al., 2011); RNT = Repetitive negative thinking; RTQ = Repetitive Thinking Questionnaire (McEvoy et al., 2010); RT = Recognition Test (based on Mathews & Mackintosh, 2000); Ruminative Response Scale (Roelofs, Muris, Huebben, & Aarts, 2006); Scrambled Sentences Test (Rude et al., 2002; Wenzlaff & Bates, 2000; Wenzlaff & Bates, 1994; Wenzlaff & Bates, 2000); TAU = treatment as usual; UK = United Kingdom; US = United States of America; VQ = Valuing Questionnaire (Smout et al., 2014); WHODAS II = World Health Organization Disability Assessment Schedule II (WHO, 2000); WLC = Waiting List Control.

* Information were also retrieved from the initial RCT (Proudfoot et al., 2013;); † Information were also retrieved from the initial RCT (Pots et al., 2016); ‡ Results were reported for multiple mediation models and were consistent in both trials in Seeley et al. (2018). The two trials in Seeley et al. (2018) were treated as separate investigations of mediators, as they were conducted with different samples and different outcome measurements for depression, although some mediator instruments were deployed in both trials; ‡ In the study of Terides et al. (2017), mediation analyses were conducted with complete outcome data only, whereas information in regard to complete data was provided in Table 5 in Terides et al. (2017) indicate that rewarding behavior was evaluated as a significant mediator, the authors discuss these significant findings in the discussion section conversely; ‡ Information were also retrieved from the initial RCT (van der Zanden et al., 2012); ‡ Study drop out was calculated as follows: participants completing posttreatment assessment divided by randomized participants across groups. These calculations are not necessarily in accordance with the study drop-out rates as reported in primary studies; ‡ Conversions were conducted as follows (rounded off to whole numbers): 1 day = 0.14 weeks, 1 month = 4.35 weeks, 1 year = 52.18 weeks; ‡** Analyzed sample; *** Participants in Newby et al. (2014) received mid-treatment measurement prior to commencing lesson 4, which was available 4 weeks after baseline; — Not reported/applicable; ↑ Increasing scores in the mediator construct were hypothesized to be associated with depressive symptom improvement; ↓ Decreasing scores in the mediator constructs were hypothesized to be associated with depressive symptom improvement; (n. s.) not significant.
examined in one study each (3.8%) were Cognitive Bias Modification (CBM), Problem-Solving Therapy (PST), interpretation training, life-review, mental health self-efficacy, monitoring, and self-help strategies for behavior change. Three studies (11.5%) included a combined treatment of Internet-based CBT (iCBT) augmented with either CBM, PST or monitoring. Just over half of the studies evaluated unguided IMIs (k = 1; 53.8%), whereas in 12 studies (46.2%) some type of guidance was evaluated. The number of modules in IMIs ranged from 4 to 13 (M = 7.41; SD = 2.59). Study drop-out rates ranged from 1% to 61%.

Wait-list control groups (WLC) were deployed in ten studies (38.5%), while further eight studies (30.8%) added one (or more) other control condition(s) to the WLC. Other comparison groups comprised different active control conditions like an alternative IMI (k = 6; 23.1%), online writing (k = 4; 15.4%), attention control (k = 4; 15.4%), face-to-face psychotherapy (k = 1; 3.8%), motivational interviewing and discussion group (k = 1; 3.8%), non-directive information (k = 1; 3.8%), treatment-as-usual (TAU; k = 1; 3.8%) or referral to other treatment (k = 1; 3.8%) or the Self-Rating Depression Scale (SDS; Zung, 1965; k = 1; 3.8%). As such, with 96.2% (k = 25) the vast majority of included studies used self-report measures to assess depression symptom severity and only one study resorted to a clinician-administered scale.

Mediation analyses were most often conducted by means of bootstrapping procedures (Hayes, 2012; Montoya & Hayes, 2017; Preacher & Hayes, 2008; k = 15; 57.7%) or structural equation modelling (Cheong, MacKinnon, & Khoo, 2003; Iacobucci, Saldanha, & Deng, 2007; MacKinnon, 2017; Muthén & Muthén, 1998-2012; Russell, Kahn, Spoth, & Altmaier, 1998; k = 8; 30.8%). Other approaches to mediation analyses deployed in one study each (3.8%) were linear mixed modelling, mixed models repeated measures and parallel multiple mediation analyses, all within the framework of the causal steps approach (Baron & Kenny, 1986).

### 3.2. Mediators

Overall, the included studies examined 64 distinct potential mediators (Table 1), of which the majority (m = 42; 65.6%) was found to be significant (Table A3 in the Appendix). Intervening variables with mixed findings (evaluated both as significant and non-significant) are indicated with the symbol "~", Mediators that were examined in studies that evaluated the temporality criterion are indexed with the letter "T", and mediators that were directly manipulated in an experiment with the letter "e". Due to the wide spectrum and limited overlap of single mediator constructs (i.e., small number of studies per mediator), meta-analytical pooling would not have been feasible.

#### 3.2.1. Cognitive mediators

Cognitive mediators represent the largest group of investigated mediators (m = 28 mediators; 43.7%). In studies with clinical samples the
following mediators were established as significant: cognitive skills, positive beliefs, functional metacognitive beliefs, psychological flexibility, mastery/perceived control, as well as changes in interpretation bias, rumination, repetitive negative thinking, negative thoughts, dysfunctional attitudes and worrying. In non-clinical samples the cognitive variables coping self-efficacy, valued living, reappraisal, cognitive fusion, perceived control and changes in repetitive negative thinking were established as significant. A total of seven cognitive mediators were found to be non-significant (Table 1).

3.2.2. Behavioral mediators

Findings on behavioral activation (m = 4; 6.3%) were mixed. Behavioral activation was evaluated twice as significant and once as non-significant in studies with clinical samples. In one study resorting to a non-clinical sample, behavioral activation significantly mediated intervention effects (Table 1).

3.2.3. Emotional and affective mediators

Emotional and affective mediators were examined in non-clinical samples only (m = 4; 6.3%). Emotion regulation was evaluated as significant in one study; whereas the findings on emotional self-awareness were mixed, and on positive affect non-significant (Table 1).

3.2.4. Skills as mediators

A total of 18 different skills (m = 18; 28.1%) were examined and these intervening variables comprise the second largest group of mediators. In clinical samples the skills of mindfulness (observing and non-judging of inner experience), CBT skills usage and changes in negative problem orientation were found to be significant; whereas four skills were evaluated as non-significant (Table 1). In non-clinical samples, acceptance, self-help strategies, mindfulness (non-judging and acting with awareness), present moment awareness and problem-solving skills were significant intervening variables; further three skills were found to be non-significant (Table 1).

3.2.5. Other mediating variables

Mediator variables that were not clearly classifiable into one of the aforementioned categories represent the third largest group of mediators (m = 10; 15.6%). In studies with clinical samples, conceptually differing mediators like ego integrity, savoring of pleasant events, disability status, or anxiety were found to be significant. In non-clinical samples the intervening variables sense of autonomy, goal re-engagement and education values were evaluated as significant. Only mental health literacy was evaluated as non-significant in this group of mediators (Table 1).

3.3. Risk of bias and quality assessment for process research

The results of the risk of bias assessment are detailed in Fig. A.1 in the Appendix. The domains with the lowest risk of bias across studies were “selective reporting” and “other bias”; whereas the criterion least met was “blinding of participants and personnel”. Of note, masking of participants and therapists is somewhat impossible to accomplish in conventional psychotherapy research, but might be generally possible in IMIs, especially in unguided interventions with two active comparison conditions (Domhardt et al., 2020). Overall, the quality of included studies can be regarded as rather high, with 20 studies (76.9%) having low or moderate risk of bias.

The results of the quality assessment for process research (i.e., evaluation of the causal specificity of studies investigating hypothesized mechanisms) are detailed for each study in Table 2. Obviously, as the criteria CBT and control group were simultaneously inclusion criteria of this review, all studies adequately adhered to these standards. The next best fulfilled criteria were theoretical foundation (k = 25; 96.2% of studies) and sample size per group (k = 21; 80.8%). Multiple mediator variables were measured in 16 studies (61.5%). Notably, only five studies (19.2%) assessed the mediator and outcome variables at least on three measurement points (temporality) and only one single study realized a direct manipulation of a mediator in an experimental design (3.8%). In fact, Hirsch et al. (2018) investigated the effects of experimentally inducing positive interpretations prior to cognitive bias modification training (CBM) by means of a priming task. Thereby the authors found that changes in interpretation bias at post-treatment partially mediated the effects of CBM on worry, rumination and anxiety at follow-up in participants with symptoms of depression or anxiety; whereas the intervening variable effect of interpretation bias on depression symptoms missed statistical significance marginally in this experiment (Hirsch et al., 2018).

Altogether, with 16 studies fulfilling five or more (out of seven) quality criteria, the overall study quality based on this rating system can be considered as rather good. An overview of the percentages of studies meeting each quality criterion is outlined in Table A.4, the number of quality criteria met by each study is illustrated in Fig. A.2 across studies (Appendix).

4. Discussion

To our knowledge, this is the first comprehensive systematic review that evaluated mediation studies in various forms of IMIs for depression. In this review, a total of 26 mediation studies with 6820 participants were included, which investigated 64 intervening variables in IMIs based on 11 different theoretical foundations. Thereby, cognitive variables were found to be the largest group of significant mediators. The next largest groups comprised a range of different behavioral and emotional mediators, as well as mindfulness and other skills. Eligible mediation studies were assessed in terms of design characteristics suitable for the convergence of causality, indicating a rather high overall study quality, with still substantial room for improvement on certain aspects within this branch of process research.

As the majority of IMIs were based on CBT-principles, it is not surprising that the majority of examined mediators are somehow related to the conceptualizations of this therapeutic orientation (Beck et al., 1979). Thereby, the cognitive and behavioral mediators examined in IMIs largely correspond to those intervening variables summarized in systematic reviews of face-to-face therapies for depression in general (Lemmens et al., 2016) and cognitive (behavioral) therapy in particular (Cristea et al., 2015; Garratt, Ingram, Rand, & Sawalani, 2007; Hoffmann, Asmundson, & Beck, 2013; Kazantzis et al., 2018; Lorenzo-Luaces et al., 2015). These findings indicate that certain cognitive and behavioral processes are consistent in psychotherapeutic interventions for depression, regardless if they are implemented face-to-face or remotely via the Internet or mobile devices. Furthermore, comparisions for cognitive mediators suggest, that these intervening variables might not be specific to CBT, but may well depict common underlying cognitive processes in psychotherapeutic interventions for depression, regardless if they are implemented face-to-face or remotely via the Internet or mobile devices. Yet, more information on specific cognitive processes and mediators not derived from CBT are needed to substantiate these findings, as knowledge from other evidence-based face-to-face psychotherapies for depression (e.g., psychodynamic, client-centered or interpersonal therapy) are largely pending in IMIs so far.

The recency of technology-delivered interventions for depression might have further influenced the findings that processes related to ACT and mindfulness are somewhat more frequently and disproportionately examined in IMIs, when compared to the evidence in face-to-face therapies (Lemmens et al., 2016). The mediators acceptance and mindfulness seem to be specific to these relatively novel therapeutic approaches, although these early indications need to be further substantiated by research directly addressing the specificity of these intervening variables. Moreover, apart from mobile-based monitoring of emotions and areas of functioning (i.e., “tracking”; Bakker et al., 2018; Kauer et al., 2012), no “digital factors” (Domhardt et al., minor revision), unique
mediators or specific mechanisms to digitalized interventions for depression were identified to this point. With the concept of “digital factors”, we refer to factors that are either common to all, or only specific to certain digital health interventions – but are essentially not a component of conventional psychotherapies on-site (Domhardt et al., minor revision). Conceivable “common digital factors” are for example the technological implementation, digital working alliance or positive outcome expectations about the digital device and intervention. Presumed “digital specific factors” inherent solely to certain interventions may be for instance mobile and passive sensing, continuous automated feedback with smartphones or wearables, as well as real-time triggered interventions based on moment-to-moment analyses of human behaviors in everyday life. Upcoming research efforts must clarify, which of these digital factors are indeed active ingredients of technology-delivered interventions, and if they initiate the same or different mechanisms of change when compared to conventional face-to-face psychotherapies. These investigations may also contribute to find effective ways how to best increase adherence rates and the engagement of patients to digital interventions (Yardley et al., 2016), which are especially low when IMIs are unguided and transferred from controlled research settings into routine health care (Domhardt et al., 2019; Graham, Lattie, & Mohr, 2019).

**Table 3** Recommendations for intervention development, clinical practice and future research.

**Intervention development**
- Studies on mediators in mobile-based interventions are scarce; these interventions need to be evaluated in RCTs and their specific processes examined in detail
- Use additional approaches to identify the active ingredients and change mechanisms of IMIs, e.g., component studies, multiple optimization strategy frameworks using factorial designs, parametric and catalytic RCTs
- Take recent expansions of psychotherapy research into account, since transdiagnostic protocols and process-oriented interventions might further contribute to the scalability of IMIs for depression and common mental disorders
- Further advancements of theories on working mechanisms can inform intervention development

**Clinical practice**
- Secure that evidence-based components are implemented, which induce established mediators and mechanisms of change essential for improved outcomes (see Table 1 and Table A.3 for (non-)significant mediators)
- Monitor processes in routine clinical care, also those related to possible negative effects and deterioration rates
- Resort to multi-modal assessments (cognitive, behavioral, emotional/affective and/or psychological/biological variables) from multiple sources (e.g., self-report, clinician ratings, passive data), with well-established, but also novel mediator instruments with sound psychometric properties
- Digital phenotyping, i.e., active and passive data collection with smartphones in situ, and ecological momentary assessment studies might provide novel insights on (digital) factors and change mechanisms in routine care and daily lives of patients

**Future research**
- Adhere to methodological recommendations for process research, especially in regard to the establishment of temporality and direct manipulation of mediators (see Table 2 and Table A.3 for details)
- Investigate presumed novel digital factors and specific mechanisms in IMIs for depression, and mediating factors not researched so far (e.g., perceived support, distancing, hopefulness, empowerment, attributional style)
- Examine processes associated with negative effects and deterioration rates
- Stick to established methodological considerations, and evaluate multiple potential mediators that can falsify rivaling theories
- Ponder on the timing of measurements (e.g., early and sudden gains might necessitate measurements early in treatment) and implement fine-grained temporal designs; ideally a session-by-session basis, and also in daily life between sessions
- Analyze moderators and subgroups, as well as trajectories of change in individual patients
- Use recent statistical methods for mediation analysis (e.g., time lagged mediation analysis or latent growth curve modelling) and evaluate the specific influence of each mediator, as well as their interactions and associations with specific and common factors
- Extend the evidence-base by means of the replication of mediation studies, but also with original component studies, in order to identify the active ingredients of IMIs for depression

Presumably the particular dynamics of Internet intervention research will take up these gaps, and apprehend to recent developments and considerations of psychotherapy research, like the tendency toward a process orientation (Hayes et al., 2019; Hofmann, & Ciarrochi, 2020; Hofmann & Hayes, 2019) or the expansions of transdiagnostic protocols to common mental disorders as well (Steele et al., 2018; Weisel et al., 2019). These current developments might also further contribute to the potential of IMIs to expand and augment existing mental health care services (Domhardt et al., 2019; Ebert et al., 2017), by means of scalable self-help interventions (Buntrock et al., 2016; Karyotaki et al., 2017), advanced models of stepped care (Domhardt & Baumeister, 2018; van Straten, Hill, Richards, & Cuijpers, 2015) and blended therapy approaches (Baumeister, Gräße, Ebert, & Krämer, 2018; Domhardt, Steubi, & Baumeister, 2018; Erbe, Eichert, Ripser, & Ebert, 2017), as well as by providing tailored interventions for underserved populations (Andersson et al., 2019; Domhardt, Steubi, & Baumeister, 2018).

Investigations of transdiagnostic, process-based and tailored interventions in future mediation studies are of particular importance, to comprehensively assess the actual capability of IMIs to extend mental health services within different models of care – for depressive and other common mental disorders alike.

The quality assessment of mediation studies in this review revealed that the majority of included studies adequately adhered to standards of psychotherapy research in general (RCT design, control group), and to some extent also for process research in particular (theoretical deduction of mediators a priori). Still, it is crucial that forthcoming studies stick to established constructs of psychotherapy process research, as outlined in the introduction and overcome persistent conceptual confusions, which ancillary obstructed the debate between proponents of common and specific factors (Mulder et al., 2017). Moreover, the quality assessment of mediation studies indicated some significant shortcomings in their aptness for the detection of change mechanisms and the determination of causality. Most importantly, only a minority of studies (19%) established the temporal ordering of changes in mediators and outcomes with at least three measurement points and were therefore principally able to determine the directionality and shape of change. Yet, the criterion of temporality is considered to be of paramount importance (Kazdin, 2007, 2009), and the establishment of the timeline that changes in the mediator occur before changes in the outcome, represents the Achilles’ heel of process research in face-to-face therapies (Kazdin & Noack, 2003) and in IMIs for depression alike. Furthermore, direct causal conclusions are largely precluded also because only one study directly manipulated a mediator variable in an experimental design so far (Hirsch et al., 2018). On the contrary, it is a remarkable strength of the majority of included mediation studies that they were capable to implement relatively large sample sizes per condition, and tested multiple concurrent mediators in order to falsify rivaling mechanisms. The comparatively large sample sizes on average (M = 262; SD = 243) might have also contributed to the finding that studies using an active comparison group similarly often established significance for examined mediators as studies deploying passive control conditions (59.1% vs. 65.9%). In contrast, the differences in significance on mediators in face-to-face psychotherapies (Lemmens et al., 2016) were markedly more pronounced in studies deploying active controls compared to passive controls (41.3% vs. 69.9%), presumably also partly due to smaller samples sizes (M = 173; SD = 145) in conventional psychotherapy process research. Noteworthy, the quality of eligible studies did not differ between studies with clinical (61/91 sum score of quality criteria met over studies) and non-clinical samples (60/91) in the current review at hand. Overall, the findings of the quality assessment suggest a rather stable quality in this novel branch of psychotherapy process research in the last five years, also when compared to mediation studies in face-to-face psychotherapies for depression (Lemmens et al., 2016).

Two of four benchmark studies (i.e., fulfilling all quality criteria apart from direct manipulation of mediators; Table 2) illustrate the...
apitude of IMIs to reveal novel insights into presumed change mechanisms. For example, van Luenen et al. (2019) revealed intervening variable effects for behavioral activation and goal re-engagement in Internet-based CBT for 188 participants with HIV on depressive symptoms when compared to WLC. In deploying bivariate autoregressive latent trajectory analysis (Bollen & Curran, 2004), the authors were able to show that the intervening variable effects were stronger for behavioral activation than for goal re-engagement, thereby providing information on the differential magnitude of mediating effects. In another benchmark study, Warmerdam et al. (2010) found in an RCT with 263 clinically depressed participants, that the positive effects of Internet-based CBT and PST on symptom severity were transmitted by the examined mediators in both interventions alike. Thereewith, the study disclosed initial evidence that both interventions might operate with shared processes (i.e., changes in dysfunctional attitudes, worrying, negative problem orientation and perceived control), although mediators were conceptually related only to one of the aforementioned theoretical orientations. Aside these findings on (non-)specificity the authors were not able to establish the temporality criterion, even though three measurements of mediators and outcomes were implemented. This was because most symptom change happened early in treatment, not allowing for the detection of later changes in mediators and outcomes after mid-treatment (week five), impeding conclusions on the direction of change. Notably, only one study found initial support for a dose–response relationship between the intervention and the mediator, and explicitly addressed the gradient criterion. In fact, Heckendorf et al. (2019) found that increased usage of the examined gratitude app with daily exercises correlated with the mediator repetitive negative thinking in a community sample of 260 adults compared to WLC on depression and anxiety symptom severity.

4.1. Limitations and strengths

This systematic review offers several strengths, including adherence to a pre-registered study protocol, comprehensive incorporation of IMIs with various theoretical foundations and treatment formats, as well as the relatively high overall quality of eligible studies. However, some important limitations need to be considered when interpreting the findings of the present study. First, the generalizability of this review might be limited, since the vast majority of mediation studies were conducted in western countries and resorted to female participants. Second, the majority of mediation studies used self-report instruments to assess depression outcomes and mediators alike; while this might represent on one side a constraint in terms of the validity and reliability of the assessment of depression symptom severity, the recourse to self-reports might be otherwise an advantage in regard to the detection of change mechanisms, as certain processes are only detectable by introspection of patients. Third, most mediation studies focused on a significance criterion when evaluating the intervening variable effects, but the emphasis on p-values might be misleading and the resort to interpretable effect sizes with estimates of accuracy more appropriate (Kraemer, 2019). Fourth, the concentration on statistical significance and the focus on English language publications might have also resulted in some degree of publication bias (Egger et al., 1997; Rothstein, Sutton, & Borenstein, 2005), whereas our systematic searches revealed no relevant non-English papers, and our searches in ICTRP indicated no unpublished registered mediation studies with non-significant findings to this point.

4.2. Future directions

Recommendations for future research on change mechanisms in IMIs for depression, intervention development and clinical practice are summarized in Table 3. Most importantly, forthcoming mediation studies should aim to amend the methodological limitations outlined in the quality assessment, in order to enable inferences with higher causal specificity. Furthermore, future studies should resort to more recent and elaborated approaches to statistical mediation analysis, like dynamic structural equation modelling (Grimm & Ram, 2018; Iacobucci et al., 2007), time lagged mediation analysis (Cole & Maxwell, 2003), latent growth curve modelling (Cheong et al., 2003; von Soest & Hagtvet, 2011) or dynamic network approaches (Hofmann et al., 2020), circumventing the limitations inherent to the causal steps approach, and capable to model the temporal ordering, as well as the actual shape of change. In these ways, the investigation of putative mediators derived from theoretical orientations other than CBT, and the establishment of mediators in mobile-based interventions are of value to enlarge the evidence base and clarify questions of (non-)specificity. Mediators not examined so far (e.g., biological variables, attributional style or social support; Domhardt, Münzer, Fegert, & Goldbeck, 2015), and mediators associated with presumed digital specific and digital common factors (Domhardt et al., minor revision), like empowerment or self-efficacy, might be of particular relevance for future intervention development, striving for IMIs with higher adherence rates and improved treatment outcomes. Complementary to these research efforts might be the application of the Multiphase Optimization Strategy (MOST; Collins, Murphy, & Strecher, 2007) and methodological advancements to analyze the differential efficacy of intervention components and mediators by innovative meta-analyses of individual participant data (e.g., Purukawa et al., 2019; Lin et al., 2019).

In addition, although mediation studies on IMIs for children and adolescents were not a decisive focus of the present review, our systematic searches revealed only two studies dedicated solely to this younger age group (Riley, Duke, Freeman, Hood, & Harris, 2015; Smith et al., 2015). This shortage of research is a substantial drawback, since empirical findings from adults with depressive disorders should not be conveyed on pediatric populations offhand; given substantial differences in treatment manuals and the consistently lower effect sizes of psychotherapies for depression in children and adolescents, irrespective if interventions are delivered face-to-face (Cuijpers et al., 2020) or technology-mediated (Domhardt, Ebert, & Baumeister, 2018; Ebert et al., 2015). Accurate evidence-based knowledge on the specific mechanisms of change in IMIs for depression in children and adolescents would be of tremendous importance, in order to inform intervention development and increase – the currently rather unsatisfying – effect sizes of psychotherapeutic programs for youth. This is even more relevant, as the utilization rates of conventional evidence-based psychotherapies are low in children and adolescents (Costello, He, Sampson, Kessler, & Merikangas, 2014; Merikangas et al., 2011), and novel digital interventions might be an auspicious opportunity to extend mental health care services, considering that technology-transmitted interventions might be particularly attractive, accessible and scalable for youth.

Finally, as all included studies in this review focused on depression symptom severity as (primary) outcome, upcoming studies should examine the change mechanisms associated with other relevant outcome domains from the patient’s perspective, like personal targets, quality of life or negative outcomes to be avoided (Cuijpers, 2019). So far, the relation of mediator variables to negative effects were not obtainable in eligible studies. While crucial evidence on the potential negative effects (Rozental, Boettcher, Andersson, Schmidt, & Carlbring, 2015) and deterioration rates in IMIs for depression are available elsewhere (Ebert et al., 2016; Karyotaki et al., 2018) – indicating no substantial differences between IMIs and face-to-face psychotherapies to date – upcoming studies should investigate the association of mediators to potential adverse outcomes carefully. Comprehensive evidence on these issues would contribute to the safety of technology-delivered interventions for depression, and inform the implementation of guided and unguided IMIs into routine clinical care.
5. Conclusion

Overall, this systematic review indicates that the well-established treatment effects of IMIs for depression are mediated by a range of primarily cognitive and emotional variables, and as well as several other behavioral and emotional variables, which are comparable to those mediators already identified in face-to-face psychotherapies. However, strong causal conclusions are not warranted to this point, because most often the central criterion of temporality is not established in mediation studies, and experimental studies with a direct manipulation of presumed mediators are largely pending. The findings and methodological recommendations of this review might guide future research, aiming to ascertain the active ingredients and mechanisms of change in technology-delivered interventions for depression, with greater clarity, scientific rigor and higher causal certainty.

Role of funding sources

This work is funded by the German Federal Ministry of Education and Research (BMBF; Grant Identification FKZ 01KG1802). The funding source BMBF had no role in the study design, collection, analysis or interpretation of the data, writing the manuscript, or the decision to submit this paper for publication.

Contributors

MD, PC and HB developed the design of the study. MD and LS conducted the systematic literature searches, extracted the data, and rated the risk of bias and methodological quality criteria of included studies. MD wrote the first draft of the manuscript. All authors (MD, LS, JB, CB, EK, DDE, PC, and HB) have contributed to the further writing, and have approved the final manuscript.

Declaration of Competing Interest

All authors declare that they have no conflicts of interest.

Acknowledgements

The authors wish to thank Helene Gelein and Roman von Rezori for their contribution to the search strings, and Sophie Engler for her assistance in editing tables and figures.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cpr.2020.101953.

References

Cuijpers, P. (2016). The future of psychotherapy research: Stop the waste and focus on causal conclusions are not warranted to this point, because most often the central criterion of temporality is not established in mediation studies, and experimental studies with a direct manipulation of presumed mediators are largely pending. The findings and methodological recommendations of this review might guide future research, aiming to ascertain the active ingredients and mechanisms of change in technology-delivered interventions for depression, with greater clarity, scientific rigor and higher causal certainty.
M. Domhardt et al.

Clinical Psychology Review 83 (2021) 101953


Matthias Domhardt, is a research associate and lecturer in the Department of Clinical Psychology and Psychotherapy at Ulm University, Germany. His main research interests include psychotherapy process research, digital mental health and pediatric psychology, as well as resilience and mental disorders specifically associated with stress. He has studied psychology at the Universities of Waikato (Hamilton, New Zealand) and Trier (Germany), and graduated as licensed child and adolescent psychotherapist at Ulm University, where he is currently completing his PhD.

Lena Steubi, is a research associate and PhD student in the Department of Clinical Psychology and Psychotherapy at Ulm University, Germany. Her research interests comprise Internet- and mobile-based interventions, mechanisms of change, and blended therapy approaches.

Johanna Boettcher, PhD, is Professor of Clinical Psychology and Psychotherapy at the Psychologische Hochschule Berlin, Germany. Her research focuses on Internet-based interventions for mental disorders, transdiagnostic therapies for affective disorders, intercultural therapy approaches, as well as mediators and mechanisms of change.

Claudia Buntrock, PhD, is a postdoctoral researcher at the Chair of Clinical Psychology and Psychotherapy at Friedrich-Alexander-University Erlangen-Nuremberg. Her research focuses on health economic evaluations and the prevention of depression through e-Mental-Health interventions.

Eirini Karyostaki, PhD, is Assistant Professor in the Department of Clinical Psychology at the Vrije Universiteit Amsterdam, the Netherlands, and Research Fellow in Global Health and Social Medicine at Harvard Medical School. Her research work is dedicated to the effectiveness of psychotherapy in treating depression and on Internet-based interventions, with a particular focus on systematic reviews with conventional and individual participant data meta-analyses.

David D. Ebert, PhD, is Associate Professor for Clinical Psychology at the Vrije Universiteit Amsterdam, the Netherlands, and Director of the WHO Collaborating Centre for Research and Dissemination of Psychological Interventions. His research focuses on randomized controlled trials and meta-analyses on prevention and psychological treatments of common mental disorders. His research work is primarily directed to the prevention of mental disorders, psychological treatments of depression and anxiety, and Internet-delivered treatments.

Pim Cuijpers, PhD, is Professor of Clinical Psychology at the Vrije Universiteit Amsterdam, the Netherlands, and Director of the WHO Collaborating Centre for Research and Dissemination of Psychological Interventions. His research focuses on randomized controlled trials and meta-analyses on prevention and psychological treatments of common mental disorders. His research work is primarily directed to the prevention of mental disorders, psychological treatments of depression and anxiety, and Internet-delivered treatments.

Harald Baumeister, PhD, is Professor and Head of the Department of Clinical Psychology and Psychotherapy and the associated psychotherapy outpatient clinic at Ulm University, Germany. His research focuses on depression research, psychotherapy research and technology-delivered diagnostics and interventions, with particular research interests in factors contributing to the differential efficacy of digital health interventions, ways for improving the acceptance and adherence of these novel psychotherapeutic interventions, as well as their respective working mechanisms.