Behaviour Research and Therapy 120 (2019) 103443

Contents lists available at ScienceDirect

Behaviour Research and Therapy

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

Dysfunctional cognition and self-efficacy as mediators of symptom change in exposure therapy for agoraphobia – Systematic review and meta-analysis

Christoph Breuninger[a,*], Brunna Tuschen-Caffier[a], Jennifer Svaldi[b]

[a] Department of Clinical Psychology and Psychotherapy, University of Freiburg, Freiburg, Germany
[b] Department of Clinical Psychology and Psychotherapy, University of Tübingen, Tübingen, Germany

ARTICLE INFO

Keywords:
Exposure therapy
Mediation
Dysfunctional cognition
Self-efficacy
Agoraphobia
CBT

ABSTRACT

Both dysfunctional and self-efficacy-related cognitions are theorized as etiological and maintaining factors in agoraphobia. Exposure therapy is an effective treatment and central component of CBT for agoraphobia, but the role of changes in these cognitions as a mechanism of action has not been established. The present review aims to evaluate (a) whether exposure without cognitive interventions elicits changes in cognitive variables and (b) whether cognitive changes mediate outcomes in exposure-based treatments. We searched PsycInfo and PubMed for studies on agoraphobia (with or without panic disorder) and exposure as a treatment component. Fifteen articles with 29 relevant study arms (N = 921) were identified for a meta-analysis of cognitive changes after exposure. Seventeen articles (N = 1881) were included in a systematic narrative review of cognitive mediation. A random effects model revealed a large effect of cognitive improvement after pure exposure treatments, d = 1.02 (95% CI 0.81–1.23). The systematic review mostly supported changes in cognition as mediators of symptom change. Improved study designs and statistical methods in future mediation studies are needed to strengthen causal interpretation. Cognitive change is a probable mechanism of action in exposure therapy, especially change in self-efficacy. The present review suggests novel ways in which cognitive interventions can augment exposure therapy.

Agoraphobia is a mental disorder characterized by fear and avoidance of diverse situations such as using public transportation, being in enclosed spaces or standing in line in shops (DSM-5; American Psychiatric Association (APA, 2013). Research on agoraphobia is complicated by its changing diagnostic status. Since DSM–III–R and up until DSM-IV-TR, agoraphobia had been considered a complication or residual of panic disorder (Wittchen, Gloster, Beesdo-Baum, Fava, & Craske, 2010). During the development of DSM-5, Wittchen et al. (2010) argued convincingly that agoraphobia can result from different pathogenic pathways and is best conceptualized as a construct separate from panic disorder. Thus, while thoughts that help would be unavailable and/or escape difficult are characteristic across agoraphobic individuals and situations (APA, 2013), individuals have different reasons to fear that help or escape will become necessary. These include suffering from panic attacks, panic-like symptoms, other incapacitating symptoms, and more idiosyncratic situational fears (Wittchen, Reed, & Kessler, 1998). Consequently, agoraphobia is now (DSM-5) classified as an independent disorder, which is coded as comorbid with panic disorder when both criteria are met. Lifetime prevalence (based on DSM-IV classification) is estimated at 1.4% for agoraphobia alone (Kessler, Berglund, et al., 2005) and an additional 1.1% for panic disorder with agoraphobia (Kessler et al., 2006). Prevalence in women is found to be approximately twice as high as in men (Goodwin et al., 2005). Agoraphobia is of considerable clinical severity itself (Kessler, Chiu, Demler, Merikangas, & Walters, 2005) and adds substantial severity and functional impairment when comorbid with panic disorder (Kessler et al., 2006). In a systematic review of moderators of improvement in CBT for panic disorder and agoraphobia (Porter & Chambless, 2015), agoraphobic avoidance emerged as the strongest predictor of poor treatment response, thus further underlining the importance of effectively treating agoraphobia.

Cognitive behavioral therapy (CBT) is supported by numerous meta-analyses as an efficacious treatment for panic disorder with or without agoraphobia (Gould, Ott, & Pollack, 1995; Mitte, 2005; Sánchez-Meca, Rosa-Alcázar, Marín-Martínez, & Gómez-Conesa, 2010). Exposure to feared situations *in vivo* is considered a crucial technique when agoraphobia is the main diagnosis (Gloster et al., 2011; Ruhmland & Margraf, 2001) and consequently forms an integral part of established treatment protocols (e.g., Craske & Barlow, 2007). A recent meta-analysis (Sánchez-Meca et al., 2010) reports very large effect sizes for

* Corresponding author. Department of Clinical Psychology and Psychotherapy, University of Freiburg, Engelbergerstr. 41, 79106, Freiburg, Germany.
E-mail address: christoph.breuninger@psychologie.uni-freiburg.de (C. Breuninger).

https://doi.org/10.1016/j.brat.2019.103443
Received 26 September 2018; Received in revised form 2 July 2019; Accepted 26 July 2019
Available online 26 July 2019
0005-7967/ © 2019 Elsevier Ltd. All rights reserved.

cognitive-behavioral interventions including exposure ($d = 1.29$), much larger than for cognitive interventions alone ($d = 0.34$). Despite these impressive average improvements, only about 50% of patients achieve full remission even after comprehensive CBT (Loerinc et al., 2015). Improving the existing interventions for the substantial number of patients who currently do not improve enough requires a deeper understanding of the mechanisms of action of exposure therapy. From these mechanisms, we could derive improved ways of implementing exposure as well as learn which cognitive techniques are most suitable to augment them.

Investigating these mechanisms goes beyond demonstrating causality (as is achieved in RCTs) and aims to explain the process of *how* a cause (such as exposure exercises) brings about its effect (e.g., symptom improvement). Mediators are variables that can explain (some of) the covariance between cause and effect in statistical models and are one important step in identifying mechanisms of action (together with aspects of research design and theoretical considerations; Kazdin, 2007). Mediators should not be confused with moderators, which refer to variables that change the size or direction of effects (Baron & Kenny, 1986; e.g., agoraphobia reducing the efficacy of CBT for panic disorder as mentioned above,; Porter & Chambless, 2015). For exposure therapy, emotional processing theory (EPT; Foa & Kozak, 1986) provided a theoretical explanation that has high face validity especially in the context of panic and agoraphobia because it stresses the importance of the psychophysiological fear response during exposure, and led many practitioners to focus on initial fear activation and habituation of fear as indicators of therapeutic change. In contrast, a review by Craske et al. (2008) concluded that neither activation nor habituation of fear are empirically supported as mediators of therapeutic outcomes after exposure for anxiety disorders. A recent meta-analysis of exposure treatments across anxiety disorders (Rupp, Doebler, Ehring, & Vossbeck-Elsebusch, 2016), as well as a methodologically rigorous study on panic disorder and agoraphobia (Meuret, Seidel, Rosenfield, Hofmann, & Rosenfield, 2012), reached the same conclusion. As a result of these negative findings, research into other theories on mechanisms of action in exposure is of renewed importance.

Cognitive factors have long been considered important in panic disorder (Clark, 1986), and an integrative review (Casey, Oei, & Newcombe, 2004) supported the role of both “negative” (i.e., dysfunctional, distorted) and “positive” (self-efficacy related) cognitions in CBT for panic disorder and agoraphobia. Conceptually, negative cognitions are linked to an increased sense of threat, while positive cognitions concern an appraisal of available coping abilities. Panic attacks (recurring as part of panic disorder) are seen as an immediate result of high negative and low positive cognitions in the face of triggering stimuli. Agoraphobic avoidance or safety-behaviors constitute dysfunctional coping strategies, which individuals apply in anticipation of future panic attacks (i.e., when cognitions of threat and inability to cope persist in the absence of panic-inducing stimuli; Casey et al., 2004). Both types of cognitions have received tentative support as mediating variables for treatment success in CBT of panic disorder with or without agoraphobia in a recent meta-analytic review (Fentz, Arendt, O'Toole, Hoffart, & Hougaard, 2014) as well as in high-quality individual studies. As such, one multi-center study of panic disorder by Hofmann et al. (2007) found changes in catastrophic beliefs to mediate changes in disorder severity for patients treated by CBT, but not for patients receiving pharmacotherapy. Fentz et al. (2013) similarly found changes in panic self-efficacy to mediate changes in anxiety symptoms using weekly assessments and multilevel modelling.

Unfortunately, most recent studies followed the DSM-IV classification and focused on panic disorder, treating agoraphobia only marginally or implicitly. Nonetheless, there are some indications that negative cognitions are linked even more closely to agoraphobic avoidance than to panic disorder (Khawaja & Oei, 1998; Telch, Brouillard, Telch, Agras, & Taylor, 1989). The same is true for positive cognitions (Casey et al., 2004), which are rooted in the treatment of phobic avoidance and were only later generalized to panic (Williams & Zane, 1997). This is in interesting contrast to the fact that purely cognitive therapy is generally found to be less effective in the treatment of agoraphobia than behavioral exposure, while the empirical status of cognitive interventions for panic disorder is strong (Sánchez-Meca et al., 2010; van den Hout, Arntz, & Hoekstra, 1994). A possible explanation is that cognitive change could be as important (and in some cases brought about more effectively) in behavioral as in cognitive interventions, as has been postulated by Hofmann (2008b) in response to a review of component studies that failed to demonstrate the importance of cognitive interventions in CBT (Longmore & Worrell, 2007). Exploring cognitive effects and mechanisms of exposure therapy for agoraphobia is thus a promising avenue to deepen our understanding of the interplay of cognitive and behavioral interventions and improve treatments for agoraphobia.

The present review aims to improve the knowledge on mechanisms of action in CBT by shedding light on the role of cognitive changes in exposure-based treatments for agoraphobia. We do this by conducting (1) a systematic review and meta-analysis regarding changes in cognitive variables in pure exposure treatments for patients with agoraphobia. We also performed meta-analyses on the differences in post-treatment cognitive outcomes of pure exposure treatments compared with (a) active cognitive or cognitive-behavioral treatments and (b) wait-list control conditions in the subset of studies that included such groups. Furthermore, we conducted (2) a systematic review regarding whether cognitive variables mediate the effects of exposure-based treatments (i.e., including exposure exercises alone or in combination with other interventions in a CBT framework) on symptoms of panic and agoraphobia. We performed sub-analyses (a) grouping studies according to the relative importance of in-vivo exposure in pure, central or some exposure as detailed in the methods section; (b) grouping studies according to whether or not their design establishes temporal precedence of the potential mediator over symptom outcomes by measuring them at different or multiple time points; and (c) grouping results for negative (dysfunctional) and positive (self-efficacy related) cognitions. We calculated tentative effect size estimates of the mediation for the overall effect and separately for negative and positive cognitions. Traditionally, mediation is often assessed by a series of (partial) correlations between the independent and dependent variables and hypothesized mediators (Baron & Kenny, 1986), which can also be aggregated meta-analytically as has been done in the review by Fentz et al. (2014). However, this excludes modern statistical approaches to mediation such as latent growth curves (e.g., Gallagher et al., 2013) or linear mixed models (e.g., Gloster et al., 2014). Furthermore, the correlational approach is usually performed on cross-sectional data and neglects important aspects of research design such as establishing whether changes in hypothesized mediators happen before changes in outcome variables (Kazdin, 2007). As an example, Bouchard and colleagues (Bouchard et al., 2007) used time series analyses of diary data to demonstrate this temporal precedence. Thus, we decided to look beyond correlational studies, which lend themselves to meta-analysis, and narratively review a broader range of studies on mediation contributing insights from innovative research designs and statistical methods. As will be elaborated in the discussion, future studies of mediators should adopt coherent design and statistical frameworks (e.g., Kraemer, Kiernan, Essex, & Kupfer, 2008) to enable a meta-analytical verification of our findings. In the light of growing support for the efficacy of virtual reality exposure therapy (Powers & Emmelkamp, 2008), we further decided to include pertinent studies alongside traditional exposure interventions based on extinction or habituation (Jansson & Öst, 1982) and guided mastery rationales (Williams, 1990).

1. Method

1.1. Identification and selection of studies

We searched the PsycInfo and PubMed databases for publications in English language published up until November 2017, requiring a match

for either panic disorder or agoraphobia in combination with a number of search terms for CBT treatments by searching title and abstract with the term (panic disorder OR agoraphobi*) AND (cognitive behavio* OR behavio* therapy OR cognitive therapy OR CBT OR exposure OR guided mastery), resulting in 3565 matches. The references provided by articles included in the review were searched for further relevant publications, yielding two more potentially relevant articles. After removing duplicates, 2679 publications were assessed for inclusion criteria.

Inclusion criteria were as follows: (a) peer-reviewed journal articles; (b) adult study populations (18–65 years) because cognitive processes could plausibly differ in youth and older adults (c) including at least one subsample with a majority of patients diagnosed with agoraphobia or panic disorder *with* agoraphobia; (d) reporting data on at least one treatment condition including exposure to agoraphobic stimuli; (e) reporting cognitive measures. Further inclusion criteria were (f1) reporting cognitive outcomes of a pure exposure condition for inclusion in our first, meta-analytic research question and/or (f2) reporting mediation analyses with cognitive measures as mediators and measures of symptom severity as outcomes for our second research question. We excluded articles that reported on previously published findings or did not contain sufficient data (the relevant means and standard deviations) for our meta-analysis.

We first screened titles and abstracts of all publications for studies that clearly did not meet inclusion criteria. Specifically, $n = 2211$ publications were identified as not reporting primary empirical studies including patients with panic disorder and/or agoraphobia treated by some variant of behavioral therapy CBT, and $n = 305$ publications were identified as reporting on study designs unsuitable for the purposes of our review (i.e., focused on comorbid conditions, treatment-resistant populations, etc.). Of the remaining 163 publications, we could obtain the full text for 157 publications ($n = 5$ excluded as unobtainable) and reviewed them for eligibility according to the inclusion criteria described above (the PRISMA flow diagram in Fig. 1 depicts the selection and inclusion process in detail). This resulted in our final sample of 30 articles, of which 15 were suitable for our first analysis of cognitive outcomes of pure exposure and 17 for the second analysis of cognitive mediation in exposure-based treatments (including two articles suitable for both analyses, see Table 1 for details of the included articles).

1.2. Coding of study characteristics

The following data were extracted from the 30 articles included in our analyses: whether the study was designed as a randomized controlled trial (RCT) or an open trial; whether agoraphobia (AG) or panic disorder with agoraphobia (PDA) was the primary diagnosis investigated in the sample or subsample relevant to our analyses; sample sizes of the total sample and of the participants receiving exposure treatment; treatment groups reported in the study together with format (individual vs. group), number of sessions and total duration of the treatment in hours; the relative importance of exposure in the treatment, either *pure* (no other psychotherapeutic interventions beyond basic psychoeducation delivered), *central* (additional interventions described as secondary to exposure, more than 50% of treatment time spent in exposure), *some* (exposure included among other interventions of equal importance) or *none* (no exposure interventions in this treatment group); time points of measurements used for the original analyses, including intermediary and follow-up assessments; cognitive measures used as mediator variables; cognitive or panic- and agoraphobia-related measures used as outcomes for our first and second analysis, respectively; whether the study summarily supports cognitive mediation in exposure treatment by demonstrating either change in cognition as a result of pure exposure treatment (first analysis) or mediation of clinical outcomes by cognitive variables in exposure-based treatments (second analysis), coded as *yes* (statistically significant effects in a large majority of variables under investigation), *mixed* (statistically significant effects in some variables under investigation), or *no* (no or only few statistically significant effects); effect size estimate as detailed below. Coding was performed by the first author, discussing ambiguous cases with the other authors.

1.3. Research synthesis

For our first analysis concerning change in cognition as a result of pure exposure treatment, we conducted a meta-analysis on the pre-to post-treatment effect sizes. Standardized mean change scores (equivalent to Hedges' *g;* Becker, 1988; Hedges, 1981) were calculated using raw pre-treatment score standardization as recommended by Morris and DeShon (2002) for unknown and assumedly heterogeneous correlations between time-points. In addition to the overall effect, we calculated summary effect sizes separately for negative (dysfunctional) and positive (self-efficacy-related) cognitions. Furthermore, we performed meta-analyses on the differences in post-treatment cognitive outcomes of pure exposure treatment compared with (a) active cognitive or cognitive-behavioral treatments and (b) wait-list control conditions in the subset of studies that included such groups. Effect sizes for these comparisons were estimated as Hedges' *g* (Hedges, 1981) standardized mean differences between post-treatment measurements of the respective groups.[1] When studies reported multiple relevant cognitive outcomes, we combined their effect sizes and corresponding standard errors using the procedure recommended by Borenstein and colleagues (Borenstein, Hedges, Higgins, & Rothstein, 2009), assuming a correlation between measures of $r = 0.7$. Similarly, if more than one treatment condition qualified for inclusion in our analysis, we combined their means and standard deviations for subsequent analytic steps.

Random effects estimation was used in all meta-analyses because the diverse studies included in our analyses cannot be assumed to share a common true effect, and because of better external generalizability (Borenstein, Hedges, Higgins, & Rothstein, 2010). We assessed the likelihood of publication bias (small studies with large effect sizes leading to overestimation of the average effect size because small effect sizes remain unpublished/"file drawer" problem) by visually inspecting funnel plots and regressing studies' effect sizes on their respective standard errors with Egger's test (Egger, Smith, Schneider, & Minder, 1997). If there were indications for possible publication bias, we used the trim and fill method (Duval & Tweedie, 2000) to check the robustness of our results when unpublished small effect sizes are assumed. Calculations were performed in R version 3.3.2 (R Core Team, 2017) using the metafor package (Viechtbauer, 2010).

For our second analysis concerning mediation of clinical outcomes by cognitive variables in exposure-based treatments, we aggregated the reported statistical significance in the primary studies following a box score approach with the three categories yes, mixed and no as detailed above. To provide more nuanced information, we performed sub-analyses (1) grouping studies according to the relative importance of in-vivo exposure in pure, central or some exposure as detailed above; (2) grouping studies according to whether or not their design establishes temporal precedence of the potential mediator over symptom outcomes by measuring them at different or multiple time points; and (3) grouping results for negative (dysfunctional) and positive (self-efficacy related) cognitions. In addition to statistical significance, assessment of effect sizes in terms of variance explained by each mediator contributes important information to judge their relative importance. Unfortunately, the statistical methods employed to test for mediation

[1] The pre-to-post within-group and post-post control-group effect sizes presented here are calculated very similarly and are comparable in general. Nonetheless, differences in magnitude are possible due to factors such as spontaneous change in the untreated control group or changes in variability over time (Morris & DeShon, 2002).

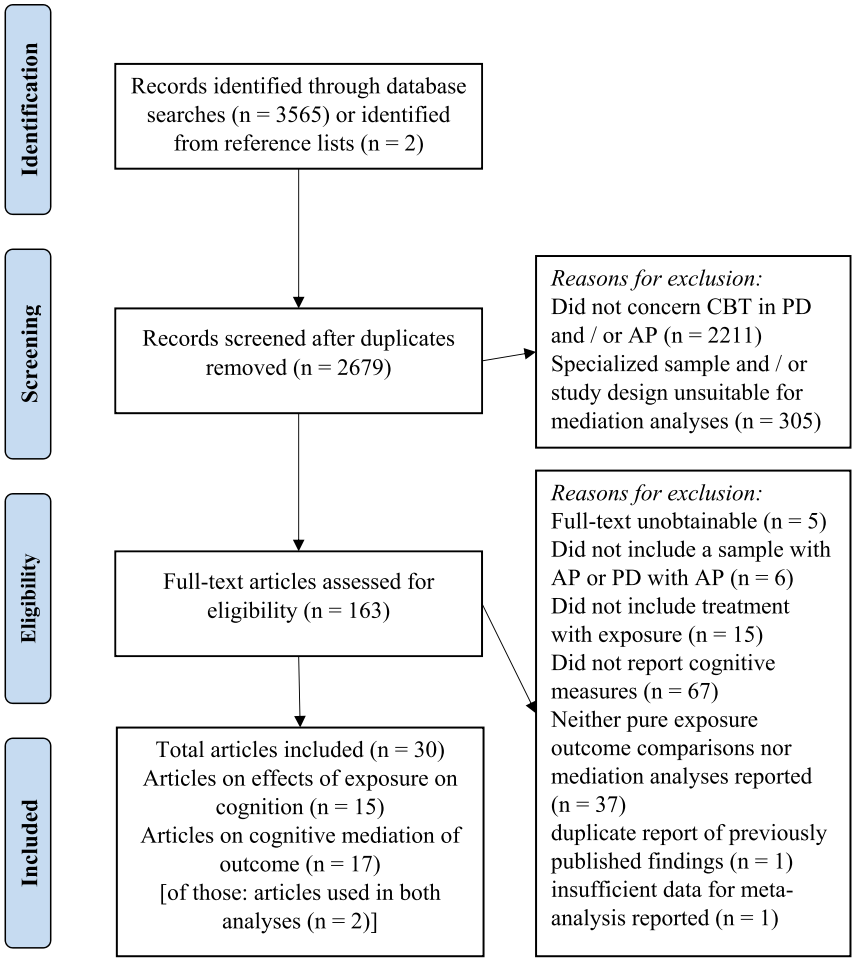

Fig. 1. Flow chart of study selection.

have evolved significantly over time, with a clear consensus on the most appropriate approach still lacking today (Preacher, 2015). The studies included in our analysis employ statistical methods ranging from partial or multiple correlation (e.g., Williams, Kinney, & Falbo, 1989; Zane & Williams, 1993) through path analysis (Hoffart, 1998) and multivariate time series (Bouchard et al., 2007) to multi-level/latent growth curve/latent difference score models (Fentz et al., 2013; Gallagher et al., 2013; Gloster et al., 2014). This makes comparing effect size measures extremely problematic. Nonetheless, we extracted a measure of explained variance (R^2) for potential mediators wherever possible. If available, the explained variance was taken from the original authors' analyses. In many older studies (e.g., Casey, Newcombe, & Oei, 2005) this was calculated from correlations between pre-to post-treatment change scores in mediators and post-treatment scores in symptom outcome. However, a range of other calculations were also used, e.g. via the correlation between actual outcome and the outcome predicted from a multi-level model (Fentz et al., 2013). These R^2 values were averaged for each of the levels and groups of aggregation mentioned above. Given the heterogeneity and the limitations of older statistical approaches, our aggregated R^2 scores can only serve as very rough estimates.

2. Results

2.1. Characteristics of included studies

Detailed characteristics of the included studies are displayed in Table 1. Our first meta-analytic analysis included 15 studies with a total of 29 study arms and 921 participants relevant to our analyses (684 participants in 15 pure exposure study arms, 55 participants in 5 waitlist control study arms, and 182 participants in 9 active cognitive or cognitive-behavioral control study arms). With the exception of the large open trial by Hahlweg and colleagues (Hahlweg, Fiegenbaum, Frank, Schroeder, & von Witzleben, 2001) all included studies were randomized controlled trials. Psychotherapy was conducted in individual format (n = 7, 46%) or group format (n = 5, 33%) in the majority of studies, with two studies (13%) using virtual reality and one study (6%) instructions for self-exposure with training and supervision by a therapist. Studies report an average of 10.4 sessions of psychotherapy (range 2–18) with an average duration of 18.7 h (range 2–48).

Our second, narrative analysis included 17 studies with a total of 1881 participants.[2] Ten of these studies (59%) were randomized controlled trials, whereas five (29%) were open trials and two (12%) report on participants included selectively from larger samples. Psychotherapy was conducted in group format in 9 studies (53%) and in individual format in 8 studies (47%). Studies report an average of 12.1 sessions of psychotherapy (range 3–21) with an average duration of 23.4 h (range 4–48).

[2] Of these, 537 participants are also included in our meta-analysis, and 169 are duplicated within the narrative review between the different publications by Hoffart and colleagues. Thus, the total number of unique participants reported on in both analyses is 2096.

Table 1
Overview of included studies (n = 30).

Nr.	Publication	Design/ statistics	Primary diagnosis (exposure N/total N)	Treatment groups (nr. of sessions, total duration)	Amount of exposure	Measurement time points	Mediator measures	Outcome measures	Support for cognitive mediation in exposure	Effect size estimate Hedges' g/R^2
Evaluated effect of exposure on cognitive outcomes (n = 15)										
1	Bouchard et al. (1996)[a]	RCT	PDA (14/28)	G-EXP/G-CT (15, 22.5 h)	pure/ none	pre, S 5, S 10, post, 6-mo FU		ACQ, SE-CPAQ	yes (CT not superior to EXP)	1.54 (pre-post G-EXP)
2	Burke, Drummond, and Johnston (1997)	RCT	AG (14/26)	EXP/CBT (10, 25 h/30 h)	pure/ central	pre, post, 6-mo FU		ACQ, CCL, PQ	yes (CBT not superior to EXP)	0.66 (pre-post EXP)
3	de Beurs, van Balkom, Lange, Koele, and van Dyck (1995)	RCT	PDA (18/76)	EXP/EXP + MED/EXP + PBO/CBT (12, 12 h + EXP HW)	pure/ central	−6 wk, pre, S 6, post,		composite of ACQ & BSQ	mixed (only EXP + MED superior to EXP)	0.69 (pre-post EXP)
4	Hahlweg et al. (2001)[b]	open trial	PDA (400)	EXP (6–9 + 4–10 exposure days, av. 30 h)	pure	pre, post +6 wk, 1-year FU		ACQ	yes (significant change both pre-post and pre-FU)	1.00 (pre-post)
5	Hoffart (1995a)[c]	RCT	PDA (23/46)	G-GM/G-CBT (6 weeks inpatient, 32 h)	pure/ central	pre, wk 1, wk 2, wk 3, post		ACQ, PAI-S, SESA alone short	mixed (CBT slightly superior to GM)	0.69 (pre-post GM)
6	Ito et al. (2001)	RCT	PDA (20*/75)	S-EXP/S-INT/S-EXP + S-INT/WL (7, 7 h + EXP HW)	pure/ other/ central	pre, S 6, post, 3-mo FU, 6-mo FU, 12-mo FU		ACQ	yes (all treated groups similar and superior to WL)	0.97 (pre-post EXP)
7	Malbos, Rapee, and Kavakli (2013)	RCT	PDA (9/18)	VR-EXP/VR-EXP + CT (8, 12 h/10, 15 h)	pure/ central	pre, post, 3-mo FU		ACQ	yes (VR-EXP + CT not superior to VR-EXP)	0.63 (pre-post VR-EXP)
8	Marchand et al. (2008)	RCT	PDA (30/122)	G-EXP/G-CBT/G-CT/G-ST [all crossed with MED vs PBO] (14, 42 h)	pure/ central/ none/ none	pre, post, 3-mo FU, 6-mo FU, 12-mo FU		ACQ	yes (no differences between treatments)	1.38 (pre-post G-EXP)
9	Michelson et al. (1996)	RCT	PDA (24/74)	G-EXP/G-CBT/G-EXP + G-RT (16, 48 h)	pure/ central/ central	pre, S 8, post, 3-mo FU, 6-mo FU, 12-mo FU		ACQ	yes (no differences between treatments, all superior to normative control)	1.22 (pre-post G-EXP)
10	Pelissolo et al. (2012)	RCT	PDA (19[33*]/63[67*])	VR-EXP/CBT/WL (12, 12 h)	pure/ some	pre, post, 6-mo FU, 12-mo FU		ACQ	yes (CBT not superior to VR-EXP)	0.60 (pre-post VR-EXP)
11	Southworth and Kirsch (1988)	RCT	AG (10 + 10/32)	EXP/EXP-ASS/WL (10, 10 h)	pure/ pure	pre, post		SE (beh.)	yes (increase in both EXP groups and not in WL)	1.80 (pre-post, avg of EXP groups)
12	Telch, Agras, Taylor, Roth, and Gallen (1985)	RCT	AG (9/29)	G-EXP + PBO/G-EXP + MED/MED + NE (6, 13.5 h)	pure/ pure/ none	pre, post, 4.5-mo FU		SE (beh.)	mixed (increase in G-EXP + PBO, but less than in G-EXP + MED, and not significantly more than MED + NE)	1.65 (pre-post G-EXP + PBO)
13	Williams and Falbo (1996)	RCT	PDA (12/48)	GM/CT/CBT/WL (8, 8 h + HW)	pure/ none/ central	pre, post, 1.5-mo-FU		SESA, PC-SE, ACQ	yes (GM equal or superior to CT, not inferior to CBT, all superior to WL)	1.51 (pre-post GM)
14	Williams and Zane (1989)	RCT	AG (10[11*] + 11[15*]/27)	EXP/GM/WL (1–3, 1–3 h)	pure/ pure	pre, post, 1-mo-FU		SE (beh.), PD (beh.)	mixed (GM superior to WL,	0.75 (pre-post, avg GM and EXP)

(*continued on next page*)

Table 1 (*continued*)

Nr.	Publication	Design/ statistics	Primary diagnosis (exposure N/total N)	Treatment groups (nr. of sessions, total duration)	Amount of exposure	Measurement time points	Mediator measures	Outcome measures	Support for cognitive mediation in exposure	Effect size estimate Hedges' g/R^2
15	Zane and Williams (1993)	RCT	PDA or AG (16[20*] + 19[22*]/45)	EXP/GM/WL (1–5, 1–7 h)	pure/ pure	pre, post, 2-mo-FU		SE, SE-ND, PD (all beh.)	EXP superior to WL only for SE) mixed (treatments superior to WL for SE, not for PD)	0.62 (pre-post, avg GM and EXP)
Evaluated cognitive mediation of symptom severity outcomes (n = 17)										
16	Bouchard et al. (2007) [a]	selective sub-sample/ multivariate time series	PDA (7/12)	G-EXP/G-CT (15, 22.5 h)	pure/ none	daily diary from -6wk to 6wk-FU	based on ACQ & SE	panic apprehension diary	mixed	–
17	Casey, Newcombe, and Oei (2005)	RCT/ regressions[†]	PDA (36/60)	CBT/WL (12, 12 h)	some	pre, post	BBSIQ, SE-CPAQ	PAS	yes (full mediation)	.59 (cor pre-post/ post)
18	Casey, Oei, and Newcombe (2005)	RCT/multiple regression	PDA (106)	CBT (6 or 12, 6 h or 12 h or 6 h + palmtop)	some	pre, wk 3, wk 6, post	BBSIQ, SE-CPAQ	PAS	yes	.18 (ΔR^2)
19	Fentz et al. (2013)	open trial/ multilevel analyses	PDA (45)	G-CBT (15, 24 h)	some	pre, post, weekly during treatment	PAI-C, CBS	PDSS, BAI	mixed (only for PAI-C, and reverse effect of BAI on PAI-C also found)	.25 (proportion of total outcome variance explained by PAI-C)
20	Gallagher et al. (2013)	open trial/ latent growth curve & latent difference score	PDA (361)	CBT (11, 11 h–13 h)	some	weekly during treatment	ASI, SE (sit.)	PDSS-SR	yes	.53
21	Gloster et al. (2014)	RCT/linear mixed models & latent difference score	PDA (301)	CBT (14, 14 h)	central	pre, wk 4, post, 6-mo-FU	ACQ, ASI, AAQ-II	PAS, CGI-S-F	yes (most variables in most intervals)	–
22	Hoffart (1995b)[c]	RCT/path analysis	PDA (23 + 23/46)	G-CBT/G-GM (6 wks, 32 h)	central/ pure	pre, post	SE (beh.), PD (beh.)	SF (beh.)	mixed (only for SE)	– (standardized beta −.56)
23	Hoffart (1998)[c]	RCT/path analysis	PDA (23 + 23/46)	G-CBT/G-GM (6 wks, 32 h)	central/ pure	pre, post, 12-mo FU	ACQ, SESA	separate composite scores for panic and avoidance	mixed (only for ACQ and some outcome variables)	–
24	Hoffart (2016)[c]	RCT/linear mixed models	PDA (23 + 23/46)	G-CBT/G-GM (6 wks, 32 h)	central/ pure	pre, wk 2, wk 3, wk 4, post	ACQ, SESA, PAI-S	STAI	yes (all mediators in individual models, SESA reduced in combined model)	–
25	Hoffart, Hedley, Svanøe, and Sexton (2016)[c]	RCT/linear mixed models	PDA (16 + 15/31)	G-CBT/G-GM (6 wks, 32 h)	central/ pure	pre, post, 1-year FU, 2-year FU, 18-year FU	ACQ, SESA	PARS, MI, BSQ, STAI	mixed (only ACQ for some outcomes)	–
26	Hoffart, Sexton, Hedley, and Martinsen (2008)[c]	open trial/ structural equation modelling	PDA (23 + 23 + 31 + 41 + 111/188)	G-CBT/G-GM (6 wks, 32 h)/G-CBT + PCT/G-CBT + SFT (6 + 6 wks, 32 h + ca. 30 h)	some	pre, post	ACQ	MI, BSQ	mixed (effect of ACQ on MI, but BSQ on ACQ)	–
9	Michelson et al. (1996)	RCT/median split ANCOVA	PDA (24 + 25 + 25/74)	G-EXP/G-CBT/G-EXP + G-RT (16, 48 h)			ACQ		mixed (yes in median split, only	.10 (adj R^2 ACQ 6-mo FU)

(*continued on next page*)

Table 1 (*continued*)

Nr.	Publication	Design/ statistics	Primary diagnosis (exposure N/total N)	Treatment groups (nr. of sessions, total duration)	Amount of exposure	Measurement time points	Mediator measures	Outcome measures	Support for cognitive mediation in exposure	Effect size estimate Hedges' g/R^2
		& multiple regression			pure/ central/ central	pre, S 8, post, 3-mo FU, 6-mo FU, 12-mo FU		GAS, composite endstate functioning	for 6-mo FU in regression)	
27	Smits, Powers, Cho, and Telch (2004)	RCT/AN(C) OVA[†]	PDA (90/130)	G-CBT (12, 24 h)	some	pre, post	composite of ASI & BSQ	TPARF, SPRAS, FQ-Ago, SDS	yes (full for SDS, partial for others)	.27 (avg diff in % var explained by treatment with FOF partialled out)
28	van Hout, Emmelkamp, and Scholing (1994)	most and least improved cases from open trial/narrative summary	PDA (8)	EXP (8–11, 12–17 h)	pure	in 30 min intervals during exposure exercises	ASQ negative subscale	composite of agoraphobic fear and avoidance scales	yes (cognitive change differentiates between extreme groups, physiological habituation does not)	–
29	Vögele et al. (2010)[b]	open trial/ linear mixed models	PDA (427)	EXP (6–9 + 4–10 exposure days, av. 30 h)	pure	pre, post +6 wk, 1-year FU	ACQ subscales physical and mental/social consequences	MI-alone	yes (partial mediation for both subscales)	.29 (subscales combined)
30	Williams et al. (1989)	open trial/ partial correlations	PDA (27)	GM (6, 9 h)	pure	pre, mid, post, 5-mo FU	SE (beh.), PD (beh.)	BAT performance	mixed (only for SE, not for PD)	.36 (cor SE with beh in treated and transfer phobias)
15	Zane and Williams (1993)	RCT/multiple correlation	PDA or AG (16[20*] + 19[22*]/45)	EXP/GM/WL (1–5, 1–7 h)	pure/ pure	pre, post, 2-mo-FU	SE, SE-ND, SE-CT, PD (all beh.)	BAT anxiety	mixed (only SE-CT when controlling for anticipatory anxiety)	.04 (Δ R^2 added to anticipatory anxiety)

Note: [a,b,] ⋯analyses with overlapping samples share the same superscript; * after WL re-randomization, [†] following the procedure recommended by Baron and Kenny (1986). Diagnosis: *N* number of treatment completers with valid data, *PDA* panic disorder with agoraphobia; *AG* agoraphobia; Treatment: *G*-therapy in group format, *S*- self-directed exposure with training, specific instructions, and supervision provided by a therapist, *VR*-exposure conducted in virtual reality environment, *CT* cognitive therapy (without behavioral elements), *CBT* cognitive behavioral therapy (including some form of exposure exercises), *EXP* exposure therapy, *EXP-ASS* exposure therapy without therapeutic expectations (framed as assessment to patients), *GM* guided mastery (performance-oriented exposure), *INT* interoceptive exposure, *HW* homework, *MED* medication, *MED + NE* medication combined with instruction to not expose oneself to feared situations, PBO placebo, *PCT* problem-focused cognitive therapy, PM panic management, *RT* relaxation training, *SFT* schema-focused therapy, *ST* supportive therapy control group, *WL* waiting list control group; Measurement time points: *pre* pre-treatment, S x after session x, *post* post-treatment, *+/− x wk* x weeks before/after, *x-mo FU* follow-up after x months; Measures: *AAQ-II* Acceptance and Action Questionnaire-II, *ACQ* Agoraphobic Cognition Questionnaire, *ASQ* Agoraphobic Self-Statements Questionnaire, *BAI* Beck Anxiety Inventory, *BAT* behavioral avoidance test, *BBSIQ* Brief Bodily Sensations Interpretations Questionnaire, *BSQ* Bodily Sensations Questionnaire, *CBS* Catastrophic Belief Scale, *CCL* Cognitions Checklist, *CGI-S-F* Clinical Global Impression Scale–severity subscale–overall functioning item, *FQ-Ago* Fear Questionnaire–Agoraphobia subscale, *GAS* Global Assessment of Severity, *MI* Mobility Inventory, *PAI* Panic Appraisal Inventory: C Panic Coping subscale –S Social subscale, *PAS* Panic and Agoraphobia Scale, *PD (beh.)* perceived danger (measured as subjective likelihood of harm during exposure or behavioral assessment), *PDSS* Panic Disorder Severity Scale, *PDSS-SR* Panic Disorder Severity Scale–self-report version, *PQ* Probability Questionnaire, *SE (beh.)* self-efficacy (measured as expected ability to perform target behavior during exposure or during behavioral assessment), *SE (sit.)* self-efficacy (measured as expected ability to perform target behavior in different situations), *SE-CT (beh.)* self-efficacy (expected ability) for controlling scary thoughts during exposure tasks, *SE-ND (beh.)* similar to SE (beh.) but performing target behavior non-defensively (without safety behaviors), *SE-CPAQ* Self-Efficacy to Control a Panic Attack, SESA Self-Efficacy Scales for Agoraphobia, *SDS* Sheehan Disability Scale, *SF (beh.)* situational fear (measured as anxiety experienced during exposure or during behavioral assessment), *SPRAS* Sheehan Patient-Rated Anxiety Scale, *STAI* State-Trait Anxiety Inventory, *TPARF* Texas Panic Attack Record Form, *PARS* Phobic Avoidance Rating Scale, *PC-SE* panic coping self-efficacy (unpublished measure); Effect size estimates: *cor* correlation, *pre-post/post* pre-post difference scores of mediators with post-scores of outcomes.

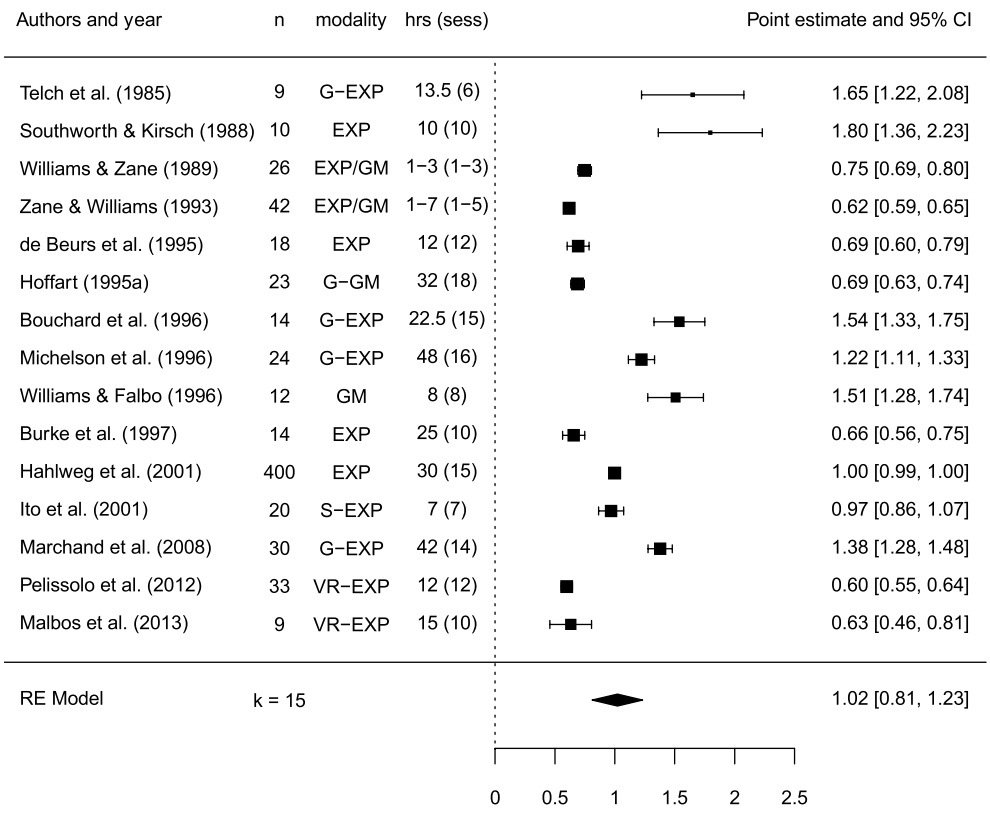

Fig. 2. Forest plot of the standardized mean change (equivalent to Hedges' *g*) from pre-to post-treatment in cognitive outcome variables for pure exposure treatments. *n* participants in exposure conditions; *modality* exposure therapy format (*G*-group, *VR*-virtual reality, *S*- self-instruction, else individual) and framework (*EXP* stimulus exposure, *GM* guided mastery); *hrs* (*sess*) hours (sessions) of exposure therapy provided. The black squares represent individual studies' standardized mean change, with the left and right whiskers representing the corresponding 95% confidence intervals. The middle of the black diamond represents the overall standardized mean change in the random effects (RE) model, with the left and right extremes of the diamond representing the corresponding 95% confidence interval. Studies are ordered by year of publication.

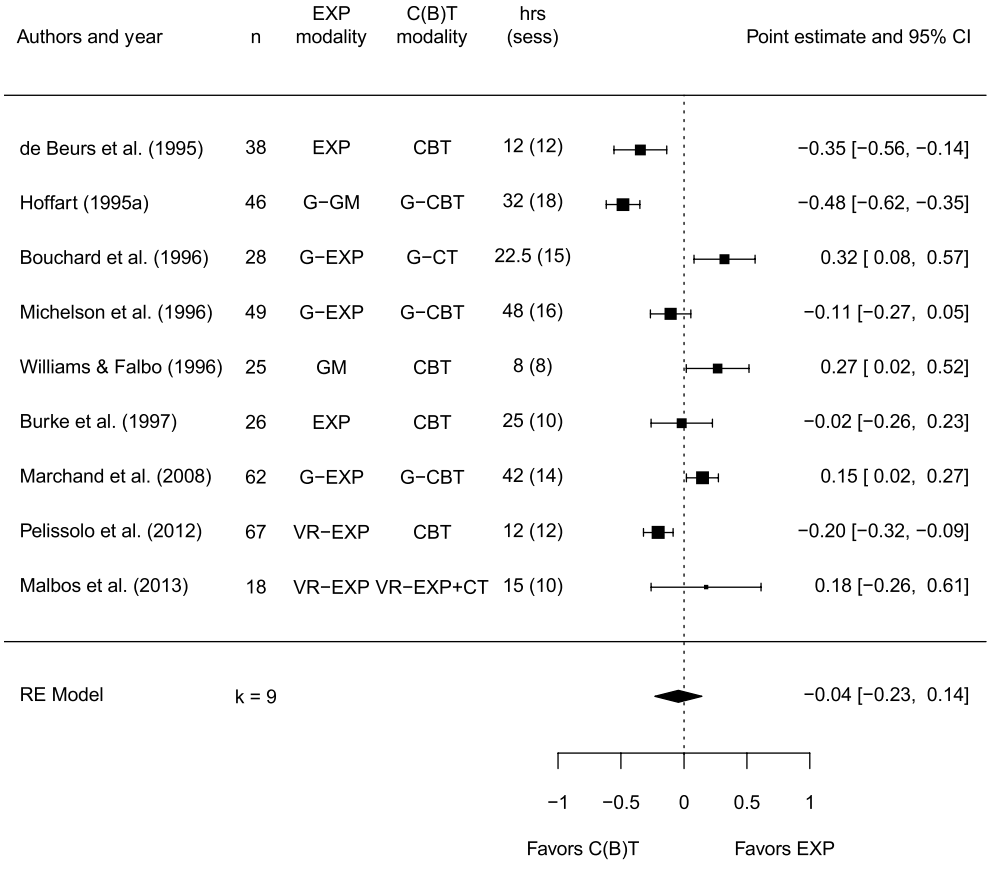

Fig. 3. Forest plot of the standardized mean differences (Hedges' *g*) in cognitive outcome variables between cognitive or cognitive-behavioral *C(B)T* and pure exposure *EXP* treatments. *n* participants combined in EXP and C(B)T conditions; *modality* therapy format (*G*-group, *VR*-virtual reality, *S*- self-instruction, else individual) and framework (*CBT* cognitive-behavioral therapy, *CT* cognitive therapy, *EXP* stimulus exposure, *GM* guided mastery); *hrs* (*sess*) hours (sessions) of therapy provided. The black squares represent individual studies' standardized mean difference, with the left and right whiskers representing the corresponding 95% confidence intervals. The middle of the black diamond represents the overall standardized mean difference in the random effects (RE) model, with the left and right extremes of the diamond representing the corresponding 95% confidence interval. Studies are ordered by year of publication.

2.2. Meta-analysis of the effect of pure exposure treatments on cognitive outcomes

The overall pre-to-post change in cognitive outcomes in pure exposure treatments from the meta-analysis of 15 studies indicated a large effect size of 1.02 (95% CI 0.81–1.23; range 0.60–1.80; $p < .001$; see Fig. 2). There was large between-study heterogeneity ($Q(14) = 1318.0$, $p < .001$, $I^2 = 99.6\%$). Egger's test indicated significant asymmetry in the funnel plot, indicating a potential file drawer problem. The trim and fill method estimated two small studies with small effect sizes to be missing, but imputing them resulted in an only marginally reduced effect size estimate of 0.94 (95% CI 0.71–1.17; range 0.07–1.80; $p < .001$). For positive (i.e., self-efficacy related) cognitive outcomes, our meta-analysis of seven studies indicated a large pre-to-post effect size of 1.21 (95% CI 0.85–1.57; range 0.61–1.80; $p < .001$). For negative (i.e., dysfunctional) cognitive outcomes 13 studies yielded a large effect size of 0.92 (95% CI 0.72–1.11; range 0.56–1.61; $p < .001$). Both sub-analyses resembled the main analysis concerning heterogeneity and the funnel plot pattern, but had no change in effect size estimates resulting from the trim and fill procedure. Five studies provided data for a between-group post-treatment comparison of pure exposure treatment with wait-list controls and resulted in a large effect size estimate of 1.02 (95% CI 0.73–1.32; range 0.67–1.45; $p < .001$) with moderately large between-study heterogeneity ($Q(4) = 19.1$, $p = .001$, $I^2 = 80.0\%$). There was no indication of publication bias for this analysis. Given that pre-to-post effect sizes are frequently criticized for confounding treatment and time effects, it is reassuring that the effect size estimates from this between-group post-treatment comparison are very similar in magnitude. This suggests that spontaneous improvement in agoraphobic cognitions is a rather negligible aspect.

Comparing pure exposure treatments with active cognitive or cognitive-behavioral control conditions was possible with post-treatment data from 9 studies (see Fig. 3) and resulted in a negative between-group effect size estimate of −0.04 (95% CI -0.23–0.14; range −0.48 to 0.32; $p = .646$), which was neither statistically significant nor meaningful in size. There was large between-study heterogeneity ($Q(8) = 76.2$, $p < .001$, $I^2 = 89.5\%$) and no sign of publication bias. Taken together, the meta-analyses suggest that pure exposure treatments yield large improvements in both positive and negative cognitions, equivalent to changes brought about by cognitive and cognitive-behavioral treatments, which target cognitions explicitly.

2.3. Narrative review of mediation of clinical outcomes by cognitive variables in exposure-based treatments

As can be seen in Table 2, eight (numbers 17,18,20,21,24,27,28,29 in Table 1) out of seventeen publications reported overall support for cognitive mediation of therapeutic gains during exposure, and nine publications (9,15,16,19,22,23,25,26,30) reported mixed support (i.e., only for some predictors or outcomes under investigation), with no publications reporting negative findings. Seven studies provided sufficient data to calculate an estimate of variance explained by the cognitive mediators, with a mean R^2 of 0.32 ($SD = 0.18$).

2.4. Comparison of studies by relative amount of exposure and cognitive interventions

When grouping studies by the role of exposure (i.e., the amount of additional cognitive interventions, if any), the support for cognitive mediation is strongest in comprehensive CBT programs, with four out of five studies (17,18,20,27) reporting positive findings and only one mixed result (19). Of the five studies with exposure as the central intervention, only two (21,24) report overall positive results for cognitive mediation, with the remaining three reporting mixed results (9,25,26). Similarly, only two out of seven studies with pure exposure treatments report overall positive results for cognitive mediation (28,29), with again all five remaining publications reporting mixed support (15,16,22,23,30). Unfortunately, effect size estimates for the mediation effect could be extracted for only one study with central and pure exposure, each, precluding any reliable generalization (see Table 2).

Being of special interest to our review, we briefly expand on the mixed findings concerning mediation in pure exposure treatments. Zane and Williams (1993; #15) used a pre-post stepwise multiple correlations design. Anticipatory anxiety (which we did not consider a cognitive measure because it is conceptually too close to symptomatic measures) was the strongest predictor of outcome, with other cognitive measures entered only if they added incremental explanatory power. In these analyses, only changes in self-efficacy to control a panic attack predicted outcome, but not other variants of self-efficacy or perceived danger (all assessed with direct behavioral reference to confronting the feared situations). All predictors were highly intercorrelated, with no theoretical reason to attribute primacy to anticipated anxiety. Thus, their results support cognitive mediation, weakened only by the unclear relationships between different aspects of cognition. Bouchard et al. (2007; #16) used daily diary assessments from twelve participants selected from a larger sample on the basis of showing therapeutic success. The roughly 200 data points from each participant were analyzed individually as multivariate time series, interpreting cross-lagged relationships as indicators of mediation. We coded their results as mixed because even though some cognitive variable significantly predicted outcome for all patients, the precise cognition differed between patients. Only in two out of seven patients treated with exposure was

Table 2
Support for cognitive mediation overall (n = 17) and stratified in different subgroups, with corresponding effect size estimates.

Subgroups	Visual results (box score)	Mean R^2 effect size estimate (SD and number of studies included)
Overall	+ + + + + + + + +[a] ? ? ? ? ? ?[a] ?[a] ?[a] ?[a]	0.32 ($SD = 0.18$; $n = 7$)
Role of exposure		
pure	+ + ? ? ? ?[a] ?[a]	0.29 ($n = 1$)
central	+ +[a] ? ?[a] ?[a]	0.10 ($n = 1$)
some (comprehensive CBT)	+ + + + ?	0.36 ($SD = 0.18$; $n = 5$)
Study design		
prospective	+ + + ? ? ? ?[a]	0.29 ($SD = 0.22$; $n = 3$)
cross-sectional	+ + + + +[a] ? ? ?[a] ?[a] ?[a]	0.33 ($SD = 0.18$; $n = 4$)
Type of cognitions		
positive (self-efficacy)	+ + + + + + + +[a] +[a] ? -[a] -[a]	0.29 ($SD = 0.24$; $n = 5$)
negative (dysfunctional)	+ + + + + + + +[a] ? ? ?[a] ?[a] ?[a] - - - -[a]	0.30 ($SD = 0.18$; $n = 3$)

Note: [a] Publications with overlapping samples (Hoffart and colleagues, #23-#27 in Table 1). Publications with positive, mixed, and negative results are indicated by +, ?, and - symbols, respectively, and presented in that order. For the overall results and sub-analyses by role of exposure and study design, each publication with relevant data contributes one symbol. For the sub-analysis by type of cognition, publications could contribute one symbol to each category if they reported results separately. SD standard deviation of effect size estimates. Effect size estimates were computed from very different research designs and statistical methods and should be interpreted with extreme caution.

outcome predicted by both self-efficacy and dysfunctional beliefs, while for two patients, only beliefs were predictive, and for three patients, only self-efficacy was predictive. The limitation in support for cognitive mediation thus again lies in unclear roles of different cognitions, with the possibility that their importance varies between patients. Hoffart reports twice on path analyses performed on different data from the same study (1995b; #22, 1998; #23). In the first publication (Hoffart, 1995b; #22), he reports on behavioral measures of self-efficacy and catastrophic beliefs recorded before performance of a behavioral avoidance task (BAT). The results suggest that changes in self-efficacy, but not catastrophic beliefs, are associated with changes in situational fear. In the second publication (Hoffart, 1998; #23), he reports on data from established questionnaires of self-efficacy (SESA) and catastrophic beliefs (ACQ) and includes 12-months follow-up data. Here, only changes in catastrophic beliefs predicted outcome at post and follow-up measurements. This contradiction for different types of measures (questions anchored in a specific situation vs. general questionnaires) in the same sample is puzzling and cannot be resolved easily. Williams and colleagues (Williams et al., 1989; #30) used measures taken during a behavioral avoidance task similar to those published by Hoffart (1995; #22), but included more measurement time points and a design that investigated generalization of treatment effects to phobic situations that were not treated directly. Again, only self-efficacy emerged as a significant predictor, but not catastrophic beliefs.

2.5. Comparison of studies with prospective versus cross-sectional designs

When grouping studies by whether or not their design was prospective and entailed at least three measurement points to establish temporal precedence between changes in supposed mediators and outcome variables, no differences became apparent. Three out of seven prospective studies report positive findings for cognitive mediation (20,21,28), with the remaining four reporting mixed findings (9,16,19,25). Similarly, five out of ten studies with only cross-sectional designs report positive findings (17,18,24,27,29), and the remaining five report mixed findings (15,22,23,26,30). Estimates of average R^2 effect sizes were also similar at about 0.3 for both cross-sectional and prospective designs (see Table 2).

2.6. Comparison of results for positive and negative cognitions as mediators

Comparing results for mediation by positive cognitions (i.e., related to self-efficacy) with those for negative cognitions (i.e., dysfunctional expectations or interpretations), a mediational role is supported more strongly for positive cognitions than for negative cognitions. Eight out of eleven studies report positive findings on mediation by positive cognitions (15,17,18,19,20,22,24,30), with one mixed (16) and two negative results (23,25). Seven out of sixteen studies report positive findings on negative cognitions (17,18,20,21,24,27,28), with five mixed (9,16,23,25,26) and four negative results (15,19,22,30). At the same time, the average R^2 effect size estimate was very similar at values of about 0.3.

3. Discussion

Knowledge about the specific processes or mechanisms through which psychotherapeutic interventions bring about change in symptoms and functioning is crucial for the further development of treatments for panic and agoraphobia (Kazdin, 2007). For exposure therapy specifically, a large number of possible mechanisms have been put forward (Tryon, 2005), with important implications for how to conduct and teach exposure therapy (Craske, Treanor, Conway, Zbozinek, & Vervliet, 2014). Cognitive factors are central to a modern understanding of panic and agoraphobia (Casey et al., 2004), yet their role as mechanisms in exposure therapy remains uncertain. Filling this gap could improve treatments for the substantial minority of patients who do not respond enough to current treatments (Loerinc et al., 2015) and has become even more important as evidence accumulates against emotional processing theory's central mechanisms, fear activation and habituation during exposure, as relevant predictors of treatment outcome (Craske et al., 2008; Rupp et al., 2016).

In the present study, we sought to clarify the role of cognitive changes as a mechanism in exposure therapy in two steps. First, we meta-analytically reviewed evidence on cognitive changes brought about by pure exposure treatments for agoraphobia. Our meta-analysis of 15 studies resulted in a large overall pre-to-post effect size of 1.02 (Hedges' *g*; see Fig. 2), indicating that pure exposure treatments substantially improve cognitive factors relevant to panic and agoraphobia. The effect was stronger for changes in positive (i.e., self-efficacy related) cognitions ($g = 1.21$, 7 studies), but still large when looking only at change in negative (i.e., dysfunctional) cognitions ($g = 0.92$, 13 studies). A meta-analysis of the nine studies that compared pure exposure treatments with cognitive or comprehensive cognitive-behavioral treatments strongly suggests that there is no meaningful difference in their overall cognitive outcomes ($g = -0.04$; 95% CI -0.23–0.14; $p = .646$). These results suggest that pure exposure treatments and comprehensive cognitive-behavioral treatments are quite comparable with regard to changes in cognitions associated with panic and agoraphobia in agoraphobic patients. This is in line with meta-analytic findings that show comparable improvement in panic and agoraphobia symptoms after exposure therapy and comprehensive CBT, both substantially larger than isolated cognitive therapy (Sánchez-Meca et al., 2010). In partial contradiction, a recent Cochrane network meta-analysis (Pompoli et al., 2016) found remission and symptomatic outcome measures of comprehensive CBT to be superior to pure behavior therapy in panic disorder with or without agoraphobia. We suspect that these inconsistent results are due to insufficient differentiation between panic and agoraphobia in both meta-analyses and heterogeneous samples in the original studies. An older meta-analysis (Ruhmland & Margraf, 2001) that separately analyzed outcomes in panic disorder with and without agoraphobia found cognitive interventions (and thus comprehensive CBT compared to pure exposure) to be more effective in panic disorder *without* agoraphobia. As we selected only studies with agoraphobic patients for our analyses presented here, the similarity of cognitive outcomes from CBT and pure exposure is consistent with this literature.

Second, we systematically reviewed studies examining these cognitive changes as mediators of symptom-related treatment outcome (such as agoraphobic fear and avoidance, frequency and intensity of panic attacks, occupational and social impairment, etc.). All 17 publications reviewed report positive findings on some form of cognitive mediation for some outcome variables, but many (nine) also report negative findings for some potential mediators or some outcome variables. These mixed findings are partly attributable to studies investigating positive and negative cognitions separately. More studies report finding a mediational role of positive (self-efficacy-related) cognitions than of negative (dysfunctional) cognitions, but there is nonetheless substantial evidence for a role of both types of cognitions. This is in line with our meta-analytical results showing a more pronounced change in positive than negative cognitions after exposure therapy, with large effect sizes for both. Estimating the effect size of mediation was possible for seven studies and resulted in an average of 32% of variance in outcome variables explained by the mediating variables (to be interpreted very cautiously due to large differences in research design and statistical procedures). This can be considered a large effect, even though it also leaves substantial variance to be explained by other factors. Evidence in favor of cognitive mediation was somewhat stronger in the studies that implemented comprehensive CBT as opposed to studies with less or no cognitive interventions added to exposure therapy. The mixed results for mediation in pure exposure treatments was, again, mostly due to negative results for negative cognitions in some studies. Thus, we can tentatively conclude that

exposure alone changes positive cognitions (self-efficacy) quite strongly and that these changes are significantly involved in symptom change. Changes in negative (dysfunctional) cognitions are also substantial, but seem less important for symptom changes elicited by exposure. Existing cognitive interventions might add to exposure mostly in changing negative cognitions, thus also making these more relevant as mechanisms to symptom change. In support of this reasoning, cognitive interventions in established CBT manuals for panic and agoraphobia focus strongly on negative rather than positive cognitions (e.g., Craske & Barlow, 2007). Nonetheless, as illustrated by the individual trajectories in the daily diary study by Bouchard et al. (2007; #16), it is possible that different types of cognitions (positive vs. negative) are functional for different individuals' therapeutic process. Furthermore, some form of reciprocal interaction between different cognitive and symptom variables is plausible (e.g., Fentz et al., 2013). It is important to note, however, that symptomatic improvement in pharmacotherapy was not associated with cognitive changes in a large and high-quality study of CBT vs. pharmacotherapy in panic disorder (Hofmann et al., 2007). Also, treatment gains are maintained much better to 1-year follow-up after CBT than pharmacotherapy (Gould et al., 1995), which could be related to cognitive changes. In combination with the subset of studies in our review that tested for temporal precedence of cognitive changes over symptom changes, we have good evidence that changes in cognitions are not simply an epiphenomenon of symptom change.

Our findings are congruent with the meta-analysis and review by Fentz et al. (2014), which reviewed studies on positive cognitions as a mediator in CBT treatments, using the correlative framework of mediation by Baron and Kenny (1986). They also found very large effect sizes for changes in overall cognitions after CBT, with even larger changes for positive than negative cognitions. The magnitude of effect sizes is somewhat larger in their meta-analysis (pre-to-post $g = 1.41$ for positive and 1.25 for negative cognitions) than in ours, but the confidence intervals overlap. Thus, our study extends these findings by demonstrating a similar pattern of cognitive change in pure exposure as in comprehensive CBT treatments. In their narrative review of mediational studies, Fentz et al. (2014) include three correlational and four longitudinal studies, finding only very tentative support for cognitive mediators in CBT for panic disorder with or without agoraphobia. Our narrative review includes nine correlational and seven longitudinal studies, thus substantially enlarging the existing review. Furthermore, we were able to draw conclusions regarding cognitive mediation in exposure treatments for agoraphobia, specifically, not only for CBT in general. Our tentative estimate of variance explained by cognitive mediators (32%) is also in line with Fentz and colleagues' (2014) meta-analytic findings regarding the correlation between changes in cognitive variables and changes in outcome variables ($r = 0.53$ for positive and .50 for negative cognitions, yielding roughly 25% of variance explained).

Our findings also support the integrated cognitive model (Casey et al., 2004), which stresses the importance of both positive and negative cognitions in the maintenance of panic disorder and agoraphobia. The authors also note that self-efficacy might be especially relevant to agoraphobic avoidance. Integrating our findings with the promising and influential literature on inhibitory/extinction learning during exposure therapy (Craske et al., 2014; Weisman & Rodebaugh, 2018) one can conclude that relevant mechanisms of therapeutic change in exposure therapy include changes in (a) relatively automatic processes of association between stimuli and threat (Craske et al., 2008) as well as (b) cognitive biases in threat perception and processing (e.g., expectancy of harm, Hofmann, 2008a) and ultimately (c) self-reflexive processes concerning one's ability to cope with threatening situations or sensations (the present manuscript). The growing evidence that treatments work better for individuals where the aspect of the disorder addressed by their main components (e.g., dysfunctional cognitions in cognitive therapy) is *not* very pronounced (Cheavens, Strunk, Lazarus, & Goldstein, 2012; Meuret, Hofmann, & Rosenfield, 2010) allows the conclusion that cognitive interventions that focus on individual patients' relative strengths are a promising avenue. As an example, resource activation techniques could be adapted as an augmentation of exposure exercises (Flückiger, Caspar, Holtforth, & Willutzki, 2009). Further considering the fact that mere exposure seems to improve self-efficacy in many but not all patients, research on "immunizing" strategies, which some patients might habitually use to protect their old, dysfunctional beliefs against new experiences of mastery, could inspire corresponding therapeutic techniques. Such research exists for negative cognitions (i.e., why expectations of harm are sometimes not changed despite contradicting experiences, Rief et al., 2015), but to our knowledge is lacking for positive cognitions (i.e., self-efficacy).

Furthermore, our findings help to clarify a long-standing debate on the importance of cognitive interventions and cognitive mechanisms of action in CBT (Hofmann, 2008b; Longmore & Worrell, 2007). Our finding of strong cognitive effects from pure exposure is in line with Longmore and Worrell's (2007) review of component studies, which failed to find large differences between treatment conditions with and without cognitive interventions. At the same time, our mediational findings are in line with Hofmann (2008b) in supporting an important role of cognitive changes in treatment response. One should note, however, that in contrast to the conclusion drawn by Longmore and Worrell (2007), our review does not imply cognitive techniques to be useless. To the contrary, there is some evidence of added effect and stronger mediation for cognitive interventions especially concerning changes in negative (dysfunctional) cognitions.

Our review identified several shortcomings in the existing literature on mediation, which future research should aim to improve upon. First, the designs and statistical methods used to investigate mechanisms of action are currently very heterogeneous. They range from correlational analyses performed on pre- and post-therapy questionnaires (e.g., Casey, Newcombe, et al., 2005, #17), to the implementation of some intermediate assessments with regression-based (e.g., Casey, Oei, & Newcombe, 2005, #18) or latent linear mixed model analyses (e.g., Gloster et al., 2014, #21) to daily diary time series analyses (e.g., Bouchard et al., 2007, #16). On the one hand, this heterogeneity is a sign of necessary progress, especially concerning the statistical models used in mediation analyses. On the other hand, it is a serious problem for meta-analytical aggregation. Thus, after rightly moving forward from the series of correlation analyses suggested by Baron and Kenny (1986), there is a need to reach a new common standard for mediation analyses.

In this vein, MacKinnon and colleagues (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002) compare a number of statistical methods to test hypothesized mediators for significance within the Baron and Kenny (1986) framework. For more nuanced analyses, there appears to be some convergence in the field towards latent change score analyses (Ferrer & McArdle, 2010), which can model jointly both change over time and dynamic interrelations between different processes. Our review includes two recent publications based on latent change scores (sometimes also called latent difference scores: Gallagher et al., 2013, #20; Gloster et al., 2014, #21). However, some statistical issues remain. For example, there is a strong argument for separating within- and between-person effects in mediation models (Curran & Bauer, 2011), as theories of psychotherapy are in fact concerned with within-person processes. Our review includes two such studies (Fentz et al., 2013, #19; Hoffart, 2016, #24). From a meta-analytic viewpoint, the lack of established effect size measures for these advanced models is a further concern. Arriving at a common standard in this regard is crucial to enable future meta-analyses of mediation effects.

A related shortcoming in the literature is that many studies do not measure outcome variables and hypothesized mediators frequently enough (or at all) between pre- and post-therapy measurements. Even though results in our review did not differ systematically between the longitudinal and the correlational studies included, future research should take care to demonstrate temporal precedence between

purported mediators and outcome variables (Kazdin, 2007; Kraemer et al., 2008). This is not only necessary to draw stronger causal conclusions, but also to discover if different mediators are important at different stages in treatment (as reported by Gloster et al., 2014). In combination with findings on rapid response (Wilson, 1999), we conclude that measurements of potential mediators should be multiple and early. Furthermore, studies that report mediation analyses rarely include a control group to ascertain that changes in mediator or outcome variables are caused by the intervention in question. In our review, only two studies (Michelson, Marchione, Greenwald, Testa, & Marchione, 1996, #9; Zane & Williams, 1993, #15) out of 31 could be included both in the meta-analyses of cognitive change due to exposure and in the review of mediational analyses. To remedy this issue, randomized controlled trials (RCTs) should routinely include in their design and analyses questions of moderation and mediation (Kraemer, Wilson, Fairburn, & Agras, 2002).

A further limitation of our review stems from the shift in diagnostic status of panic disorder and agoraphobia. While DSM-5 lists both as independent (even if frequently comorbid) disorders (American Psychiatric Association, 2013), all studies included in our review rely on diagnoses according to DSM-IV-R or older, in which agoraphobia was considered secondary to panic disorder when a person suffers from both (American Psychiatric Association, 2000). Research has demonstrated that both are indeed independent in many cases (Wittchen et al., 2008, 2010). Thus, the mixed diagnosis of panic disorder with agoraphobia, which was used in the majority of studies in our review, could possibly hide important differences in the clinical presentation of the participants. This issue is especially relevant because in-vivo exposure therapy—the focus of our review—is aimed largely at agoraphobic avoidance, whereas panic disorder without agoraphobia is better addressed by cognitive interventions or specific interoceptive exposure exercises (Ruhmland & Margraf, 2001). Consequently, a replication of our findings should be attempted once there are enough studies based on the DSM-5 distinction between agoraphobia and panic disorder. Nonetheless, it is mostly the assumed relationship between panic and agoraphobia which has changed with the revisions of the DSM, with the criteria for agoraphobia as a syndrome/qualifier or disorder relatively constant (Wittchen et al., 2010). Thus, one can assume that participants of different studies included in our review differ in the amount of panic they experienced, but are rather similar in their agoraphobic features.

Further, the conclusions from the present review are limited by some uncertainties concerning how treatment was implemented in the included studies. While in most cases, treatment was delivered based on treatment manuals, adherence ratings have rarely been performed. Even on a conceptual level, some ambiguity remains. As such, it is to debate whether therapist encouragement (an important part of “guided mastery” exposure therapy) constitutes a cognitive intervention. Likewise, many behavioral treatments include psychoeducation about the disorder, parts of which could be seen as cognitive interventions (e.g., telling patients that one does not die of panic attacks). Thus, we cannot ultimately know how “purely” exposure therapy was implemented in the treatments we classified as pure exposure. Nonetheless, both reports on the therapy time spent in exposure exercises as opposed to other therapeutic interventions and the fact that a substantial number of studies in our review explicitly contrasted a pure exposure condition with conditions including cognitive interventions support our interpretation that exposure was used in a relatively isolated way in those studies.

In summary, the present review and meta-analysis supports an important role of both negative (dysfunctional) and especially positive (self-efficacy-related) cognitions as mechanisms of exposure therapy in agoraphobia. Future research should aim to corroborate these findings with more longitudinal designs in combination with randomized controlled trials, frequent measurement points and modern statistical analyses. Ultimately, a meta-analysis of mediation studies will be an important goal. Clinically, our findings suggest that drawing increased self-efficacy from exposure exercises could be at least equally important as reducing the expectancy of harm. Thus, clinicians might want to monitor patients’ self-efficacy in the process of exposure therapy and consider changes in their approach if self-efficacy remains low. Specific cognitive interventions to strengthen self-efficacy before, during or after exposure exercises could be a promising add-on to the existing interventions aimed at reducing dysfunctional cognitions or maximizing inhibitory learning and are a promising topic for future research.

Role of funding sources

This work was supported in part through a dissertation grant from Heinrich Böll Stiftung, Germany, to CB in the form of funds for personal and material expenses. HBS had no role in the study design, collection, analysis or interpretation of the data, writing the manuscript, or the decision to submit the paper for publication.

Conflicts of interest

All authors declare that they have no conflicts of interest.

Acknowledgements

The authors wish to thank Ann-Marie Küchler for performing a preliminary literature search and overview of the topic, and Johannes Rath for background research on relevant questionnaires and help with data management.

References

References included in the systematic review (either in the meta-analysis or the narrative review) are marked with an.

American Psychiatric Association (2000). *DSM-IV-TR: Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC, US: American Psychiatric Publishing text revision.

American Psychiatric Association (2013). (5th ed.). *Diagnostic and statistical manual of mental disordersVol. 5*. Washington, DC, US: American Psychiatric Publishing.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173.

Becker, B. J. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology, 41*(2), 257–278. https://doi.org/10.1111/j.2044-8317.1988.tb00901.x.

*de Beurs, E., van Balkom, A. J. L. M., Lange, A., Koele, P., & van Dyck, R. (1995). Treatment of panic disorder with agoraphobia: Comparison of fluvoxamine, placebo, and psychological panic management combined with exposure and of exposure in vivo alone. *American Journal of Psychiatry, 152*(5), 683–691. https://doi.org/10.1176/ajp.152.5.683.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* Chichester, UK: John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470743386.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods, 1*(2), 97–111. https://doi.org/10.1002/jrsm.12.

*Bouchard, S., Gauthier, J., Laberge, B., French, D., Pelletier, M.-H., & Godbout, C. (1996). Exposure versus cognitive restructuring in the treatment of panic disorder with agoraphobia. *Behaviour Research and Therapy, 34*(3), 213–224. https://doi.org/10.1016/0005-7967(95)00077-1.

*Bouchard, S., Gauthier, J., Nouwen, A., Ivers, H., Vallières, A., Simard, S., et al. (2007). Temporal relationship between dysfunctional beliefs, self-efficacy and panic apprehension in the treatment of panic disorder with agoraphobia. *Journal of Behavior Therapy and Experimental Psychiatry, 38*(3), 275–292. https://doi.org/10.1016/j.jbtep.2006.08.002.

*Burke, M., Drummond, L. M., & Johnston, D. W. (1997). Treatment choice for agoraphobic women: Exposure or cognitive-behaviour therapy? *British Journal of Clinical Psychology, 36*(3), 409–420. https://doi.org/10.1111/j.2044-8260.1997.tb01248.x.

*Casey, L. M., Newcombe, P. A., & Oei, T. P. S. (2005). Cognitive mediation of panic severity: The role of catastrophic misinterpretation of bodily sensations and panic self-efficacy. *Cognitive Therapy and Research, 29*(2), 187–200. https://doi.org/10.1007/s10608-005-3164-3.

Casey, L. M., Oei, T. P. S., & Newcombe, P. A. (2004). An integrated cognitive model of panic disorder: The role of positive and negative cognitions. *Clinical Psychology Review, 24*(5), 529–555. https://doi.org/10.1016/j.cpr.2004.01.005.

*Casey, L. M., Oei, T. P. S., & Newcombe, P. A. (2005). Looking beyond the negatives: A

time period analysis of positive cognitions, negative cognitions, and working alliance in cognitive–behavior therapy for panic disorder. *Psychotherapy Research, 15*(1–2), 55–68. https://doi.org/10.1080/10503300512331327038.

Cheavens, J. S., Strunk, D. R., Lazarus, S. A., & Goldstein, L. A. (2012). The compensation and capitalization models: A test of two approaches to individualizing the treatment of depression. *Behaviour Research and Therapy, 50*(11), 699–706. https://doi.org/10.1016/j.brat.2012.08.002.

Clark, D. M. (1986). A cognitive approach to panic. *Behaviour Research and Therapy, 24*(4), 461–470. https://doi.org/10.1016/0005-7967(86)90011-2.

Craske, M. G., & Barlow, D. H. (2007). *Mastery of your anxiety and panic: Therapist guide (treatments that work)* (4th ed.). New York: Oxford University Press.

Craske, M. G., Kircanski, K., Zelikowsky, M., Mystkowski, J., Chowdhury, N., & Baker, A. (2008). Optimizing inhibitory learning during exposure therapy. *Behaviour Research and Therapy, 46*(1), 5–27. https://doi.org/10.1016/j.brat.2007.10.003.

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy, 58*, 10–23. https://doi.org/10.1016/j.brat.2014.04.006.

Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology, 62*(1), 583–619. https://doi.org/10.1146/annurev.psych.093008.100356.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*, 629–634. https://doi.org/10.1136/bmj.315.7109.629.

Fentz, H. N., Arendt, M., O'Toole, M. S., Hoffart, A., & Hougaard, E. (2014). The mediational role of panic self-efficacy in cognitive behavioral therapy for panic disorder: A systematic review and meta-analysis. *Behaviour Research and Therapy, 60*, 23–33. https://doi.org/10.1016/j.brat.2014.06.003.

*Fentz, H. N., Hoffart, A., Jensen, M. B., Arendt, M., O'Toole, M. S., Rosenberg, N. K., et al. (2013). Mechanisms of change in cognitive behaviour therapy for panic disorder: The role of panic self-efficacy and catastrophic misinterpretations. *Behaviour Research and Therapy, 51*(9), 579–587. https://doi.org/10.1016/j.brat.2013.06.002.

Ferrer, E., & McArdle, J. J. (2010). Longitudinal modeling of developmental changes in psychological research. *Current Directions in Psychological Science, 19*(3), 149–154. https://doi.org/10.1177/0963721410370300.

Flückiger, C., Caspar, F., Holtforth, M. G., & Willutzki, U. (2009). Working with patients' strengths: A microprocess approach. *Psychotherapy Research, 19*(2), 213–223. https://doi.org/10.1080/10503300902755300.

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin, 99*(1), 20–35. https://doi.org/10.1037/0033-2909.99.1.20.

*Gallagher, M. W., Payne, L. A., White, K. S., Shear, K. M., Woods, S. W., Gorman, J. M., et al. (2013). Mechanisms of change in cognitive behavioral therapy for panic disorder: The unique effects of self-efficacy and anxiety sensitivity. *Behaviour Research and Therapy, 51*(11), 767–777. https://doi.org/10.1016/j.brat.2013.09.001.

*Gloster, A. T., Klotsche, J., Gerlach, A. L., Hamm, A. O., Ströhle, A., Gauggel, S., et al. (2014). Timing matters: Change depends on the stage of treatment in cognitive behavioral therapy for panic disorder with agoraphobia. *Journal of Consulting and Clinical Psychology, 82*(1), 141–153. https://doi.org/10.1037/a0034555.

Gloster, A. T., Wittchen, H.-U., Einsle, F., Lang, T., Helbig-Lang, S., Fydrich, T., et al. (2011). Psychological treatment for panic disorder with agoraphobia: A randomized controlled trial to examine the role of therapist-guided exposure in situ in CBT. *Journal of Consulting and Clinical Psychology, 79*(3), 406–420. https://doi.org/10.1037/a0023584.

Goodwin, R. D., Faravelli, C., Rosi, S., Cosci, F., Truglia, E., de Graaf, R., et al. (2005). The epidemiology of panic disorder and agoraphobia in Europe. *European Neuropsychopharmacology, 15*(4), 435–443. https://doi.org/10.1016/j.euroneuro.2005.04.006.

Gould, R. A., Ott, M. W., & Pollack, M. H. (1995). A meta-analysis of treatment outcome for panic disorder. *Clinical Psychology Review, 15*(8), 819–844. https://doi.org/10.1016/0272-7358(95)00048-8.

*Hahlweg, K., Fiegenbaum, W., Frank, M., Schroeder, B., & von Witzleben, I. (2001). Short- and long-term effectiveness of an empirically supported treatment for agoraphobia. *Journal of Consulting and Clinical Psychology, 69*(3), 375–382. https://doi.org/10.1037/0022-006X.69.3.375.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational and Behavioral Statistics, 6*(2), 107–128. https://doi.org/10.3102/10769986006002107.

*Hoffart, A. (1995a). A comparison of cognitive and guided mastery therapy of agoraphobia. *Behaviour Research and Therapy, 33*(4), 423–434.

*Hoffart, A. (1995b). Cognitive mediators of situational fear in agoraphobia. *Journal of Behavior Therapy and Experimental Psychiatry, 26*(4), 313–320. https://doi.org/10.1016/0005-7916(95)00043-7.

*Hoffart, A. (1998). Cognitive and guided mastery therapy of agoraphobia: Long-term outcome and mechanisms of change. *Cognitive Therapy and Research, 22*(3), 195–207. https://doi.org/10.1023/A:1018736624372.

*Hoffart, A. (2016). Cognitive models for panic disorder with agoraphobia: A study of disaggregated within-person effects. *Journal of Consulting and Clinical Psychology, 84*(9), 839–844. https://doi.org/10.1037/ccp0000114.

*Hoffart, A., Hedley, L. M., Svanøe, K., & Sexton, H. (2016). Cognitive and guided mastery therapies for panic disorder with agoraphobia: 18-year long-term outcome and predictors of long-term change. *Clinical Psychology & Psychotherapy, 23*(1), 1–13. https://doi.org/10.1002/cpp.1934.

*Hoffart, A., Sexton, H., Hedley, L. M., & Martinsen, E. W. (2008). Mechanisms of change in cognitive therapy for panic disorder with agoraphobia. *Journal of Behavior Therapy and Experimental Psychiatry, 39*(3), 262–275. https://doi.org/10.1016/j.jbtep.2007.07.006.

Hofmann, S. G. (2008a). Cognitive processes during fear acquisition and extinction in animals and humans: Implications for exposure therapy of anxiety disorders. *Clinical Psychology Review, 28*(2), 199–210. https://doi.org/10.1016/j.cpr.2007.04.009.

Hofmann, S. G. (2008b). Common misconceptions about cognitive mediation of treatment change: A commentary to Longmore and Worrell (2007). *Clinical Psychology Review, 28*(1), 67–70. https://doi.org/10.1016/j.cpr.2007.03.003.

Hofmann, S. G., Meuret, A. E., Rosenfield, D., Suvak, M. K., Barlow, D. H., Gorman, J. M., et al. (2007). Preliminary evidence for cognitive mediation during cognitive-behavioral therapy of panic disorder. *Journal of Consulting and Clinical Psychology, 75*(3), 374–379. https://doi.org/10.1037/0022-006X.75.3.374.

van den Hout, M., Arntz, A., & Hoekstra, R. (1994). Exposure reduced agoraphobia but not panic, and cognitive therapy reduced panic but not agoraphobia. *Behaviour Research and Therapy, 32*(4), 447–451. https://doi.org/10.1016/0005-7967(94)90008-6.

*van Hout, W. J. P. J., Emmelkamp, P. M. G., & Scholing, A. (1994). The role of negative self-statements during exposure in vivo: A process study of eight panic disorder patients with agoraphobia. *Behavior Modification, 18*(4), 389–410. https://doi.org/10.1177/01454455940184002.

*Ito, L. M., Araujo, L. A. De, Tess, V. L. C., de Barros-Neto, T. P., Asbahr, F. R., & Marks, I. (2001). Self-exposure therapy for panic disorder with agoraphobia: Randomised controlled study of external v. interoceptive self-exposure. *British Journal of Psychiatry, 178*(4), 331–336. https://doi.org/10.1192/bjp.178.4.331.

Jansson, L., & Öst, L.-G. (1982). Behavioral treatments for agoraphobia: An evaluative review. *Clinical Psychology Review, 2*(3), 311–336. https://doi.org/10.1016/0272-7358(82)90017-4.

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology, 3*(1), 1–27. https://doi.org/10.1146/annurev.clinpsy.3.022806.091432.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Archives of General Psychiatry, 62*(6), 593–602. https://doi.org/10.1001/archpsyc.62.6.593.

Kessler, R. C., Chiu, W. T., Demler, O., Merikangas, K. R., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the national comorbidity survey replication. *Archives of General Psychiatry, 62*(6), 617–627. https://doi.org/10.1001/archpsyc.62.6.617.

Kessler, R. C., Chiu, W. T., Jin, R., Ruscio, A. M., Shear, K., & Walters, E. E. (2006). The epidemiology of panic attacks, panic disorder, and agoraphobia in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 63*(4), 415–424. https://doi.org/10.1001/archpsyc.63.4.415.

Khawaja, N. G., & Oei, T. P. S. (1998). Catastrophic cognitions in panic disorder with and without agoraphobia. *Clinical Psychology Review, 18*(3), 341–365. https://doi.org/10.1016/S0272-7358(97)00088-3.

Kraemer, H. C., Kiernan, M., Essex, M., & Kupfer, D. J. (2008). How and why criteria defining moderators and mediators differ between the Baron & Kenny and MacArthur approaches. *Health Psychology, 27*(2), S101–S108. https://doi.org/10.1037/0278-6133.27.2(Suppl.).S101.

Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry, 59*(10), 877. https://doi.org/10.1001/archpsyc.59.10.877.

Loerinc, A. G., Meuret, A. E., Twohig, M. P., Rosenfield, D., Bluett, E. J., & Craske, M. G. (2015). Response rates for CBT for anxiety disorders: Need for standardized criteria. *Clinical Psychology Review, 42*, 72–82. https://doi.org/10.1016/j.cpr.2015.08.004.

Longmore, R. J., & Worrell, M. (2007). Do we need to challenge thoughts in cognitive behavior therapy? *Clinical Psychology Review, 27*(2), 173–187. https://doi.org/10.1016/j.cpr.2006.08.001.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7*(1), 83–104. https://doi.org/10.1037//1082-989X.7.1.83.

*Malbos, E., Rapee, R. M., & Kavakli, M. (2013). A controlled study of agoraphobia and the independent effect of virtual reality exposure therapy. *Australian and New Zealand Journal of Psychiatry, 47*(2), 160–168. https://doi.org/10.1177/0004867412453626.

*Marchand, A., Coutu, M.-F., Dupuis, G., Fleet, R., Borgeat, F., Todorov, C., et al. (2008). Treatment of panic disorder with agoraphobia: Randomized placebo-controlled trial of four psychosocial treatments combined with imipramine or placebo. *Cognitive Behaviour Therapy, 37*(3), 146–159. https://doi.org/10.1080/16506070701743120.

Meuret, A. E., Hofmann, S. G., & Rosenfield, D. (2010). Catastrophic appraisal and perceived control as moderators of treatment response in panic disorder. *International Journal of Cognitive Therapy, 3*(3), 262–277. https://doi.org/10.1521/ijct.2010.3.3.262.

Meuret, A. E., Seidel, A., Rosenfield, B., Hofmann, S. G., & Rosenfield, D. (2012). Does fear reactivity during exposure predict panic symptom reduction? *Journal of Consulting and Clinical Psychology, 80*(5), 773–785. https://doi.org/10.1037/a0028032.

*Michelson, L. K., Marchione, K. E., Greenwald, M., Testa, S., & Marchione, N. J. (1996). A comparative outcome and follow-up investigation of panic disorder with agoraphobia: The relative and combined efficacy of cognitive therapy, relaxation training, and therapist-assisted exposure. *Journal of Anxiety Disorders, 10*(5), 297–330. https://doi.org/10.1016/0887-6185(96)00001-1.

Mitte, K. (2005). A meta-analysis of the efficacy of psycho- and pharmacotherapy in panic disorder with and without agoraphobia. *Journal of Affective Disorders, 88*(1), 27–45. https://doi.org/10.1016/j.jad.2005.05.003.

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis

with repeated measures and independent-groups designs. *Psychological Methods, 7*(1), 105–125. https://doi.org/10.1037/1082-989X.7.1.105.

*Pelissolo, A., Zaoui, M., Aguayo, G., Yao, S. N., Roche, S., Ecochard, R., et al. (2012). Virtual reality exposure therapy versus cognitive behavior therapy for panic disorder with agoraphobia: A randomized comparison study. *Journal of CyberTherapy and Rehabilitation, 5*(1), 35–43.

Pompoli, A., Furukawa, T. A., Imai, H., Tajika, A., Efthimiou, O., & Salanti, G. (2016). Psychological therapies for panic disorder with or without agoraphobia in adults: A network meta-analysis. *Cochrane Database of Systematic Reviews, 4*https://doi.org/10.1002/14651858.CD011004.pub2.

Porter, E., & Chambless, D. L. (2015). A systematic review of predictors and moderators of improvement in cognitive-behavioral therapy for panic disorder and agoraphobia. *Clinical Psychology Review, 42*, 179–192. https://doi.org/10.1016/j.cpr.2015.09.004.

Powers, M. B., & Emmelkamp, P. M. G. (2008). Virtual reality exposure therapy for anxiety disorders: A meta-analysis. *Journal of Anxiety Disorders, 22*(3), 561–569. https://doi.org/10.1016/j.janxdis.2007.04.006.

Preacher, K. J. (2015). Advances in mediation analysis: A survey and synthesis of new developments. *Annual Review of Psychology, 66*(1), 825–852. https://doi.org/10.1146/annurev-psych-010814-015258.

R Core Team (2017). *R: A language and environment for statistical computing. Vienna, Austria*.

Rief, W., Glombiewski, J. A., Gollwitzer, M., Schubö, A., Schwarting, R., & Thorwart, A. (2015). Expectancies as core features of mental disorders. *Current Opinion in Psychiatry, 28*(5), 378–385. https://doi.org/10.1097/YCO.0000000000000184.

Ruhmland, M., & Margraf, J. (2001). Effektivität psychologischer Therapien von Panik und Agoraphobie: Meta-Analysen auf Störungsebene. *Verhaltenstherapie, 11*(1), 41–53. https://doi.org/10.1159/000050323.

Rupp, C., Doebler, P., Ehring, T., & Vossbeck-Elsebusch, A. N. (2016). Emotional processing theory put to test: A meta- analysis on the association between process and outcome measures in exposure therapy. *Clinical Psychology & Psychotherapy, 24*(3), 697–711. https://doi.org/10.1002/cpp.2039.

Sánchez-Meca, J., Rosa-Alcázar, A. I., Marín-Martínez, F., & Gómez-Conesa, A. (2010). Psychological treatment of panic disorder with or without agoraphobia: A meta-analysis. *Clinical Psychology Review, 30*(1), 37–50. https://doi.org/10.1016/j.cpr.2009.08.011.

*Smits, J. a J., Powers, M. B., Cho, Y., & Telch, M. J. (2004). Mechanism of change in cognitive-behavioral treatment of panic disorder: Evidence for the fear of fear mediational hypothesis. *Journal of Consulting and Clinical Psychology, 72*(4), 646–652. https://doi.org/10.1037/0022-006X.72.4.646.

*Southworth, S., & Kirsch, I. (1988). The role of expectancy in exposure-generated fear reduction in agoraphobia. *Behaviour Research and Therapy, 26*(2), 113–120. https://doi.org/10.1016/0005-7967(88)90110-6.

*Telch, M. J., Agras, W. S., Taylor, C. B., Roth, W. T., & Gallen, C. C. (1985). Combined pharmacological and behavioral treatment for agoraphobia. *Behaviour Research and Therapy, 23*(3), 325–335. https://doi.org/10.1016/0005-7967(85)90011-7.

Telch, M. J., Brouillard, M., Telch, C. F., Agras, W. S., & Taylor, C. B. (1989). Role of cognitive appraisal in panic-related avoidance. *Behaviour Research and Therapy, 27*(4), 373–383. https://doi.org/10.1016/0005-7967(89)90007-7.

Tryon, W. W. (2005). Possible mechanisms for why desensitization and exposure therapy work. *Clinical Psychology Review, 25*(1), 67–95. https://doi.org/10.1016/j.cpr.2004.08.005.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48.

*Vögele, C., Ehlers, A., Meyer, A. H., Frank, M., Hahlweg, K., & Margraf, J. (2010). Cognitive mediation of clinical improvement after intensive exposure therapy of agoraphobia and social phobia. *Depression and Anxiety, 27*(3), 294–301. https://doi.org/10.1002/da.20651.

Weisman, J. S., & Rodebaugh, T. L. (2018). Exposure therapy augmentation: A review and extension of techniques informed by an inhibitory learning approach. *Clinical Psychology Review, 59*, 41–51. https://doi.org/10.1016/j.cpr.2017.10.010.

Williams, S. L. (1990). Guided mastery treatment of agoraphobia: Beyond stimulus exposure. *Progress in Behavior Modification, 26*, 89–121.

*Williams, S. L., & Falbo, J. (1996). Cognitive and performance-based treatments for panic attacks in people with varying degrees of agoraphobic disability. *Behaviour Research and Therapy, 34*(3), 253–264. https://doi.org/10.1016/0005-7967(95)00063-1.

*Williams, S. L., Kinney, P. J., & Falbo, J. (1989). Generalization of therapeutic changes in agoraphobia: The role of perceived self-efficacy. *Journal of Consulting and Clinical Psychology, 57*(3), 436–442. https://doi.org/10.1037/0022-006X.57.3.436.

*Williams, S. L., & Zane, G. (1989). Guided mastery and stimulus exposure treatments for severe performance anxiety in agoraphobics. *Behaviour Research and Therapy, 27*(3), 237–245. https://doi.org/10.1016/0005-7967(89)90042-9.

Williams, S. L., & Zane, G. (1997). Guided mastery treatment of phobias. *Clinical Psychologist, 50*, 13–15.

Wilson, G. T. (1999). Rapid response to cognitive behavior therapy. *Clinical Psychology: Science and Practice, 6*(3), 289–292. https://doi.org/10.1093/clipsy.6.3.289.

Wittchen, H.-U., Gloster, A. T., Beesdo-Baum, K., Fava, G. A., & Craske, M. G. (2010). Agoraphobia: A review of the diagnostic classificatory position and criteria. *Depression and Anxiety, 27*(2), 113–133. https://doi.org/10.1002/da.20646.

Wittchen, H.-U., Nocon, A., Beesdo, K., Pine, D. S., Hofler, M., Lieb, R., et al. (2008). Agoraphobia and panic. Prospective-longitudinal relations suggest a rethinking of diagnostic concepts. *Psychotherapy and Psychosomatics, 77*(3), 147–157. https://doi.org/10.1159/000116608.

Wittchen, H.-U., Reed, V., & Kessler, R. C. (1998). The relationship of agoraphobia and panic in a community sample of adolescents and young adults. *Archives of General Psychiatry, 55*(11), 1017–1024. https://doi.org/10.1001/archpsyc.55.11.1017.

*Zane, G., & Williams, S. L. (1993). Performance-related anxiety in agoraphobia: Treatment procedures and cognitive mechanisms of change. *Behavior Therapy, 24*(4), 625–643. https://doi.org/10.1016/S0005-7894(05)80322-3.