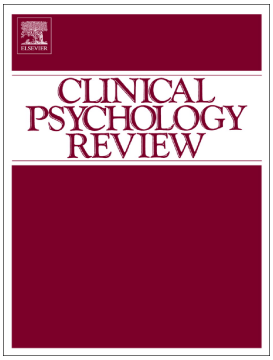


Self-reported cognitive biases in depression: A meta-analysis

Ines Nieto, Elena Robles, Carmelo Vazquez

PII: S0272-7358(20)30122-7

DOI: https://doi.org/10.1016/j.cpr.2020.101934

Reference: CPR 101934

To appear in: *Clinical Psychology Review*

Received date: 16 April 2020

Revised date: 18 September 2020

Accepted date: 13 October 2020

Please cite this article as: I. Nieto, E. Robles and C. Vazquez, Self-reported cognitive biases in depression: A meta-analysis, *Clinical Psychology Review* (2020), https://doi.org/10.1016/j.cpr.2020.101934

Self-reported Cognitive Biases in Depression: A Meta-analysis

Ines Nieto[a], Elena Robles[a], & Carmelo Vazquez[a,*]

[a]Department of Clinical Psychology, Complutense University of Madrid

Acknowledgements: This work was partially supported by national grants PSI2015-69253-R and PID2019-108711GB-I00 to CV and CT17/17-CT18/17 UCM predoctoral fellowhip to IN. We also thank Jamie O'Grady for his help in editing the paper and Jonas Everaert for his technical advice when conducting this meta-analysis.

Declaration of interest: none

Trial registration: This systematic review and meta-analysis was prospectively registered on the 20th of November 2018 with PROSPERO number CRD42018115365.

*__Corresponding author:__
Carmelo Vázquez
School of Psychology
Complutense University of Madrid
28223 Madrid (Spain)
E-mail: cvazquez@psi.ucm.es

Abstract

Despite the influence of Beck's cognitive models of depression, the presence and magnitude of the specific proposed cognitive biases have not been systematically investigated. After a systematic search in PsycInfo and PubMED, studies reporting self-reported outcomes on cognitive biases and depressive symptoms in depressed and/or healthy groups were included. From a total of 4,840 records, two different meta-analyses were conducted. 23 studies on 4,865 participants provided data about catastrophising and depression (g = 0.95, 95% CI [0.64; 1.26]) and 40 studies on 4,678 participants provided data about interpretation bias in depression (g = 0.78, 95% CI [0.43; 1.13]). Moderation analyses showed that the relationship between catastrophising and depression was higher in studies with more women, when the corresponding author was from a Western country, and when the instrument to measure depression was the DSM criteria, the SCL-90, the BDI, or the DASS. The relationship between interpretation bias and depressive symptoms was significant only in studies comparing depressed and healthy groups, and when using specific instruments to measure symptoms (DSM/RDC criteria plus a scale cut-off score) and cognitive bias (CDQ/CBQ, SCT, AST-D, other). Some limitations are acknowledged, but risk of publication bias was found to be low, and these results support the utility of some self-reported measures of cognitive biases in depression.

Keywords: cognitive bias; depression; cognitive model; catastrophizing; interpretation bias

Introduction

Depression is considered one of the leading causes of disability in the world and is associated with great social and economic costs (Whiteford et al., 2013). More than 200 million people suffer from major depression (World Health Organization, 2017) and the relapse rate is around 85% (McIntyre & O'Donovan, 2004). Given the impact of this disorder, it is important to understand the variables that influence its development, maintenance and recurrence. Several theories have been proposed to explain the causes of this disorder (Gotlib & Hammen, 2014). Among them, cognitive theories have been very influential both in research and applied fields. These theories point out to dysfunctional thinking as a key causal factor related to the onset and maintenance of these emotional disorders (Beck, 1967, 1976; Barlow et al., 2011).

Beck's seminal cognitive model of depression (1967, 1976) provided some of the concepts and explanatory heuristics that have been incorporated in most of the current cognitive models of depression. According to this model (see Figure 1), individuals may develop, early in life, latent cognitive schemas that get activated when facing environmental stressors that are pertinent to the contents of those schemas (e.g., themes of loss). Cognitive schemas are rather abstract representations of the world and determine the way in which information is processed and how events and stimuli are interpreted in a given context (Dalgleish & Power, 2000). When the negative contents of these cognitive schemas are triggered by internal or external events, psychological processes (i.e., memory, interpretation, and attention) operate following negative cognitive biases, like dichotomous thinking or arbitrary inferences (see Table 1). These processes then lead to biased mental products or thoughts about the self, the world and the future (i.e. the cognitive triad). It is important to note the difference between cognitive schemas and cognitive biases. Following Ingram and

Kendall's (1986) cognitive taxonomy, cognitive schemas would be structural variables of the system (i.e., broad cognitive frames through which information is filtered, represented and organized), whereas cognitive biases would function as operational variables, or mechanisms, by which cognitive structures work. Although schemas would be the most distal causes of depression (Panzarella, Alloy & Whitehouse, 2006), biased cognitive operations would also have a critical causal role as they are the action mechanisms used to support and validate the schemas. Overall, this cognitive machinery might be considered as an antecedent of the presence of negative automatic thoughts, negative beliefs (about oneself, the world and/or the future), and ultimately depressive symptoms, which would all be the products (i.e., the tangible outputs) with which clinicians typically work with their clients.

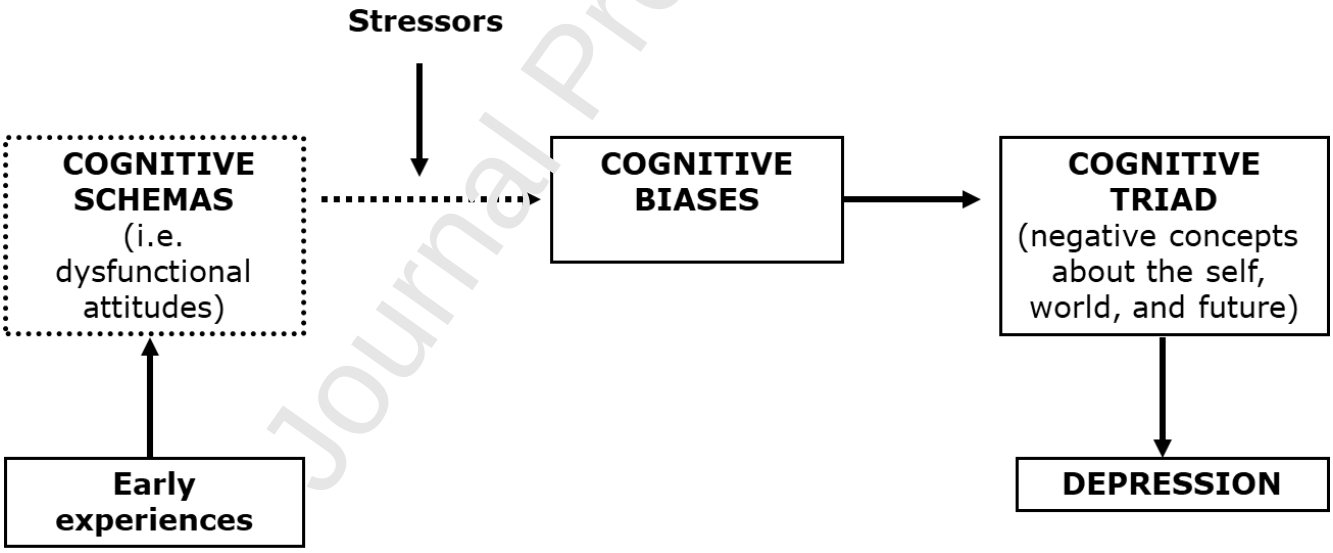


Figure 1. Beck's cognitive model. See further descriptions in the text.

Beck's theoretical account has had a profound impact on both clinical and research fields. For instance, cognitive-behavior therapy (CBT) for depression follows the rationale of the cognitive model (Beck et al., 1979; Greenberg & Padesky, 2015). One of the main aims in CBT is to modify biased thinking (Beck, 1976) as a path to improve symptoms such as

negative affect and anhedonia (Joorman & Gotlib, 2010). Also, extensive research on cognition and depression has analysed the relationships between these constructs and depressive symptoms in a variety of samples regarding culture, gender, age and diagnostic severity (e.g. Hu et al., 2016; Cristea et al., 2015; Vîslă et al., 2016). Moreover, Beck's proposal has inspired the creation of widely used instruments to measure both symptoms (e.g., the Beck Depression Inventory-II, BDI-II, Beck, Steer, & Brown, 1996) and cognitive constructs like the Cognitive Errors Questionnaire (CEQ, Lefebvre, 1981) or the Cognitive Bias Questionnaire (CBQ, Krantz & Hammen, 1979). In sum, Beck's model has been extraordinarily influential both in research and applied fields in clinical psychology.

The role of cognitive biases in Beck's model

Beck's original theory has been further developed during recent decades. Empirical research has supported and refined the model by clarifying the nature of stressors contributing to form early negative schemas (Hammen, 2005), distinguishing different subtypes of schemas (Clark & Beck, 1991), or lately adding new mechanistic processes, like rumination (Nolen-Hoeksema, Wisco & Lyubomirsky 2008) or overgeneral autobiographical memory (Williams et al., 2007), that have been incorporated in the depression literature for some decades after Beck's initial formulation. Likewise, there have been efforts to inquire into the neurological basis of Beck's theoretical account (Beck, 2008). For instance, researchers have found that cognitive biases seem to be characterized by a maladaptive bottom-up system at the subcortical level, which is reinforced by an attenuated cortical cognitive control unable to regulate them (Disner et al., 2011).

Despite the relevance of Beck's cognitive model of depression, the specific cognitive biases proposed in it (see Table 1) have not been subjected to a systematic review, except for

causal attribution bias (Hu, Zhang & Yang, 2015). Although the model is still inspiring research on cognitive processes (Gotlib & Joormann, 2010) and has become one of the best validated and most frequently used therapeutic interventions (Cuijpers et al., 2019), the validity of the cognitive biases explicitly formulated in the model is relatively unknown.

Table 1. *Cognitive biases of depression according to Beck's cognitive model.*

Cognitive Biases	Definition	Example
Catastrophizing or Fortune telling	Anticipating negative extreme events and outcomes.	"I am going out of my mind"
Personalization or Internal causal attribution	Always interpreting events from a guilt-point of view. Exaggerated guilt.	"He is disgusted by me" "When something goes wrong, it is my fault"
Dichotomous thinking	Extreme thinking where things are interpreted like all-or-nothing, black-or-white.	"If I can´t do it perfect, I won´t do it at all"
Selective abstraction	Paying attention in a selective manner to negative events or outcomes and missing the rest of the information.	"She only said the job was 'good'"
Arbitrary inference	Making negative interpretations with no evidence that supports them, or with contrary evidence.	"It will go bad for sure"
Overgeneralization	Making general interpretations and conclusions taking into account only one piece of information.	"Everything will always be this bad"
Minimization/Discounting positives	Disqualifying positive events or experiences as non-relevant	"Winning the price was not that good"
Labelling	Naming negatively the self and others based on the occurrence of just one mistake.	"I am a loser"
Mindreading	Systematically thinking that others are reacting negatively to you.	"He thinks I am a failure"
Emotional reasoning	Assuming negative emotions represent how things are in reality.	"I feel guilty. Therefore, I must have done something bad"

Note. These categories of cognitive bias have been proposed in different books and articles by Beck and collaborators. For the present meta-analysis the following documents were revised: Beck (1963, 1967, 1976, 1979) and Burns (1980).

Whereas some meta-analyses have approached the study of cognitive constructs such as irrational beliefs, dysfunctional thinking, or automatic thoughts (Cristea et al., 2015; Vîslă et al., 2016), the magnitude of most cognitive biases in depressed individuals has not been synthesised in the literature. Also, it remains unclear whether these processes are exclusive for depressed individuals, which makes it important to study their relationship with depressive symptoms itself (dimensional studies) but also to compare different types of populations (categorical studies). Thus, the aim of this study was to conduct a meta-analysis of those cognitive biases of depression, as specified in Beck's cognitive model, that have not been meta-analysed in previous studies (i.e. those presented in Table 1 except for Personalization/Internal causal attributions). Random-effect models were used (since sampling variability was expected, Riley, Higgins, & Deeks, 2011) to meta-analyse each cognitive bias, using the standardized mean difference with Hedge's correction (*g*) as the effect size (ES). The meta-analysis was pre-registered in PROSPERO (CRD42018115365). The first hypothesis was that there would be significantly larger levels of cognitive biases in depression groups than in other comparison groups, such as healthy participants, or individuals with other psychological symptoms, or subclinical levels of depression. Also, it was expected that several variables could moderate this effect. The year of the study and geographic location of the corresponding author were coded to study potential spatial and temporal effects. Based on previous findings, sample characteristics such as age (Reed, Chan, & Mikels, 2014), gender (Kessler & Bromet, 2013), or type of sample (Clark, Beck & Alford, 1999) were investigated although the direction of these potential moderation effects was not anticipated. Finally, methodological variables such as sample size, type of measure, and the

psychometric characteristics of the instruments were also included as potential moderators given previous meta-analytic results (Everaert, Podina, & Koster, 2017).

Method

Eligibility and search criteria

A systematic search was conducted on the databases of PsycINFO and PubMed, until February 2020, combining terms related to the spectrum of depression (depress* OR dysphor* OR mood OR “affective disorder” OR “sad mood” OR sadness), comparison groups based on DSM-5 (American Psychology Association, 2013) categories and type of population (delirium OR dementia OR "neurodevelopmental disorder" OR schizophrenia OR "psychotic disorder" OR "delusional disorder" OR "bipolar disorder" OR "anxiety disorder" OR "somatoform disorder" OR "factitious disorder" OR "dissociative disorder" OR "obsessive-compulsive disorder" OR "trauma-related disorder" OR "stressor-related disorder" OR "somatic symptom disorder" OR "eating disorder" OR "elimination disorder" OR "sleep disorder" OR "sexual dysfunction" OR "gender identity disorder" OR "disruptive disorder" OR "impulse-control disorder" OR "conduct disorder" OR "substance-related disorder" OR "addictive disorder" OR "alcohol-use disorder" OR "neurocognitive disorder" OR "personality disorder" OR "paraphilic disorder" OR "intellectual disability" OR "developmental disorder" OR "autistic disorder" OR "oppositional defiant disorder" OR "attention deficit-hyperactivity disorder" OR "normal population" OR "general population" OR student* OR control OR healthy OR group* OR nondepressed OR nondysphoric OR “never depressed” OR compar* OR differ*) and the list of cognitive biases described in Beck’s model of depression: catastrophizing, dichotomous thinking, selective abstraction, arbitrary inference, overgeneralization, minimization, labelling, mindreading, and emotional reasoning. This selection was made after

revising different documents by Beck and collaborators (Beck, 1963, 1967, 1976, 1979; Burns, 1980). Different terms were used to search for the results of each specific cognitive bias (e.g. “polarized thinking” OR “dichotomous thinking” OR “bipolar thinking” OR “all-or-nothing thinking”) given that these constructs have been named heterogeneously in the literature. Moreover, more general terms (‘negative thinking’, ‘interpretation bias’, ‘thinking error’, and ‘cognitive distortion’) were also included in the search to fully cover the research in the area. Samples with physical or medical conditions were excluded. Also, personalization, or causal attribution bias, were not included in the search as it has already been meta-analysed (Hu et al., 2015). Two complementary approaches were used to complete the inclusion of studies. First, all the studies cited in the meta-analyses found in our search were individually inspected to add possible relevant references. Second, a complementary search was performed using the terms (cognitive distortion* OR “cognitive bias” OR cognitive style* OR cognitive intrusion* OR cognitive error* OR maladaptive cognition* OR thinking error* OR heuristic*) AND ("meta-analysis" OR "meta-analyses" OR "review") to find meta-analyses and reviews in the area and inspect the primary studies included.

The inclusion criteria were a) empirical categorical and dimensional studies; b) measuring cognitive biases and depression symptoms; c) with self-report questionnaires; d) in adults (> 18 years); e) published in English; f) peer-reviewed; and g) providing the necessary data to calculate effect sizes. Studies were excluded if they measured cognitive constructs different from cognitive biases (e.g. Dysfunctional Attitude Scale, DAS, Hammen & Krantz, 1976; Automatic Thoughts Questionnaire, ATQ, Hollon & Kendall, 1980), or measured cognitive biases with experimental tasks, interviews or non-quantitative methods. Studies including only samples with comorbidity or with disorders other than depression were also excluded. Multiple measures of the same cognitive bias or depressive symptoms were handled

either by choosing the most common instrument among the included studies (to improve comparison) or by calculating the mean for that result.

Data collection process

The search, selection and codification processes were made by two different PhD students (IN and ER) independently. The inter-rater agreement for the selection of studies was excellent (kappa= 0.80, k=274). All the potential variables relevant to the analyses were coded (see Table 1 in Supplementary Material). The following moderators were included: year of the study, geographic location of the author of correspondence -Western (Europe, North-America, Australia and New Zealand), non-Western (Asia)- sample size, percentage of women in the total sample, type of sample (subclinical or clinical, general population, students), measure of depression symptoms (used to select the groups in categorical studies and to calculate the correlation with cognitive bias in dimensional studies), measure of cognitive bias (no pre-specified categories), number of items, and reliability of the cognitive bias measure. The first and second authors coded all studies independently, being kappa inter-rater reliability 0.77. Disagreements were discussed to reach consensus among all the authors.

Analytic plan

Summary measures

Random-effects meta-analyses for all the included cognitive biases were conducted, in SPSS 20 and R 3.5.0 (metafor package Viechtbauer 2010), using the standardized mean difference ($d= \bar{X}_1-\bar{X}_2/S_{pooled}$) with Hedge's correction ($g=c(m)*d$) as the effect size (ES). Positive values reflect a higher level of cognitive bias in the experimental group compared to the control group. Hedge's *g* values can be categorized as small (0.2–0.5), medium (0.5–0.8), or large (>0.8) (Cohen, 1988). All studies were included in the analyses. Thus, the ES was

calculated based on data comparing groups (depressed group vs. any other condition) and correlations between cognitive biases and quantitative measures of depression symptoms. This strategy allowed us to include a larger number of studies and make the results more reliable. Means, standard deviations, sample sizes and correlations were used to calculate the ES. If between-group data were not available, then reported t-values or between-group F-values and sample sizes were used. For dimensional studies, Pearson correlations were first standardized using Fisher transformation ($Z_r=1/2*\ln*((1+r)/(1-r))$). Then, these values were converted to standardised mean differences ($d=2r/\sqrt{1-r^2}$; $v_d = 4v / (1-r^2)^3$), followed by Hedge's correction (*g).* In all cases the pooled standard deviation was used to calculate *g*, the method to weight studies was the inverse of the sampling variance of g, and confidence intervals (CI) were used as reflection of the estimate precision, both within and between studies.

Homogeneity

Homogeneity was assessed using two indicators: the Q-value of the test of heterogeneity, and the I^2 index. When the Q value has a significance level lower than .05 the null hypothesis of homogeneity is rejected, leading to the conclusion that all the studies are not estimating the same parameter. Regarding the I^2 values, categorization is established as follows: 0% = no heterogeneity; equal to or higher than 25% = low heterogeneity; equal to or higher than 50% = moderate heterogeneity; and equal to or higher than 75% = high heterogeneity (Borenstein et al., 2009ster).

Moderation analyses

For moderation analyses, mixed-effects models were used to test categorical variables and meta-regression (DerSimonian-Laird's estimator) was used to test continuous variables (Borenstein et al., 2009).

Risk of bias

Publication bias was analysed using different methods. First, the Orwin's (1983) fail-safe number (Ns) was calculated to know the number of studies with a medium effect equal to zero that would make the ES in the meta-analysis become non significant. The threat is considered to be real when Ns ≤ (5k + 10). Second, the funnel plot was inspected. It represents the precision of each primary study (standard error, SE) against its individual effect size. Without publication bias, the shape and density should be symmetric. Two different methods can be used to test the symmetry of the funnel plot. Kendall's Rank correlation tests the null hypothesis that the ES and the SE are independent. The Eager's regression test for funnel plot asymmetry (Sterne, Becker, & Egger, 2006) tests the null hypothesis that there is perfect symmetry in the plot (starting point of the regression line equal to 0). Finally, the trim-and-fill procedure (Duval & Tweedie, 2000) is used as a sensitive analysis to calculate the ESs and confidence intervals of the individual studies accounting for the missing values reflected in the asymmetry of the funnel plot.

All the analyses were repeated after removing outliers. These were defined as those primary studies with both sides of their 95% confidence interval outside the 95% confidence interval of the pooled studies.

Quality of studies

The quality of primary studies was assessed using an adaptation of the Downs and Black's checklist (Downs & Black, 1998). This scale has previously been used in a meta-

analysis on interpretation biases (Everaert et al., 2017) and evaluates standards of methodological quality, such as quality of report (e.g., "*Is the hypothesis/aim/objective of the study clearly described?*"), external validity (e.g., "*Were those subjects who were prepared to participate representative of the entire population from which they were recruited?*"), bias and confounding variables (e.g., "*If any of the results of the study were based on "data dredging", was this made clear?",* "*Was there adequate adjustment for confounding in the analyses from which the main findings were drawn? ")*, and statistical power ("*Did the study have sufficient power to detect an effect where the probability value for a difference being due to chance is less than 5%?"*). Ratings were made by the first and second author independently and an inter-rater agreement of kappa = 0.94 was reached. Disagreements were resolved with discussion until reaching consensus. The quality of the meta-analysis itself was also assessed following standard sets of recommendations (Moher et al., 2015; American Psychology Association, 2018).

Results

Study selection

The process of selection and inclusion of studies is shown in Figure 2. From a total of 4,840 records (1,320 duplicates) 3,131 were excluded based on the screening of the title or abstract, while 461 were excluded after a full-text reading. The main reason for exclusion was the lack of a measure of bias (i.e., many studies were focused on constructs different from cognitive biases, such as cognitive schemas or automatic thoughts -see Figure 1). Many studies were also excluded because they did not use self-report questionnaires to measure cognitive biases. Finally, a total of 63 studies were included in the quantitative analyses. 21 studies fulfilled the inclusion criteria but could not be included in the meta-analysis. This was due to the lack of

commonality in the type of comparison group and the type of cognitive bias measured by a large enough number of studies (see Table 2 in Supplementary Materials).

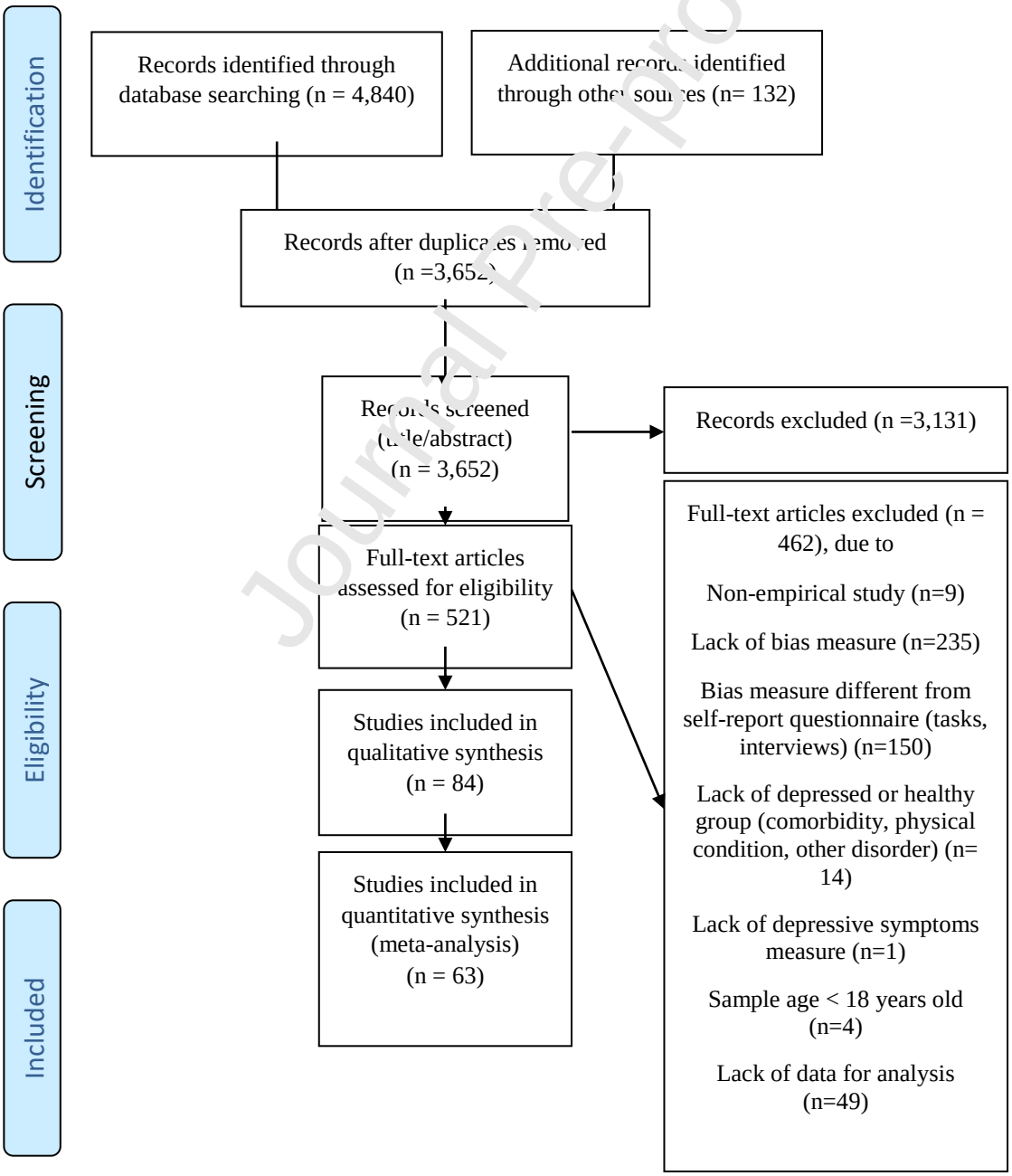


Figure 2. PRISMA flow diagram

Study characteristics

Studies included in the meta-analysis were published from 1979 to 2017, covering almost four decades of research. Total sample size ranged from 19 to 611 participants, with mean percentage of female participants being 64.4% (from 28.1% to 100%) and a mean age of 32.5 years old (18.9-71.6).

From the different searches that were performed (one per specific cognitive bias name plus the different searches with general terms), studies fulfilling all inclusion criteria were only found in five of them: 'catastrophizing', 'negative thinking', 'interpretation bias', 'thinking error' and 'cognitive distortion'. Given the lack of studies in the other categories, the content of the instruments used to measure cognitive bias in the selected studies was examined. It was found that only a small group of studies provided results for a specific type of cognitive bias (i.e. catastrophizing). All the other articles provided results for the general concept of biased cognition. Thus, two different analyses were conducted. The first one (k=23) included results evaluating catastrophizing bias in relation to depression. Some of these studies conceptualized catastrophizing as an emotion regulation strategy and used the

Cognitive Emotion Regulation Questionnaire (CERQ) to measure it; other series of studies conceptualized it as the probability of occurrence of negative events (e.g., Corcoran et al. (2006) and measured catastrophizing by asking participants to predict the likelihood that some negative events happened to them. The second analysis included 40 studies using a broader conceptualization of biased processing. The measurement methods used in this second group of studies was varied, but all of them instructed participants to interpret ambiguous information applied to themselves. For instance, the Cognitive Bias Questionnaire presents potentially problematic daily life situations and participants are asked to select from four response options the one that best represents their interpretation of the situation; the Ambiguous Scenarios Test-D is composed by ambiguous scenarios which participants imagine that they have happened to them and are asked to rate how pleasant they perceive it, based on their interpretation of such scenarios, and the Sentence Completion Test presents participants with short sentence stems formed by agent–verb combinations (e.g. 'I think ...') or simple nouns (e.g. 'the world …') which participants need to complete based on their first interpretation).

Only a small number of categorical studies (i.e. studies with groups or subgroups of individuals) included more than one comparison group, making the analyses comparing depressed participants and other diagnostic groups not possible. Thus, all the analyses were performed based on the comparison of depressed individuals with healthy individuals. The group of studies measuring catastrophizing bias included 7 categorical studies: 6 with a clinical sample and 1 with subclinical individuals (Beck Depression Inventory, BDI, Beck et al., 1996, total score ≥16). The group of studies measuring general cognitive bias included 28 categorical studies: 18 with a clinical sample diagnosed with major depression, 2 with a clinical sample diagnosed with either major depression or dysthymia, 8 with a subclinical

sample (using BDI cut-off scores), and 1 study performing a sad mood induction in the experimental group. All the dimensional studies (k=16 for catastrophizing and k=11 for interpretation bias) were based on healthy, or general population participants, in which depression symptoms were measured with a self-report instrument.

Quality of primary studies

The average quality of primary studies was low (53.34%, SD = 8.83) when compared with Everaert et al. (2017) meta-analytic study on interpretation bias where the average quality, also using the Downs and Black's checklist, was 75.60% (SD=9.43). Items measuring external and internal validity (e.g. representativeness of the sample, reliability and validity of the measures, time of recruitment, or randomization) and power (item 19) showed the lowest quality scores. Table 3 (in Supplementary Material) presents the mean and standard deviation for each item of the adapted version of the scale.

Synthesis of results

Overall effect sizes

Effect sizes for catastrophizing bias and interpretation bias (see Figure 3 and Figure 4) were large and moderate (g=0.95, p< 0.001 and g=0.78, p<0.001, respectively). Heterogeneity was significant and high in all cases, with the I^2 value around 90% (see Table 2 and 3). Significant effect sizes were maintained after the removal of 6 outliers for the catastrophising bias analysis, and 13 for interpretation bias. These sensitivity analyses also showed a reduction in heterogeneity . Details for sensitivity analyses can be found in the Supplementary Material (Tables 4 and 5).

Moderator analyses

a) Catastrophizing bias. For the analyses on catastrophizing and depression (see Table 2), it was found that the variables year of study, sample characteristics (size, type or mean age) and measure of bias (type, number of items or reliability) did not significantly moderate the ES.

Percentage of women, nationality of the corresponding author, and instrument to measure depressive symptoms significantly moderated the ES. The relationship between catastrophizing bias and depression was higher in those studies with a higher percentage of women (meta-regression *estimate*= 0.04, *SE*=0.01, *p*=0.001). Also, the ES was significant for Western but not for non-Western nationality of authors of correspondence and also in studies using DSM criteria, the Symptom Checklist-90, the Beck Depression Inventory or the Depression, Anxiety and Stress Scale. The ES did not reach significance for those studies using the Self-rating Depression Scale or the Centre for Epidemiological Studies–Depression scale. The 'Other' category for the moderator type of depressive symptoms measure also showed a large and significant effect. However, due to the lack of studies, it was not possible to make independent groups to clarify the direction of this significant effect. The instrument to measure depressive symptoms was the only significant moderator in the sensitivity analyses (see Supplementary Material, Table 4).

Table 2. *Analyses on catastrophizing and depression.*

Catastrophizing	N	k	g **[95%CI]**	p	Q (df)p	I^2
Overall ES	4,865	23	0.95 [0.64; 1.26]	< 0.001	311.67 (22) < 0.001	92.94%
Moderators			Beta coefficient/ Mean ES [95%CI]			

Year	4,865	23	0.00 [-0.05; 0.05]	.96	0.0031 (1) .96
Nationality	4,865	23			4.08 (1) .04
Western		17	1.15 [0.78; 1.52]	< 0.001	16.03 (16) .45
Non-western		6	0.41 [-0.21; 1.02]	.19	13.26 (5) .02
N	4,865	23	0.00 [-0.001; 0.003]	.18	1.84 (1) .18
% women	3,204	17	0.04 [0.01; 0.06]	< 0.01	10.56 (1) < 0.01
Mean age	3,204	17	0.00 [-0.03; 0.03]	.95	0.00 (1) .95
Sample type	4,603	22			1.59 (2) .45
GP		4	1.02 [0.25; 1.80]	.01	1.35 (3) .72
Students		12	0.75 [0.30; 1.20]	<0.001	17.21 (12) .14
GP and CP		6	1.25 [0.60; 1.89]	<0.001	13.21 (5) .02
Measure of depression	4,865	23			12.93 (6) .04
DSM		6	1.24 [0.66; 1.82]	< 0.001	15.56 (5) .01
SCL-90		2	1.37 [0.37; 2.37]	.01	0.04 (1) .84
SDS		2	0.79 [-0.17; 1.74]	.11	0.41 (1) .52
BDI		5	1.23 [0.62; 1.84]	< 0.001	0.11 (4) .10
CES-D		3	-0.40 [-1.23; 0.42]	.34	6.95 (2) .03
DASS21		2	1.03 [0.04; 2.01]	.04	0 (1) .99
Other[1]		3	0.98 [0.17; 1.79]	.02	3.15 (2) .21
Catastrophising bias measure	4,865	23			3.03 (1) .08
CERQ-C		15	1.14 [0.78; 1.51]	<0.001	4.92 (14) .99

Other[2]		8	0.57 [0.05; 1.09]	.03	28.58 (7) <0.01
Number of items	4,865	23	0.01 [-0.04; 0.07]	.63	0.23 (1) .63
Reliability	4,078	18	3.64 [-0.61; 7.89]	.09	2.82 (1) .09

N=number of participants. k=number of studies. *g [*95% CI*]* = standardized mean difference with Hedge's correction with confidence interval. Beta coefficient/Mean ES [95% CI] = estimate values for meta-regression analyses with continuous variables, and mean effect sizes for mixed models with categorical variables in moderation analyses. Q (df) p= test of homogeneity statistic, degrees of freedom and p-value. I^2 = percentage in which the observed variability exceeds the expected by chance. Type of sample: CP = clinical population; GP = general population. Measure of depression: BDI = Beck Depression Inventory*;* CES-D = Centre for Epidemiological Studies–Depression; DASS21 = Depression Anxiety Stress Scale; DSM = Diagnostic and Statistical Manual of Mental Disorders criteria to select, at least, the experimental group in categorical studies*;* SCL-90 = Symptom checklist-90*;* SDS = Self-rating Depression Scale. Catastrophizing bias measure: CERQ-C= Cognitive Emotion Regulation Questionnaire-catastrophizing.

[1]Other measures of depression: Brief Symptom Inventory; Patient Health Questionnaire-9; Profile of Mood Scale.

[2]Other catastrophizing bias measures: Cognitive Error Questionnaire-catastrophizing subscale; Future Events Questionnaire (FEQ)-estimate the likelihood-negative subscale; Grief cognitions questionnaire-catastrophic misinterpretations subscale; Likelihood estimation measure- negative events-self subscale; Probability-Cost-Questionnaire (PCQ)-depression-probability subscale; Refraining from catastrophic thinking; The availability test-negative future-self subscale.

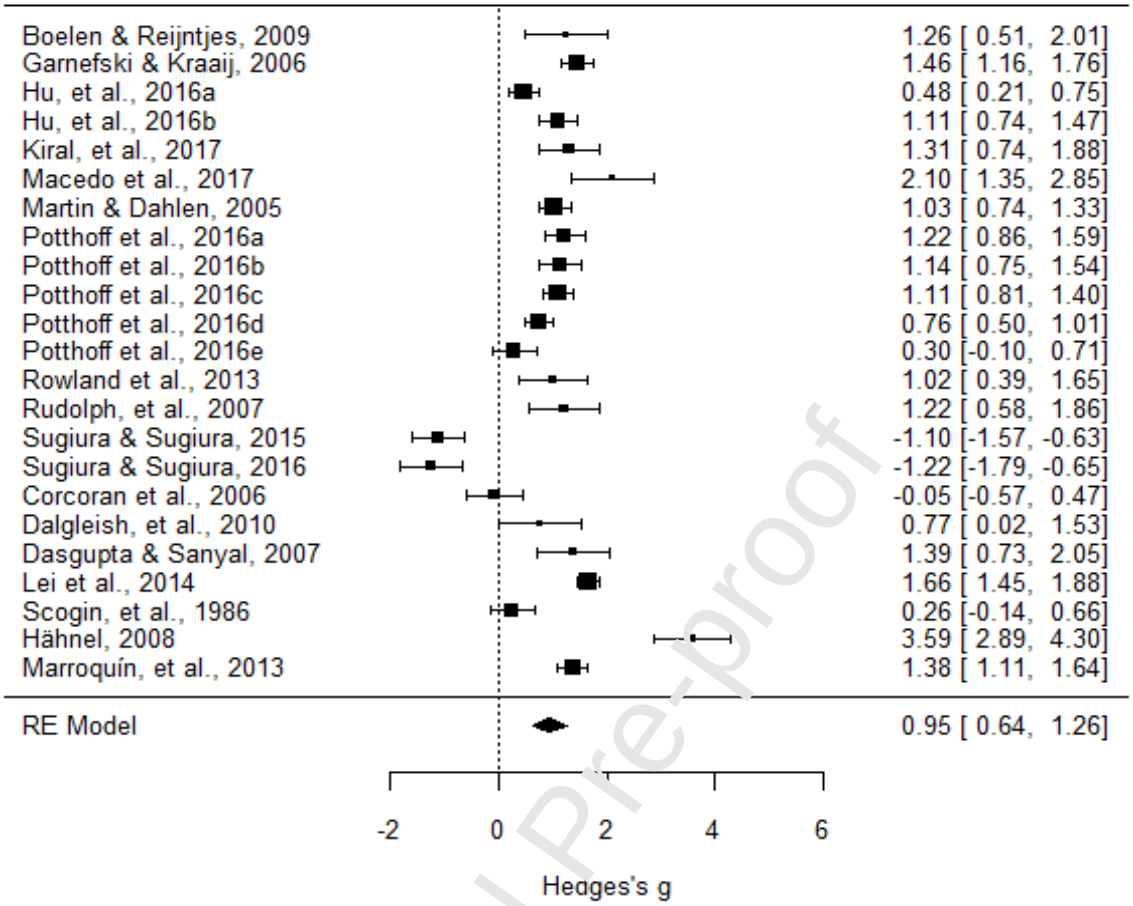


Figure 3. Forest plot for catastrophizing studies.

b) Interpretation bias. Meta-regression and mixed-effects models showed that the variables year of study and the nationality of corresponding author did not significantly moderate the relationship between interpretation bias and depressive symptoms (see Table 3).

Regarding the characteristics of the sample (size, percentage of women and mean age) they did not have a moderating effect, but the type of sample did. Categorical studies comparing depressed and healthy groups were the only category of type of sample with a significant mean ES. The effect in student samples (categorical studies comparing groups above and below a cut-off score, or dimensional studies with only healthy or (sub)clinical

groups) or dimensional studies with only a clinical population did not show a significant relationship.

Finally, the type of instruments used to measure interpretation bias and depression seemed to be relevant. ES was significant only in categorical studies using both a diagnostic instrument (based on DSM or RDC criteria) and a cut-off point in a questionnaire (BDI, Hamilton Depression Rating Scale (HDRS), LSAS), or using just a diagnostic instrument based on DSM criteria but not in those using only RDC criteria or questionnaires (i.e., Beck Depression Inventory, Centre for Epidemiological Studies Depression). Regarding the measure of interpretation bias, ES was significant and large when using the Cognitive Distortion/Bias Questionnaire, the Sentence Completion Test and the Ambiguous Scenarios Test for Depression but not when using the Cognitive Error Questionnaire, or the Interpretation Bias Questionnaire. The group of other measures of distortion also had a significant moderating effect, but the small number of studies using each of these instruments did not allow for the conducting of further analyses within the category. Number of items or reliability of the bias measure did not have a significant effect on ES.

None of these moderation effects were significant in the sensitivity analyses after the removal of outliers (see Supplementary Material, Table 5).

Table 3. *Analysis on interpretation bias and depression.*

Interpretation bias	N	k	g **[95%CI]**	P	Q (df)p	I^2
Overall ES	4,678	40	0,78 [0.43; 1.13]	<0.0001	742.45 (39) <0.0001	94.75%
Moderators			Beta coefficient/ Mean ES[95%CI]			

Year	4,678	40	-0.02 [-0.05; 0.01]	.11	2.53 (1) .11
Nationality	4,678	40			0.25 (1) .62
Western		38	0.76 [0.40; 1.12]	<0.001	90.18 (37) <0.0001
Non- Western		2	1.17 [-0.35; 2.69]	.13	0.53 (1) .47
N	4,678	40	0.00 [-0.00; 0.00]	.32	0.98 (1) .32
% women	3,983	34	0.00 [-0.02; 0.02]	.10	0.00 (1) .10
Mean age	2,847	26	0.02 [-0.01; 0.06]	.19	.69 (1) .19
Sample type	4,678	40			9.49 (3) .02
Other		2	-0.07 [-1.57; 1.43]	.93	7.19 (1) .01
Students		17	0.34 [-0.12; 0.92]	.13	12.95 (16) .68
GP and CP		19	1.34 [0.83; 1.85]	<0.001	60.74 (18) <0.001
CP		2	-0.33 [-2.05;1.39]	.71	1.89 (1) .17
Measure of depression	4575	39			15.13 (4) <0.01
DSM or RDC criteria + cut-off point		9	1.30 [0.58; 2.03]	<0.001	11.78 (8) .16
DSM interview		4	2.49 [1.33; 3.64]	<0.001	41.36 (3) <0.001
RDC		5	0.86 [-0.14; 1.96]	.09	1.58 (4) .81
BDI		17	0.36 [-0.17; 0.89]	.18	17.51 (16) .35
CES-D		4	-0.04 [-1.08; 1.01]	.94	5.31 (3) .15
Interpretation bias measure		40			30.59 (5) <0.001
CEQ		6	0.49 [-0.25; 1.22]	.19	8.34 (5) .14
CDQ/CBQ		12	0.90 [0.37; 1.44]	<0.001	3.23 (11) .99

SCT		3	1.53 [0.46; 2.59]	.001	7.24 (2) .03
AST-D		3	-1.25 [-2.28; -0.21]	.02	0.18 (2) .91
IBQ		2	-1.07 [-2.31; 0.17]	.09	2.71 (1) .10
Other		14	1.33 [0.83; 1.84]	<0.001	64.57 (13) <0.001
Number of items	3143	33	-0.01 [-0.05; 0.03]	.72	0.13 (1) .72
Reliability	1580	11	-3.64 [-11.53; 4.26]	.37	0.82 (1) .37

N=number of participants. k=number of studies. *g [*95%CI*]* = standardized mean difference with Hedge's correction with confidence interval. Beta coefficient/Mean ES [95% CI] = estimate values for meta-regression analyses with continuous variables, and mean effect sizes for mixed models with categorical variables in moderation analyses. Q (df) p= test of homogeneity statistic, degrees of freedom and p-value. I^2 = percentage in which the observed variability exceeds the expected by chance. Type of sample: CP = clinical population; GP = general population; Other =one dimensional study with a sample of students and university staff and one categorical study with groups of the general population divided based on a BDI cut-off score. Measure of depression: BDI= Beck Depression Inventory, CES-D=Centre for Epidemiological Studies–Depression; DSM = Diagnostic and Statistical Manual of mental disorders; RDC = Research Diagnostic Criteria; cut-off point = using questionnaires Beck Depression Inventory (BDI), Hamilton Depression Rating Scale (HDRS) or Liebowitz Social Anxiety Scale (LSAS). Measure of interpretation bias: AST-D = Ambiguous Scenarios Test for Depression-pleasantness rating; CEQ = Cognitive Error Questionnaire*;* CDQ/CBQ (Krantz & Hammen, 1979) = Cognitive Distortion/Bias Questionnaire-depression distortion subscale*;* IBQ = Interpretation Bias Questionnaire; SCT = Sentence Completion Test for Depression-negative statements subscale.

[1]Other measures of interpretation bias: Ambiguous Social Situations Interpretation Questionnaire (ASSIQ)*;* Ambiguous/Unambiguous Situations Diary (AUSD)*;* Sentence Completion Test for Depression-total score; Cognitive Bias Questionnaire-total score; Cognitive Distortion Questionnaire (Burns, Shaw, & Croker, 1987); Cognitions Questionnaire; Interpretations of Depression Questionnaire (IDQ); Interpretation of Events Measure (IEM); Interpretation Inventory (II); Interpretation and Judgmental Questionnaire (IJQ)-multiple choice-cross situations subscale*;* Interpretation Questionnaire for Social Phobia and Depression (IQSD)- depression distortion subscale; Negative and Positive Cognitive Error Questionnaire-negative (NPCEQ); and two measures without name (Drennen, 1991; Nunn, Mathews, & Trower, 1997).

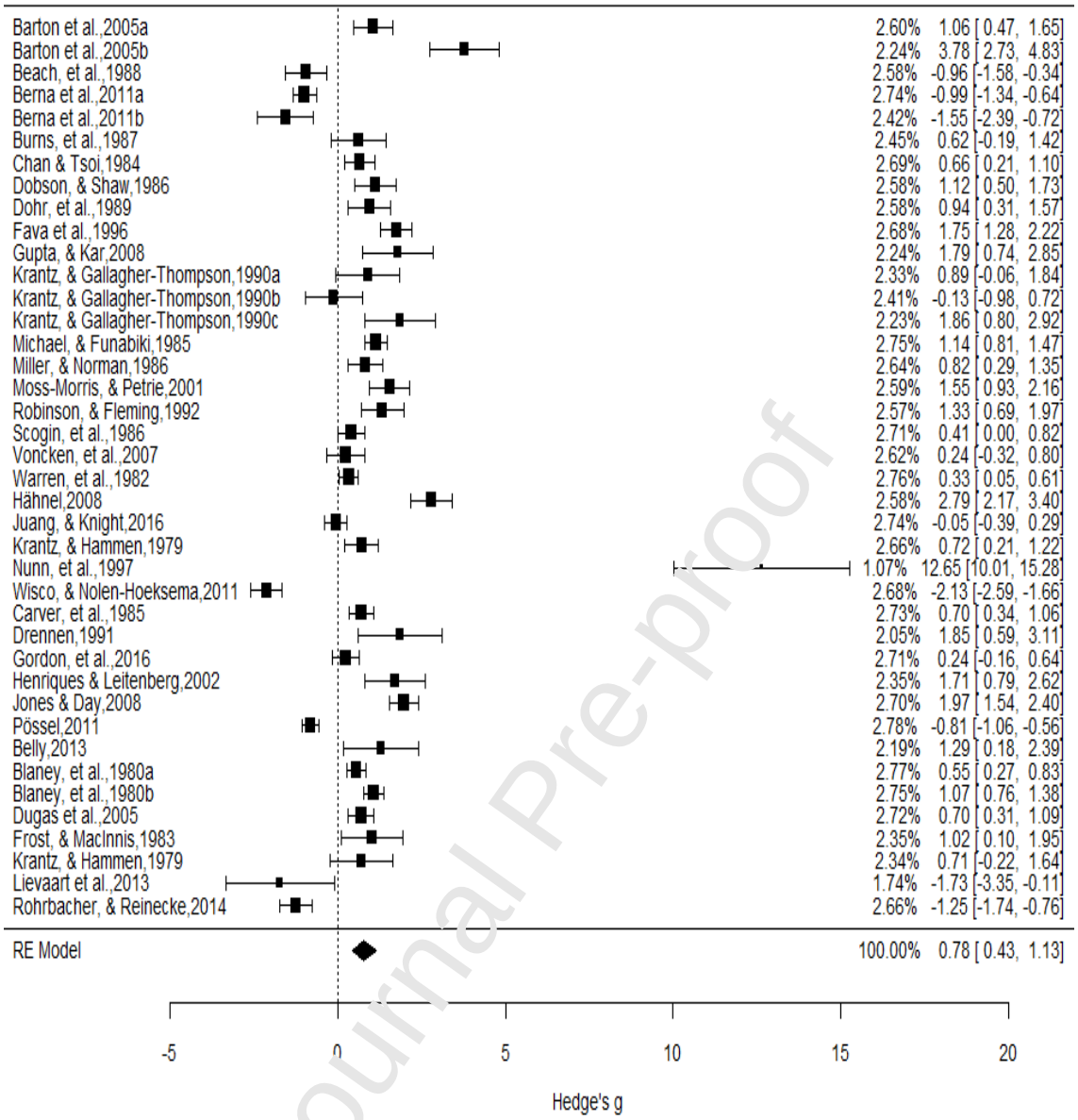


Figure 4. Forest plot for interpretation bias studies.

Risk of bias.

a) Catastrophizing studies. First, the Orwin's fail-safe number was 421, showing no threat of publication bias (Ns = 421 >5k+10). Second, an inspection of the funnel plot (see Figure 5) showed that the shape and density of the funnel plot seemed to be symmetric, which was confirmed with the rank correlation test (Kendall's tau = 0.00, p = 1) and the Eager's

regression test (z = 0.81, p = 0.42). Given that asymmetry was not found, the trim-and-fill procedure was not conducted.

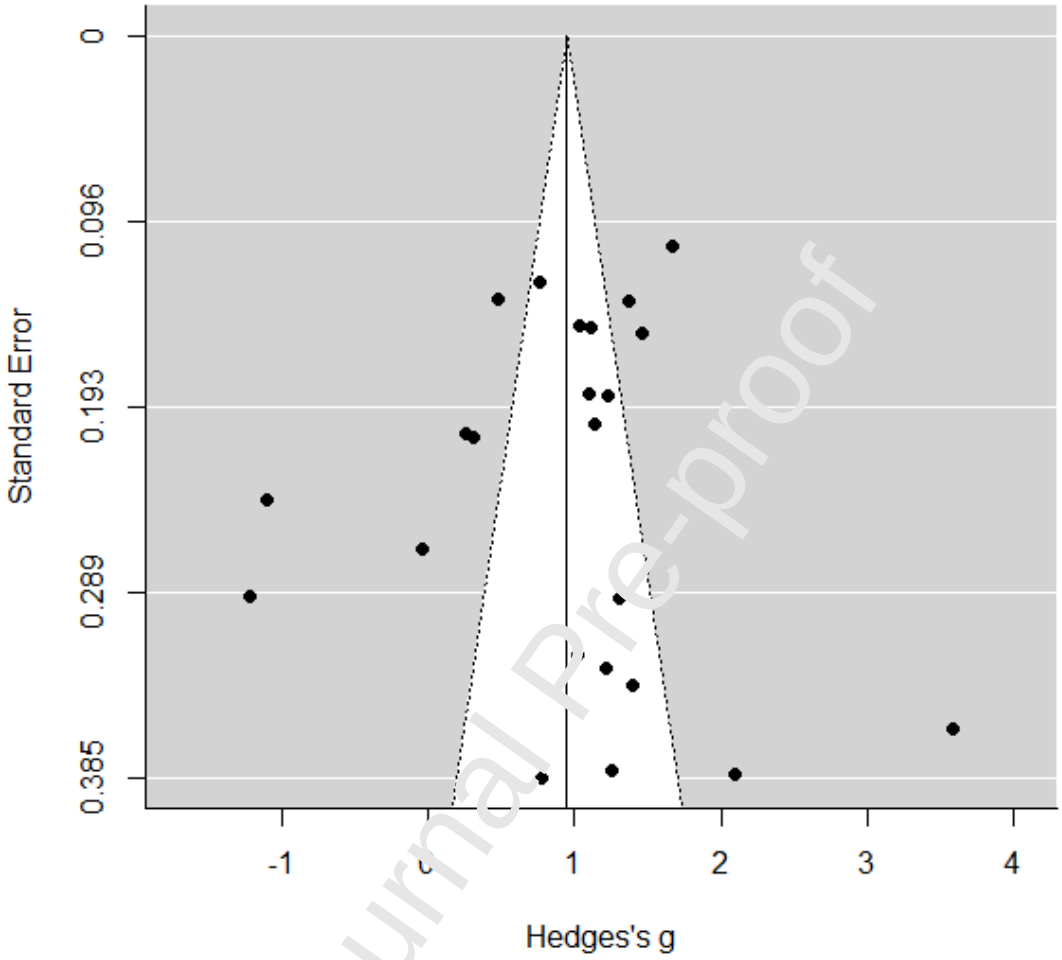


Figure 5. Funnel plot for catastrophizing studies.

b) Interpretation bias studies. First, the Orwin's fail-safe number showed there was not a threat of publication bias (Ns = 780 > 5k + 10). Second, the funnel plot was inspected with and without the outlier (see Figure 6).

The rank correlation test indicated a lack of publication bias (Kendall's tau = 0.19, p = 0.08), while the regression test for funnel plot asymmetry was significant (z = 5.50, p <

.0001). However, sensitivity analyses showed that both the rank correlation test (Kendall's tau = 0.15, p = 0.18) and the regression test for asymmetry (z = 1.80, p = 0.07) indicated the absence of publication bias when removing the outlier.

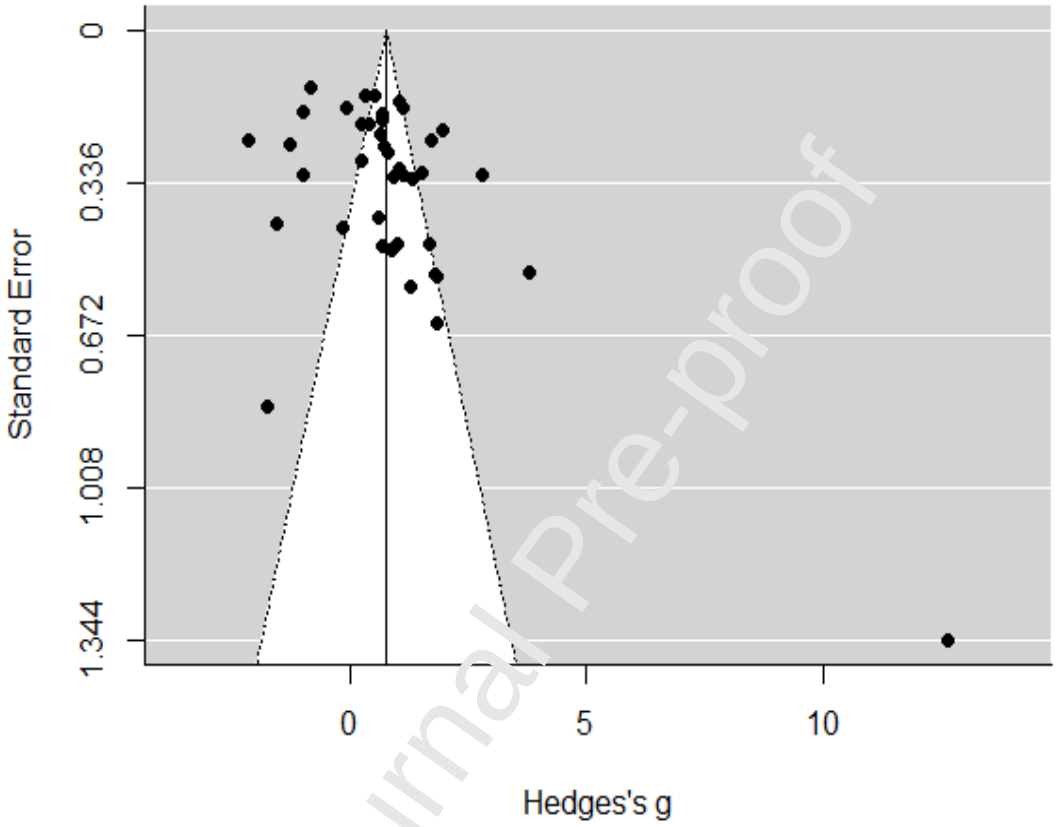


Figure 6. Funnel plot for interpretation bias studies.

Discussion

The aim of this study was to quantify the evidence of self-reported cognitive biases in depressed individuals as compared to other groups of participants. Based on the relatively scarce amount of empirical studies found in our search (k= 63), analyses only included two categories of bias (catastrophizing and interpretation bias) and comparisons between groups of depressive participants and non-symptomatic groups.

Given that the search covered a large period of time (no limit-2020), it seems surprising that only 63 studies could be included in the final review. For example, Hu et al. (2015), in their meta-analysis on attribution bias and depression, reviewed 86 studies in the past 20 years. One of the main reasons for this relatively small number of studies is that the majority of articles initially screened for inclusion focused on cognitive constructs such as maladaptive cognitive schemas and negative automatic thoughts, which do not reflect the concept of information processing biases proposed in Beck's cognitive conceptualization (Beck, 1967, 1979). Yet, a clear distinction among different components of cognitive psychopathology seems to be crucial to further understand the nature of depression (Ingram & Kendall, 1986) and therefore, efforts should be aimed at clarifying the role of different mechanisms and pathways in psychopathology. In this respect, our study focuses on the notion of cognitive biases as defined in Beck's model. Also, the number of studies was limited due to the exclusion of studies using experimental tasks. Despite some recent criticisms on the reliability of tasks commonly used in experimental psychopathology (Hedge, Powell, & Sumner, 2018; Price et al., 2015), there is a vast amount of research consistently showing cognitive biases in a diversity of domains of processing like attention (Peckham, McHugh & Otto, 2013; Winer & Salem, 2016), memory (Matt, Vázquez & Campbell, 1992), and interpretation (Everaert et al., 2017) using experimental tasks. In fact, Everaert et al. (2017) conducted a meta-analysis on interpretation bias in depression using experimental tasks. They included a total of 84 studies and found an overall ES of g=0.72 (95%-CI: [0.62;0.82]), which is very similar to the result of this meta-analysis. Yet, the aim of our study was precisely to focus on self-report methods as these are still the basic instruments that are used in clinical research and applied settings. There has also been no systematic review on these widely used clinical tools in depression.

The results supported the idea that cognitive biases are stronger in depression when compared to healthy groups. Specifically, we found a large effect for catastrophizing (g = 0.95 [0.64; 1.26] ($p<0.001$) and medium effect (g = 0,78 [0.43; 1.13] $p<0.001$) for interpretation bias. The large effect associated with catastrophic thoughts may reveal that depression is strongly related to negative expectations of the future, although this type of negative anticipatory thought seems to also be present in other conditions like chronic pain (Crombez et al., 2013), health anxiety (Todd et al., 2018) or pathological worrying (Dash, Meeten & Davey, 2013). In the case of depression, this catastrophic bias might lead to building an overall negative view of oneself, the future and one's own experience, a cognitive product known as the 'cognitive triad' (Beck, 1979). Regarding interpretation bias, results indicate a higher level of negative interpretation bias in depression when compared to healthy individuals. Specifically, depressive groups present a higher tendency to interpret ambiguous scenarios negatively when putting themselves in that situation. This finding is in line with previous findings showing that depression is characterized by an interpretation bias when using self-referential stimuli (Everaert et al., 2017).

Despite the restriction of including only self-report instruments, a high degree of heterogeneity was found. Geographic location, percentage of women and the measure of depressive symptoms explained part of the variability in catastrophizing results, while type of sample, and the instruments to measure both depression and bias moderated the ES in interpretation bias.

Regarding sample characteristics, the percentage of women moderated the relationship between catastrophizing and depression (i.e., the higher the percentage of women, the higher the effect), which could be related to a small but significantly higher presence of cognitive factors, like rumination, in women than in men (Johnson & Whisman, 2013). Although this

result is also in line with previous findings showing that depression is more prevalent in women than men (Kessler & Bromet, 2013), the moderation effect of gender for the analysis of interpretation bias was not significant. This difference could be influenced by the variance in the number of studies included in the analysis for catastrophizing bias (k=17) and for interpretation bias (k=34). Also, this result should be taken with caution given that the moderation effect disappeared in the sensitivity analysis when removing outliers. Thus, the result related to gender could have been biased by extreme values. Moreover, type of sample significantly explained part of the variance when measuring interpretation bias. The effect was only found in categorical studies comparing clinical and non-symptomatic groups. Neither categorical studies with subclinical samples, nor dimensional studies measuring symptoms of depression in healthy individuals, showed an effect. This finding points out to the possibility that interpretation bias is associated with other unique characteristics of clinically depressed patients. For example, emotion regulation strategies such as rumination are associated with depression and have also been proposed to be influenced by cognitive biases (Joormann & Vanderlind, 2014). Yet, these results should be carefully interpreted since the categories of general population and clinical patients included only one and two primary studies, respectively.

The geographic location of the corresponding author and the measure of depressive symptoms also explained part of the variance in the analysis of catastrophizing bias. The effect was significant only when the corresponding author was from a Western country. It should be noted that this result does not necessarily reflect the relationship between depression and catastrophizing in different country populations since we could only code the nationality of the corresponding author, and not of the sample, which was unfrequently specified in the primary studies. Also, this result should be interpreted with caution since

heterogeneity was high within the non-Western group of studies (Q-within=13.26, p=0.02) and the moderating effect disappeared when excluding outliers (see Supplementary Material). In fact, the only variable left to explain heterogeneity in this analysis was the type of measure of depressive symptoms. This moderation effect should not come as a surprise as there is a controversy on which are the best instruments to identify depression (Choi et al., 2014). Finally, the type of measure of interpretation bias also has a moderating effect, which could be explained by the high variability in the instruments used. For example, some of them allow open-ended responses (SCT, IBQ, Burns's CDQ, IJQ), while others provide multiple choice answers (CDQ/CBQ, IDQ, ASSIQ) or Likert scales to rate how often or how likely participants would respond to the situation in a given manner -CEQ, IQSD, II (Warren, Stake & McKee, 1982), Nunn et al., 1997- or how positively/negatively they perceive the situation (ASTD, IEM, AUSD). Nevertheless, these effects also disappeared when excluding outliers (see Supplementary Material). This variability was also reflected in the instruments measuring catastrophizing. Although this variable did not moderate the ES for catastrophizing bias, it is remarkable the difference in the result obtained by studies using the CERQ-C (g=1.14, 95CI [0.78; 1.51]) in comparison to the rest of the instruments (g=0.57, 95CI [0.05; 1.09]). This finding could actually reflect the heterogeneity in the instructions and contents of the instruments used in the rest of the studies (see Supplementary Table 6 for more details). Congruently with our findings, Everaert et al. (2017) found that the significant ES of interpretation bias in depression was moderated by the type of measurement used, i.e. only significant when measured with direct, but not with indirect, measures. These results may question the validity of some instruments to measure cognitive biases. Although we coded the psychometric features of the different measures used to analyse cognitive biases, very few studies provided data regarding the reliability and validity of the instruments. Also, in many

cases, the subscales, and not the results for the whole questionnaire, were considered to calculate effect sizes (see Supplementary Table 6 for details). All these aspects collectively may have influenced the results. Further research on the convergences and discrepancies of the modalities of measurement are needed to clarify the issue of finding gold standards of measurement of cognitive biases both in the clinical and experimental fields.

While the risk of publication bias was found to be low, the overall quality of primary studies was not optimal. More than one third of studies (26/63) met less than half the criteria evaluated. 'Quality' is a multi-faceted construct and it is likely that some of these criteria, like those referred to the quality of report (e.g., *Have actual probability values been reported for the main outcomes except where the probability value is less than 0.001 or 0.05*?), may have had a minor impact on the resulting effect sizes. However, criteria evaluating aspects related to the external and internal validity of the studies may have had a direct impact on the findings. For example, representativeness of the sample is a criterion directly affecting the conclusions that can be drawn from qualitative results. Also, the psychometric characteristics of the measures in each sample is essential to know the reliability of the findings. This is a matter of concern in the assessment of self-reported cognitions in the clinical field (e.g. Samtani & Moulds, 2017), and it could have contributed to the heterogeneity of the results of this meta-analysis. Moreover, power calculations to detect the effects are almost absent in the literature selected for this study, and both too large and too small sample sizes can lead to erroneous statistical results. However, it is important to remark that the year of publication range from 1979 to 2019, demonstrating one third of the studies published in the last century. While it is still a warning signal for future investigations, quality of research has also increased during recent decades with the development of new standards such as APA standards (APA, 2018) or PRISMA (Moher et al., 2015). Also related to the year of

publication, it is surprising that only 12 of the studies have been conducted during the last five years, which could be related to the tendency to use experimental procedures rather than self-report questionnaires as a more appropriate way to analyse cognitive biases. Nevertheless, it must be kept in mind that the use of questionnaires is still the default choice in clinical practice and clinical trials comparing therapies, which makes it highly important to keep scrutinizing the validity of these assessment methods.

Based on the findings of this meta-analysis, some recommendations to move this field forward are suggested. The assumption that some specific cognitive biases are present in depression still needs empirical evidence. While processes such as selective abstraction or overgeneralization were theoretically established based on clinical practice (Beck, 1963), this meta-analysis was not able, several decades later, to address its magnitude given the lack of empirical data. As mentioned before, the analysis was split into two categories. While the first one allowed us to conclude that catastrophizing bias exists and is stronger in depression groups than in healthy individuals, the second group of studies focused on a broader concept of cognitive biases that can be understood under the general umbrella of 'interpretation' biases. The limited number of studies within this category has made it impossible to draw conclusions about the specific categories of bias proposed in Beck's model (see Table 1).

While experimental research in the field of cognitive bias has developed refined methodologies to measure specific cognitive biases related to interpretation (Holmes, Lang & Shah, 2009), attention (Duque & Vazquez, 2015), and memory (Matt et al., 1992), and the interplay between these processes (Everaert, Koster & Deraksh, 2012), main clinical models are still anchored on theoretical concepts and assessment methods (e.g., self-report questionnaires) that require further support. Fortunately, there is evidence that both the development of theoretical models of depression (e.g., Gotlib & Joormann, 2010), and

interventions aimed at correcting biases (e.g., Jones & Sharpe, 2017), benefit from those newer experimental psychopathology approaches. Although clinical cognitive models of depression, and in particular Beck's model, are still invaluable as heuristic tools, some of their ingredients, like 'cognitive biases', have been addressed using self-report questionnaires which is problematic. On the one hand, as the present meta-analysis has revealed, there has been no consensus on the most reliable and valid self-report instruments to assess these biases. In some cases, the instruments have only been used occasionally, or only by their original authors; in other cases, the instruments require a laborious correction process by independent coders (e.g., the SCT) which is not practical for clinical purposes. On the other hand, and more importantly, self-report methodologies are likely inadequate indicators of true underlying cognitive processing biases, a problem that was identified several decades ago (Nisbett & Wilson, 1977). Also, theoretical models of cognition and the accompanying experimental methods have significantly evolved since Beck's model was formulated (Chipman, 2017) and the field is much more mature and sophisticated than when initial cognitive models of depression were formulated. For instance, experimental research on selective attention in depression (which is conceptually related to the 'selective abstraction' bias in Beck's model, see Table 1), has shown that depressed individuals have difficulties in some, but not others, modalities of attention, which were not described at the time when Beck's model was formulated. In particular, depressed individuals have difficulties to disengage from negative stimuli once they have fixated their attention on them, but they have no difficulties in orienting their initial attention towards emotional stimuli (Armstrong & Olatunji, 2012).

Whilst this is the first study to systematically investigate the presence of the self-reported cognitive biases proposed in Beck's theory, several limitations should also be

mentioned. Our study has not addressed some other types of cognitive biases related to depression, such as thought action fusion (Berle & Starcevic, 2005) or jumping to conclusions (Wittorf et al, 2012) that have also been described in the literature although not included in Beck's seminal theory. Yet, taking Beck's proposal as the theoretical framework of reference seemed to be a good approach for different reasons. First, authors using other cognitive theories (e.g., the hopelessness model of depression) acknowledge the conceptual similarities between models (Panzarella et al., 2006). Second, Beck's model of depression is one of the most used in clinical practice (Greenberg & Padesky, 2015) and empirical results supporting the theory are essential for optimal treatment. Third, the lack of empirical studies fulfilling the inclusion criteria has limited the number of cognitive biases included in the review and the variety of group comparisons to analyse. Another limitation of our results is that it is unknown whether catastrophizing and interpretation bias are specific to depressive individuals or also related to other disorders. Given the surprisingly low number of studies found in our search comparing groups with different psychopathologies, the specificity of these biases could not be assessed, which is a major proposition in cognitive models of depression (Beck & Perkins, 2001).

In sum, this systematic review has found a clear support for the existence of some self-reported cognitive biases in depression as described in Beck's theory, which particularly appear under certain circumstances (e.g., selection of samples using DSM criteria and questionnaires like the CERQ-C). Yet, the rather general low quality of the studies and the lack of sound psychometric properties of the instruments utilized allows for an understanding that the conclusions must be taken with caution. Also, given the scarce number of studies comparing samples with different psychological problems, the specificity of the findings is

unknown. Yet, these findings may shed some light on the robustness of some cognitive biases in depression and the best ways to explore them.

References

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

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, *73*(1), 3-25.

Armstrong, T., & Olatunji, B. O. (2012). Eye tracking of attention in the affective disorders: A meta-analytic review and synthesis. *Clinical Psychology Review*, *32*(8), 704-723.

Barlow, D. H., Farchione, T. J., Fairholme, C. P., Ellard, K. K., Boisseau, C. L., Allen, L. B., & Ehrenreich-May, J. (2011). *The unified protocol for transdiagnostic treatment of emotional disorders: Therapist guide*. New York, NY: Oxford University Press.

Beck, A. T. (1963). Thinking and depression: I. Idiosyncratic content and cognitive distortions. *Archives of General Psychiatry*, *9*(4), 324-333.

Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. Pennsylvania: University of Pennsylvania Press.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: Meridiam.

Beck, A. T. (2008). The evolution of the cognitive model of depression and its neurobiological correlates. *American Journal of Psychiatry*, *165*(8), 969-977.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression.* New York: The Guilford Press.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996*). Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.

Beck, R., & Perkins, T. S. (2001). Cognitive content-specificity for anxiety and depression: A meta-analysis. *Cognitive Therapy and Research, 25,* 651-663.

Berle, D., & Starcevic, V. (2005). Thought–action fusion: Review of the literature and future directions. *Clinical Psychology Review*, *25*(3), 263-284.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester: John Wiley & Sons.

Burns, R. B. (1980). *Feeling good: The new mood therapy*. New York: Penguin Books.

Campbell-Sills, L., & Barlow, D. H. (2007). Incorporating emotion regulation into conceptualizations and treatments of anxiety and mood disorders. In J. J. Gross, (Ed.), *Handbook of emotion regulation* (pp. 542-559). New York: Guilford Press.

Chipman, S. E. (Ed.). (2017). *The Oxford handbook of cognitive science.* New York: Oxford University Press.

Choi, S.W., Schalet, B., Cook, K.F., & Cella, D. (2014). Establishing a common metric for depressive symptoms: Linking the BDI-II, CES-D, and PHQ-9 to PROMIS depression. *Psychological Assessment. 26*(2),513–527.

Clark, D. A., & Beck, A. T. (1991). Personality factors in dysphoria: A psychometric refinement of Beck's Sociotropy-Autonomy Scale. *Journal of Psychopathology and Behavioral Assessment, 13*(4), 369-388.

Clark, D. A., Beck, A. T., & Alford, B. A. (1999). *Scientific foundations of cognitive theory and therapy of depression.* New York: John Wiley & Sons.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Lawrence Earlbam Associates.

Cristea, I. A., Huibers, M. J., David, D., Hollon, S. D., Andersson, G., & Cuijpers, P. (2015). The effects of cognitive behavior therapy for adult depression on dysfunctional thinking: A meta-analysis. *Clinical Psychology Review*, *42*, 62-71.

Crombez, G., Van Ryckeghem, D. M., Eccleston, C., & Van Damme, S. (2013). Attentional bias to pain-related information: A meta-analysis. *Pain*, *154*(4), 497-510.

Cuijpers, P., Noma, H., Karyotaki, E., Cipriani, A., & Furukawa, T. A. (2019). Effectiveness and acceptability of cognitive behavior therapy delivery formats in adults with depression: A network meta-analysis. *JAMA Psychiatry*. *76*(7), 700–707.

Dalgleish, T., & Power, M. (Eds.). (2000). *Handbook of cognition and emotion*. New York: John Wiley & Sons.

Dash, S. R., Meeten, F., & Davey, G. C. L. (2013). Systematic information processing style and perseverative worry. *Clinical Psychology Review*, *33*(8), 1041–1056.

Disner, S. G., Beevers, C. G., Haigh, E. A., & Beck, A. T. (2011). Neural mechanisms of the cognitive model of depression. *Nature Reviews Neuroscience*, *12*(8), 467-477.

Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomized and non-randomized studies of health care interventions. *Journal of Epidemiology & Community Health*, *52*(6), 377-384.

Duque, A. & Vazquez, C. (2015). Double attention bias for positive and negative emotional fa ces in clinical depression: Evidence from an eye-tracking study. *Journal of Behavior T herapy and Experimental Psychiatry*, *46,* 107-114.

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.

Everaert, J., Koster, E. H., & Derakshan, N. (2012). The combined cognitive bias hypothesis in depression. *Clinical Psychology Review*, *32*(5), 413-424.

Everaert, J., Podina, I. R., & Koster, E. H. (2017). A comprehensive meta-analysis of interpretation biases in depression. *Clinical Psychology Review*, *58*, 33-48.

Garnefski, N., Kraaij, V., & Spinhoven, P. (2001). Negative life events, cognitive emotion regulation and emotional problems. *Personality and Individual differences*, *30*(8), 1311-1327.

Gotlib, I. H., & Hammen, C. L. (Eds.). (2014). *Handbook of depression (3nd edition)*. New York: Guilford Press.

Gotlib, I. H., & Joormann, J. (2010). Cognition and depression: current status and future directions. *Annual Review of Clinical Psychology*, *6*, 285-312.

Greenberg, D., & Padesky, C. A. (2015). *Mind over mood (2nd ed.): Change how you feel by changing the way you think***.** New York: Guilford Press**.**

Hammen, C. (2005). Stress and Depression. *Annual Review of Clinical Psychology*, *1*(1), 293–319.

Hammen, C. L., & Krantz, S. (1976). Effect of success and failure on depressive cognitions. *J ournal of Abnormal Psychology*, *85*(6), 577-586.

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive ta sks do not produce reliable individual differences. *Behavior Research and Methods, 5 0*(3), 1166–1186.

Hollon, S. D., & Kendall, P. C. (1980). Cognitive self-statements in depression: Development of an automatic thoughts questionnaire. *Cognitive Therapy and Research, 4*(4), 383-3 95.

Holmes, E. A., Lang, T. J., & Shah, D. M. (2009). Developing interpretation bias modificatio n as a" cognitive vaccine" for depressed mood: imagining positive events makes you f eel better than thinking about them verbally. *Journal of abnormal psychology, 118*(1), 76-88.

Hu, T., Zhang, D., & Yang, Z. (2015). The relationship between attributional style for negativ e outcomes and depression: A meta-analysis. *Journal of Social and Clinical Psycholog y*, *34*(4), 304–321.

Ingram, R. E., & Kendall, P. C. (1986). Cognitive clinical psychology: Implications of an information processing perspective. In R. E. Ingram (Ed.), *Personality, psychopathology, and psychotherapy series. Information processing approaches to clinical psychology* (p. 3–21). London: Academic Press.

Johnson, D. P., & Whisman, M. A. (2013). Gender differences in rumination: A meta-analysis. *Personality and Individual Differences*, *55*(4), 367–374.

Jones, E. B., & Sharpe, L. (2017). Cognitive bias modification: A review of meta-analyses. *Journal of Affective Disorders, 223*, 175-183.

Joormann, J., & Gotlib, I. H. (2010). Emotion regulation in depression: Relation to cognitive inhibition. *Cognition and Emotion, 24*(2), 281-298.

Kessler, R. C., & Bromet, E. J. (2013). The epidemiology of depression across cultures. *Annual Review of Public Health, 34*(1), 119–138.

Krantz, S., & Hammen, C. L. (1979). Assessment of cognitive bias in depression. *Journal of Abnormal Psychology, 88*(6), 611-619.

Lefebvre, M. F. (1981). Cognitive distortion and cognitive errors in depressed psychiatric and low back pain patients. *Journal of Consulting and Clinical Psychology, 49*(4), 517-525.

Matt, G. E., Vázquez, C., & Campbell, W. K. (1992). Mood-congruent recall of affectively toned stimuli: A meta-analytic review. *Clinical Psychology Review, 12*(2), 227-255.

McIntyre, R. S., & O'Donovan, C. (2004). The human cost of not achieving full remission in depression. *Canadian Journal of Psychiatry, 49*(1), 10-16.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*(3), 231–259

Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking rumination. *Perspectives on Psychological Science, 3*(5), 400–424

Orwin, R. G. (1983). A fail-safe N for effect size in meta-analysis. *Journal of Educational Statistics, 8*(2), 157-159.

Panzarella, C., Alloy, L. B., & Whitehouse, W. G. (2006). Expanded hopelessness theory of depression: On the mechanisms by which social support protects against depression. *Cognitive Therapy and Research, 30*(3), 307-333.

Peckham, A. D., McHugh, R. K., & Otto, M. W. (2013). A meta-analysis of the magnitude of biased attention in depression. *Depression and Anxiety*, *30*(4), 1135-1142

Price, R.B., Kuckertz, J.M., Siegle, G.J., Ladouceur, C.D., Silk, J.S., Ryan, N.D., Dahl, R.E., & Amir, N. (2015). Empirical recommendations for improving the stability of the dot-probe task in clinical research. *Psychological Assessment, 27(2),* 365-376.

Reed, A. E., Chan, L., & Mikels, J. A. (2014). Meta-analysis of the age-related positivity effect: age differences in preferences for positive over negative information. *Psychology and Aging*, *29*(1), 1-15.

Riley, R. D., Higgins, J. P. T., & Deeks, J. J. (2011). Interpretation of random effects meta-analyses. *British Medical Journal, 342*(7804), d549.

Samtani, S., & Moulds, M. L. (2017). Assessing maladaptive repetitive thought in clinical disorders: A critical review of existing measures. *Clinical Psychology Review*, *53*, 14–28.

Sterne, J. A., Becker, B. J., & Egger, M. (2006). The funnel plot. In *Publication bias in meta-analysis: Prevention, assessment and adjustments*, (pp. 75-98). Chichester: John Wiley & sons.

Todd, J., van Ryckeghem, D. M. L., Sharpe, L., & Crombez, G. (2018). Attentional bias to pain-related information: a meta-analysis of dot-probe studies. *Health Psychology Review*, *12*(4), 419–436.

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

Vîslă, A., Flückiger, C., Grosse Holtforth, M., & David, D. (2016). Irrational beliefs and psychological distress: A meta-analysis. *Psychotherapy and Psychosomatics, 85*(1), 8-15

Whiteford, H. A., Degenhardt, L., Rehm, J., Baxter, A. J., Ferrari, A. J., Erskine, H. E., … Vos, T. (2013). Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *The Lancet, 382*(9904), 1575–1586

Williams, J. M. G., Barnhofer, T., Crane, C., Herman, D., Raes, F., Watkins, E., & Dalgleish, T. (2007). Autobiographical memory specificity and emotional disorder. *Psychological Bulletin*, *133*(1), 122–148.

Winer, E. S., & Salem, T. (2016). Reward devaluation: Dot-probe meta-analytic evidence of avoidance of positive information in depressed persons. *Psychological Bulletin, 142*(1), 18–78.

Wittorf, A., Giel, K. E., Hautzinger, M., Rapp, A., Schönenberg, M., Wolkenstein, L., ... & Klingberg, S. (2012). Specificity of jumping to conclusions and attributional biases: a comparison between patients with schizophrenia, depression, and anorexia nervosa. *Cognitive Neuropsychiatry*, *17*(3), 262-286

World Health Organization. (2017). *Depression and other common mental disorders: global health estimates* (No. WHO/MSD/MER/2017.2). Recovered from https://www.who.int/news-room/fact-sheets/detail/depression

SUPPLEMENTARY MATERIALS

Self-reported Cognitive Biases in Depression: A meta-analysis

Inés Nieto[1], Elena Robles[1], & Carmelo Vazquez[1]*

[1]Department of Clinical Psychology, Complutense University of Madrid

Running head: Cognitive biases meta-analysis

***Corresponding author:**
Carmelo Vázquez
School of Psychology
Complutense University of Madrid
28223 Madrid (Spain)
E-mail: cvazquez@psi.ucm.es

Supplementary Table 1. *Characteristics of included studies.*

	Author	TE (*g*)	Year	Type of study	Geographic location	N	% women	Mean age (years)	Type of sample	Measure of depression	Measure of cognitive bias
Catastrophizing bias											
1	Boelen & Reijntjes	1.26	2009	D	The Netherlands	79	83.5	21.5	S	SCL-90	GCQ-c
2	Corcoran et al.	-0.05	2006	C	UK	57	28.07	38.7	CG	DSM	AT
3	Dalgleish et al.	0.77	2010	C	UK	29	58.62	48.55	CG	DSM	LEM
4	Dasgupta & Sanyal	1.39	2007	C	India	48	47.92	49.2	CG	DSM	CERQ-c
5	Garnefski & Kraaij	1.46		D	The Netherlands	611	60%	41.92	GP	SCL-90	CERQ-c
6	Hähnel	3.59	2008	C	The Netherlands	8[illegible]	100	22.45	CG	DSM	PCQ
7	Hu et al. a	0.48	2016	D	China	255	32.94	21.9	S	SDS	CERQ-c
8	Hu et al. b	1.11	2016	D	China	262	52.67	21.8	S	SDS	CERQ-c
9	Kirall et al.	1.31	2015	D	Turkey	141	77%	59.74	GP	BDI	CERQ-c
10	Lei et al.	1.66	2014	C	China	458	55.02	38.37	CG	DSM	CERQ-c
11	Macedo et al.	2.10	2017	D	Portugal	258	79.8	19.26	S	POMS	CERQ-c
12	Martin & Dahlen	1.03	2005	D	USA	362	79.01	20.46	S	DASS21	CERQ-c
13	Marroquín et al.	1.38	2013	C	USA	262	-	-	S	BDI	FEQ

	Author	TE (*g*)	Year	Type of study	Geographic location	N	% women	Mean age (years)	Type of sample	Measure of depression	Measure of cognitive bias
Catastrophizing bias (cont.)											
14	Potthoff et al. a	1.22	2016	D	UK	301	81.80	21.69	S	BDI	CERQ-c
15	Potthoff et al. b	1.14	2016	D	UK	235	50.20	24.19	S	BDI	CERQ-c
16	Potthoff et al. c	1.11	2016	D	UK	394	82.20	29.9	S	BDI	CERQ-c
17	Potthoff et al. d	0.76	2016	D	UK	367	62.10	22.24	S	BSI	CERQ-c
18	Potthoff et al. e	0.30	2016	D	UK	102	63.70	28.18	GP	PHQ-9	CERQ-c
19	Rowland et al.	1.02	2013	D	Australia	81	54.3	44.65	GP	DASS21	CERQ-c
20	Rudolph et al.	1.22	2007	D	Canada	100	92	20.2	S	CES-D	CERQ-c
21	Scogin et al.	0.26	1986	C	USA	96	58.33	71.6	CG	DSM	CEQ-c
22	Sugiura & Sugiura	-1.10	2015	D	Japan	157	45.86	19.5	S	CES-D	CCS-Ref
23	Sugiura & Sugiura	-1.22	2016	D	Japan	126	55	18.9	S and GP	CES-D	Ref

	Author	TE (*g*)	Year	Type of study	Geographic location	N	% women	Mean age (years)	Type of sample	Measure of depression	Measure of cognitive bias
Interpretation bias											
1	Barton et al. a	1.06	2005	C	UK	50	70	39	CG	RCA and BDI	SCD
2	Barton et al. b	3.78	2005	C	UK	40	55	37.8	CG	RCA and BDI	SCD
3	Beach et al.	-0.96	1988	C	USA	45	100	38	CG	DSM and BDI	CEQ
4	Belli	1.29	2013	D	UK	40	57.5	20.2	S	BDI	ASSIQ
5	Berna et al. a	-0.99	2011	C	UK	144	69.44	22	S	BDI	AST-D
6	Berna et al. b	-1.55	2011	C	UK	33	45.45	23	S	BDI	AST-D
7	Blaney et al. 1	0.55	1980	D	USA	251	49.8	-	S	BDI	CBQ
8	Blaney et al. 2	1.07	1980	D	USA	350.5	53.56	-	S	BDI	CBQ
9	Burns et al.	0.62	1987	C	USA	25	100	40	CG	RDC	CDQ
10	Carver et al.	0.70	1985	D	USA	175	52	-	S	BDI	CBQ
11	Chan & Tsoi	0.66	1984	C	China	82	50	-	S	BDI	CDQ
12	Dobson & Shaw	1.12	1986	C	Canada	52	61.54	41	CG	RDC	II
13	Dohr et al.	0.94	1989	C	USA	44	56.82	42.5	CG	RDC and HDRS	IEM
14	Drennen	1.85	1991	D	USA	67	50.75	-	S	BDI	CD
15	Dugas et al.	0.70	2005	D	Canada	148	77.03	22.5	S	BDI	AUSD
16	Fava et al.	1.75	1996	C	USA	123	52.84	41.8	CG	DSM	CQ

	Author	TE (*g*)	Year	Type of study	Geographic location	N	% women	Mean age (years)	Type of sample	Measure of depression	Measure of cognitive bias
Interpretation bias (cont.)											
17	Frost & Maclnnis	1.02	1983	D	UK	40	100	-	S	BDI	CBQ
18	Gordon et al.	0.24	2016	D	Australia	103	56	19	S	MASQ	SCT
19	Gupta & Kar	1.79	2008	C	India	30	-	42	CG	DSM and HDRS	CBQ
20	Hähnel	2.79	2008	C	The Netherlands	84	100	22.45	CG	DSM and LSAS	IQSD
21	Henriques & Leitenberg	1.71	2002	D	USA	117	76 (64.96%)	18.8	S	BDI	NPCEQ
22	Jones & Day	1.97	2008	D	UK	231	183 (79.22%)	28.52 (10.27)	S and GP	CES-D	IDQ
23	Juang & Knight	-0.05	2016	C	USA	136	69.23	46.00	CG	CES-D	AST
24	Krantz & Gallagher-Thompson, a	0.89	1990	C	USA	19	-	-	CG	RDC	(EV) CBQ
25	Krantz & Gallagher-Thompson, b	-0.13	1990	C	USA	22	-	-	CG	RDC	(EV) CBQ
26	Krantz & Gallagher-Thompson, c	1.86	1990	C	USA	21	-	-	CG	RDC	(EV) CBQ
27	Krantz & Hammen, a	0.71	1979	D	USA	29	51,72	26	CP	DSM	CDQ
28	Krantz & Hammen, b	0.72	1979	C	USA	65	100	-	S	BDI	CDQ
29	Lievaart et al.	-1.73	2013	D	The Netherlands	36	63.89	41.7	CP	DSM	SCT-total
	Author	TE (*g*)	Year	Type of study	Geographic location	N	% women	Mean age (years)	Type of sample	Measure of depression	Measure of cognitive bias

Interpretation bias (cont.)											
30	Michael & Funabiki	1.14	1985	C	USA	555	-	-	S	BDI	CDQ
31	Miller & Norman	0.82	1986	C	USA	60	60	31,3	CG	DSM and BDI	CBQ
35	Moss-Morris & Petrie	1.55	2001	C	New Zealand	58	70.69	41,9	CG	DSM and BDI	CEQ
33	Nunn et al.	12.65	1997	C	UK	48	-	-	CG	DSM	IT
34	Pössel	-0.81	2011	D	USA	397	80,35	25,27	S	CES-D	CEQ
35	Robinson & Fleming	1.33	1992	C	Canada	47	68.09	-	CG	DSM	CEQ
36	Rohrbacher & Reinecke	-1.25	2014	D	UK	176	77.27	24	S	CES-D	AST-D
37	Scogin et al.	0.41	1986	C	USA	96	58,33	71,6	CG	DSM and HDRS	CEQ
38	Voncken et al.	0.24	2007	C	The Netherlands	53	62.26	42,3	CG	DSM	IJQ
39	Warren et al.	0.33	1982	C	USA	471	58,81	21,2	S	BDI	II
40	Wisco, & Nolen-Hoeksema	-2.13	2011	C	USA	110	61,3	21,8	GP	BDI	IBI

Note: Studies are ordered alphabetically based on the name of the first author. Geographic location: of the corresponding author. Type of study: C=categorical; D=dimensional. N = total sample size of the groups included in the analyses. % women = total percentage of women in the groups included in the analyses. Mean age in years of the groups included in the analyses. Type of sample: CG= different groups of clinical and general population; CP=clinical sample; GP= general population S=students. Measure of depression: BDI= Beck Depression Inventory; BSI= Brief Symptom Inventory; CES-D= Centre for Epidemiological Studies–depression; DASS= Depression Anxiety Stress Scale; DSM= Diagnostic and Statistical Manual of Mental Disorders instruments; PHQ-9= Patient Health Questionnaire-9; POMS= Profile Of Mood Scale; RCA= Routine Clinical Assessment; RDC= Research Diagnostic Criteria; SCL-90= Symptom checklist-90, SDS= Self-rating Depression Scale. Measure of cognitive bias: AST-D: Ambiguous Scenarios Test- Depression- pleasantness; AST =Ambiguous situation task adapted from the IBQ; ASSIQ= Ambiguous Social Situations Interpretation Questionnaire; AT= The availability test-Negative future self; AUSD= Ambiguous/Unambiguous Situations Diary; CBQ=Cognitive Bias Questionnaire-depressed-distorted; (EV) CBQ= Elders' Valenced Cognitive Bias Questionnaire-depressed-distorted; CD= Burns' cognitive distortions definitions; CDQ= Cognitive Distortion Questionnaire; CDM= Cognitive Distortion Measure; CEQ= Cognitive Error Questionnaire; CERQ-c = Cognitive

Emotion Regulation Quest (CERQ)-catastrophizing; CQ= Cognitions Questionnaire; FEQ= Future Events Questionnaire-estimate the likelihood-negative; GCQ-c= Grief cognitions questionnaire- Catastrophic misinterpretations; IBI= Interpretation bias imagery; IDQ= Interpretations of Depression Questionnaire-negative self-dispositional; IEM= Interpretation of Events Measure; II= Interpretation Inventory; IJQ= Interpretation and Judgmental Questionnaire-Part II (multiple choice)-across situations; IQSD= Interpretation Questionnaire for Social Phobia and Depression-depression-dysfunctional; IT= Interpretation Task; Likelihood estimation measure (LEM)-negative events-self; NPCEQ= Negative and Positive Cognitive Error Questionnaire-negative; PCQ= Probability-Cost-Questionnaire-depression-probability; Ref= Refraining from catastrophic thinking; SCD= Sentence Completion Test for Depression-negative statements; SCT-total= Sentence Completion Test-positive, negative, and neutral statements. Number items included in the measure of cognitive bias. α= Cronbach's alpha for the measure of cognitive bias.

Supplementary Table 2. *Studies fulfilling the inclusion criteria but not included in the meta-analysis.*

Reason for exclusion	**Citation**
a) Lack of studies to include comparison groups different from healthy participants	
Schizophrenia (current and in remission)	Corcoran et al. (2006)
Anxiety	Dasgupta & Sanyal (2007)
Suicide	Marroquin, Nolen-Hoeksema, & Miranda (2013, st2)
Bipolarity and depression in remission	Wolkenstein, Zwick, Hautzinger, & Joormann, J. (2014)
Anxiety	Hähnel (2008)
Marital discord	Beach, Nelson, & O'Leary, (1988)
General psychiatric population	Dobson & Shaw (1986)
Depression in remission	Dobson, Rush, & Bernstein (1989)
Depression in remission	Gupta, R., & Kar, B. R. (2008)
General psychiatric population	Miller, I. W., & Norman, W. H. (1986). St1
General psychiatric population	Miller, I. W., & Norman, W. H. (1986). St2
Chronic fatigue syndrome (with and without depression)	Moss-Morris, R., & Petrie, K. J. (2001).
Bereavement (with and without depression) and general psychiatric population	Robinson, P. J., & Fleming
Social phobia (with and without depression)	Voncken, M. J., Bögels, S. M., & Peeters, F. (2007)
Anxiety	Hähnel, A. (2008).
b) Lack of studies to conduct analyses on other cognitive biases[1]	
Positive bias (SCT)	Teasdale, J. D., Lloyd, C. A., & Hutton, J. M. (1998)[a]
Positive bias (SCT)	Teasdale, J. D., Taylor, M. J., Cooper, Z., Hayhurst, H., & Paykel, E. S. (1995) [b]
Overgeneralization	Scogin, F., Hamblin, D., & Beutler, L. (1986) [c]
Overgeneralization	Van den Heuvel, T. J., Derksen, J. J., Eling, P. A., & van der Staak, C. P. (2012) [d]

Overgeneralization	Kernis, M. H., Whisenhunt, C. R., Waschull, S. B., Greenier, K. D., Berry, A. J., Herlocker, C. E., & Anderson, C. A. (1998) [e]
Overgeneralization	Mitchell, S., & Campbell, E. A. (1988) [f]

[1]The studies in this category were excluded because they measured a type of cognitive bias non-comparable to the ones included in the main analyses. The instruments were:

[a b] Sentence Completion Test (SCT; Teasdale et al., 1995) provided an index of positive cognitive bias, which was outside the scope of this meta-analysis.

[c] Cognitive Error Questionnaire (CEQ; Lefebvre, 1981). This study provided data on the overgeneralization subscale alone but could not be included. However, the study also reported data on the catastrophizing subscale and the full scale separately, which were used in the main analyses.

[d] Overgeneralization subscale of the Attitudes Toward Self Scale (ATSS; Carver & Ganellen, 1983). Item example: *How I feel about myself overall, is easily influenced by a single mistake.* Overgeneralization test (OGT; (Klar et al., 1997). Item example: *When you turned on the television yesterday, there was a break in transmission, as a result of which you couldn't watch your favorite show. How probable is it that the next time you want to watch that show there will be a break in transmission again?*

[e] Overgeneralization subscale of the Attitudes Toward Self Scale (ATSS; Carver & Ganellen, 1983). Item example: *Noticing one fault of mine makes me think more and more about other faults.*

[f] Cognitions Questionnaire (CQ; Fennell & Campbell, 1984) measures generalization across time and situations. Examples of items are not provided by the authors.

Supplementary Table 3. *Means and standard deviations for the adapted Downs and Black's Checklist for Measuring Quality.*

Item		M	SD
1	Is the hypothesis/aim/objective of the study clearly described?	1	0
2	Are all primary outcomes to be measured clearly described in the Introduction or Methods section?	0.87	0.34
3	Are the characteristics of the participants included in the study clearly described?	0.52	0.50
4	Are the measures of biases clearly described?	1	0
5	Are the distributions of principal confounders in each group of subjects to be compared clearly described?	0.71	0.46
6	Are the main findings of the study clearly described?	0.62	0.49
7	Does the study provide estimates of the random variability in the data for the main outcomes?	0.74	0.44
8	Have actual probability values been reported for the main outcomes except where the probability value is less than 0.001 or 0.05?	0.38	0.49
9	Were withdrawals and dropouts reported in terms of numbers and/or reasons per group?	0.39	0.49
10	Were the subjects asked to participate in the study representative of the entire population from which they were recruited?	0.23	0.42
11	Were those subjects who were prepared to participate representative of the entire population from which they were recruited?	0	0
12	Were the main outcome measures used accurate (valid and reliable)?	0.15	0.36
13	Were the statistical tests used to assess the main outcomes appropriate?	1	0
14	If any of the results of the study were based on 'data dredging', was this made clear?	1	0
15	Were the participants recruited from the same population?	0.61	0.49
16	Were study participants recruited over the same time?	0.13	0.34
17	Was there adequate adjustment for confounding variables in the analyses?	0.65	0.48
18	Were study subjects randomized to groups?	0	0
19	Did the study have sufficient power to detect a clinically important effect?	0.02	0.13

Sensitivity analyses

Catastrophizing bias

Six outliers were found: Macedo et al., 2017; Sugiura & Sugiura, 2015; Sugiura & Sugiura, 2016; Corcoran et al., 2006; Lei et al., 2014; Hahnel, 2008. After their removal, sensitivity analyses still showed a large ES (g=1 [0.81; 1.19]), while heterogeneity was reduced from 92.94% to 75.32% which was higher than expected.

For moderation analyses, nationality of corresponding author, percentage of women and bias measure became non-significant. The instrument to measure depressive symptoms was the only moderator left to explain the variability of the results. All the instruments included in the analysis showed a significant effect, being larger when using the SCL-90, the BDI or the DASS21. Regarding the two instruments that were found to have a non-significant effect in the main analysis, when removing outliers, the SDS showed a medium effect and the CES-D could not be included in the analysis since just one study used this measure.

Supplementary Table 4. *Analyses on catastrophizing and depression without outliers.*

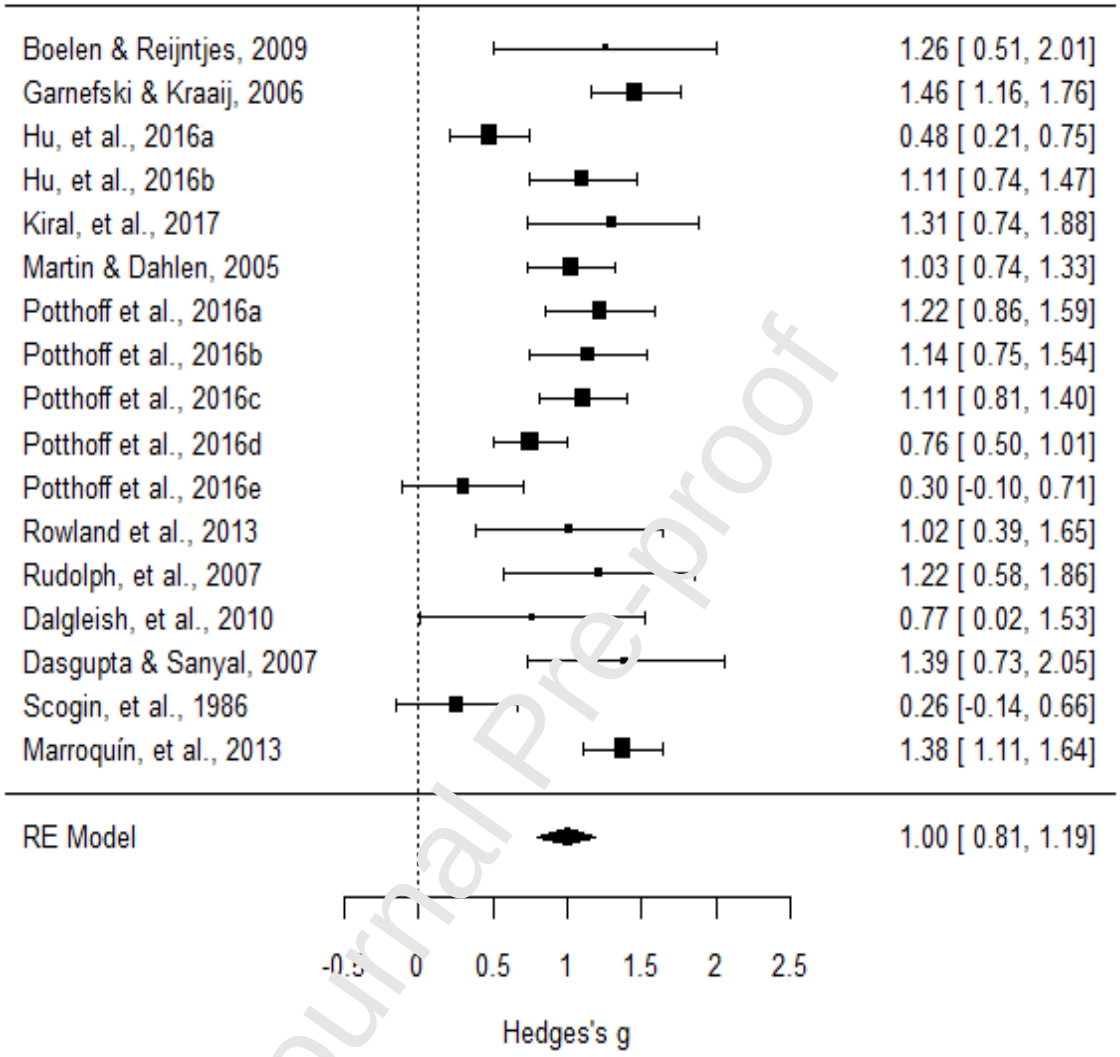
Catastrophizing without outliers	**N**	**k**	**g [95%CI]**	**P**	**Q (df)p**	I^2
Overall ES	3,725	17	1 [0.81; 1.19]	< 0.001	64.84 (16) < 0.001	75.32%
Moderators			Beta coefficient/ Mean ES [95%CI]			
Year	3,725	17	0.01 [-0.01; 0.04]	0.38	0.77 (1) 0.38	
Nationality		17			0.14 (1) 0.71	
Western		14	1.01 [0.80; 1.23]	< 0.001	11.48 (13) 0.57	
Non-western		3	0.92 [0.47; 1.37]	< 0.001	2.70 (2) 0.26	
N		17	0.00 [-0.00; 0.00]	0.17	1.92 (1) 0.17	
% women		11	0.01 [-0.01; 0.03]	0.19	1.72 (1) 0.19	
Mean age		11	0.00 [-0.02; 0.01]	0.57	0.32 (1) 0.57	
Sample type		16			0.99 (2) 0.61	
GP		4	1.02 [0.62; 1.42]	<0.001	5.28 (3) 0.15	
Students		9	1.01 [0.75; 1.26]	<0.001	3.74 (8) 0.88	
GP and CP		3	0.73 [0.22; 1.24]	<0.01	3.44 (2) 0.18	
Measure of depression		16			15.07 (5) 0.01	
DSM		3	0.69 [0.28; 1.10]	0.001	5.16 (2) 0.08	
SCL-90		2	1.41 [0.96; 1.85]	<0.001	0.16 (1) 0.69	
SDS		2	0.76 [0.39; 1.13]	<0.001	2.65 (1) 0.10	
BDI		5	1.23 [0.98; 1.48]	< 0.001	0.67 (4) 0.95	
DASS21		2	1.03 [0.60; 1.46]	< 0.001	0 (1) 0.97	
Other[1]		2	0.57 [0.19; 0.98]	<0.01	1.33 (1) 0.25	

Catastrophising bias measure	17			0.17 (1) 0.68
CERQ-C	13	1.02 [0.80; 1.24]	<0.001	8.53 (12) 0.74
Other[2]	4	0.92 [0.48; 1.35]	<0.001	4.58 (3) 0.21
Number of items	17	0.00 [-0.03; 0.03]	0.93	0.01 (1) 0.93
Reliability	13	1.47 [-1.04; 3.99]	0.25	1.32 (1) 0.25

N=number of participants. k=number of studies. *g* [95%CI] = standardized mean difference with Hedge's correction with confidence interval. Beta coefficient/Mean ES [95%CI] = estimate values for meta-regression analyses with continuous variables, and mean effect sizes for mixed models with categorical variables in moderation analyses. Q (df) p= test of homogeneity statistic, degrees of freedom and p-value. I^2 = percentage in which the observed variability exceeds the expected by chance. Type of sample: CP = clinical population; GP = general population. Measure of depression: BDI = Beck Depression Inventory; DASS21 = Depression Anxiety Stress Scale; DSM = Diagnostic and Statistical Manual of Mental Disorders criteria to select, at least, the experimental group in categorical studies; SCL-90 = Symptom checklist-90; SDS = Self-rating Depression Scale. Measure of catastrophizing bias: CERQ-C= Cognitive Emotion Regulation Questionnaire-catastrophizing.

[1]Other measures of depression: Brief Symptom Inventory, Patient Health Questionnaire-9
Profile Of Mood Scale.

[2]Other catastrophizing bias measures; Cognitive Error Questionnaire-catastrophizing subscale; Future Events Questionnaire (FEQ)-estimate the likelihood-negative subscale; Grief cognitions questionnaire-catastrophic misinterpretations subscale; Likelihood estimation measure- negative events-self subscale; Probability-Cost-Questionnaire (PCQ)-depression-probability subscale; Refraining from catastrophic thinking; The availability test-negative future-self subscale.



Supplementary *Figure 1*. Forest plot for catastrophizing studies without outliers

Interpretation bias

After removing 13 outliers (Barton et al. 2005b; Beach et al., 1988; Berna et al., 2011a; Berna et al., 2011b; Fava et al., 1996; Hahnel, 2008; Juang & Knight, 2016; Nunn et al., 1997; Wisco & Nolen-Hoeksema, 2011; Jones & Day, 2008; Possel, 2011; Lievaart et al., 2013; Rohrbacher & Reinecke, 2014), sensitivity analyses showed that the main ES remained significant, although it changed from medium *g*= 0.78 to large *g*= 0.83. Heterogeneity was reduced from 94.75% to 60.40% of variability which was higher than expected.

The moderating effect of the variable type of sample, type of measure of depressive symptoms and measure of interpretation bias became non-significant. Thus, when excluding outliers, none of the variables explained the variability of the results. Some categories of the moderators (other and CP in sample type, CES-D in measure of depression, and AST-D and IBQ in measure of bias) could not be included in the analysis as just one study used this measure.

Supplementary Table 5. *Analyses on interpretation bias and depression without outliers.*

Interpretation bias without outliers	N	k	g **[95%CI]**	P	Q (df)p	I^2
Overall ES						
All studies	3,075	27	0.83 [0.67; 1.00]	<0.0001	65.66 (39) <0.0001	60.40%
Moderators			Beta coefficient/ Mean ES[95%CI]			
Year	3,075	27	0.00 [-0.01; 0.02]	0.74	0.11 (1) 0.74	
Nationality	3,075	27			0.21 (1) 0.65	
Western		25	0.82 [0.65; 1.00]	<0.001	24.96 (24) 0.41	
Non- Western		2	0.98 [0.32; 1.63]	<0.001	2.36 (1) 0.12	
N	3,075	27	0.00 [-0.01; 0.00]	0.64	0.22 (1) 0.64	
% women	2,428	22	0.00 [-0.01; 0.01]	0.57	0.32 (1) 0.57	
Mean age	1,376	15	0.00 [-0.02; 0.02]	0.97	0.00 (1) 0.97	
Sample type	3,046	26			0.44 (1) 0.51	
Students		13	0.79 [0.56; 1.01]	<0.001	11.46 (12) 0.49	
GP and CP		13	0.90 [0.64; 1.16]	<0.001	15.51 (12) 0.21	
Measure of depression	2972	26			0.70 (3) 0.87	
DSM or RDC criteria + cut-off point		6	0.99 [0.63; 1.35]	<0.001	5.62 (5) 0.34	
DSM interview		2	0.75 [0.13; 1.36]	0.02	2.98 (1) 0.08	
RDC		5	0.84 [0.37; 1.31]	<0.001	6.39 (4) 0.17	
BDI		13	0.83 [0.60; 1.07]	<0.001	9.30 (12) 0.68	

Interpretation bias measure	3075	27			2.87 (3) 0.41
CEQ		4	1.12 [0.68; 1.56]	<0.001	5,67 (3) 0,13
CDQ/CBQ		13	0.85 [0.61; 1.10]	<0.001	10,69 (12) 0,56
SCT		2	0.59 [0.03; 1.15]	0.04	2 (1) 0,16
Other		8	0.73 [0.43; 1.03]	<0.001	6,80 (7) 0,45
Number of items	1874	23	-0.01 [-0.03; 0.01]	0.30	1.09 (1) 0.30
Reliability	1224	8	-1.95 [-6.93; 3.03]	0.44	0.59 (1) 0.44

N=number of participants. k=number of studies. g [95%CI] = standardized mean difference with Hedge's correction with confidence interval. Beta coefficient/Mean ES [95% CI] = estimate values for meta-regression analyses with continuous variables, and mean effect sizes for mixed models with categorical variables in moderation analyses. Q (df) p= test of homogeneity statistic, degrees of freedom and p-value. I2 = percentage in which the observed variability exceeds the expected by chance. Type of sample: CP= clinical population; GP = general population. Measure of depression: BDI= Beck Depression Inventory; DSM = Diagnostic and Statistical Manual of mental disorders; RDC = Research Diagnostic Criteria; cut-off point = using questionnaires Beck Depression Inventory (BDI), Hamilton Depression Rating Scale (HDRS) or Liebowitz Social Anxiety Scale (LSAS). Measure of interpretation bias: CDQ/CBQ (Krantz & Hammen, 1979) = Cognitive Distortion/Bias Questionnaire-depression, distortion subscale; CEQ = Cognitive Error Questionnaire; SCT = Sentence Completion Test for Depression-negative statements subscale.

[1]Other measures of interpretation bias: Ambiguous Social Situations Interpretation Questionnaire (ASSIQ)*;* Ambiguous/Unambiguous Situations Diary (AUSD)*;* Cognitive Bias Questionnaire-total score; Cognitive Distortion Questionnaire (Burns, Shaw, & Croker, 1987); Cognitions Questionnaire (CQ); Interpretations of Depression Questionnaire (IDQ)*;* Interpretation of Events Measure (IEM)*;* Interpretation Inventory (II); Interpretation and Judgmental Questionnaire (IJQ)-multiple choice-cross situations subscale*;* Interpretation Questionnaire for Social Phobia and Depression (IQSD)- depression, distortion subscale*;* Sentence Completion Test for Depression-total score; and two measures without name (Drennen, 1991; Nunn et al., 1997).

Supplementary Table 6. *Description of the cognitive bias measures. (In italics, examples of items when made available by authors).*

Measure	Instruction to participants	Item example
Catastrophising		
CEQ-C	Rating the similarity of the catastrophising thought to the one that they would have had in a similar circumstance on a 5-point scale ranging from "almost exactly like I would think" to "not at all like I would think".	Daily scenarios followed by a catastrophising thought about that situation.
CERQ-C	Rating the use of catastrophising in response to threatening or stressful life events on a 5-point Likert scale ranging from 1 ((almost) never) to 5 ((almost) always).	*I often think that what I have experienced is much worse than what others have experienced.*
FEQ-sub	Estimation of the likelihood that a negative event will happen to them in the future, on a scale ranging from -5 (certain that it will not happen) to +5 (certain that it will happen).	*Being rejected by a significant other.*
GCQ-sub	Rating the tendency to make catastrophic misinterpretations of their own feelings after loss.	Examples are not provided by the author(s).
LEM-sub	Estimating the likelihood that a negative event will happen to them in the future, on a 9-point scale ranging from 'not at all likely' to 'extremely likely'.	*What is the likelihood that, if you surprised a burglar in your home, he would attack you?*
PCQ- sub	Estimation of the likelihood that a negative event will happen to them in the future, on a scale ranging from 0 (not at all) to 10 (extremely likely).	*You can't find any pleasure in your life anymore.*
REF	Rating of perceived ability to do what is described in each item when they feel anxious, on a 4-point scale from 1 (I absolutely cannot) to 4 (I definitely can).	*Even if bad consequences of a problem come to mind, I can reassure myself that they are nothing more than my imagination.*
The availability test- sub	Estimation of the likelihood that a negative event will happen to them in the coming week, on a scale ranging from 1 (not at all) to 7 (very likely).	Examples are not provided by the author(s).
Interpretation bias		
ASSIQ	Choosing what they think is the most likely explanation (negative, neutral, or positive) to a given situation. The total number of negative explanations chosen is summed.	*Someone you know passes you on the street and does not stop to greet you. Why not?* *(a) They were ignoring you* *(b) They didn't see you*

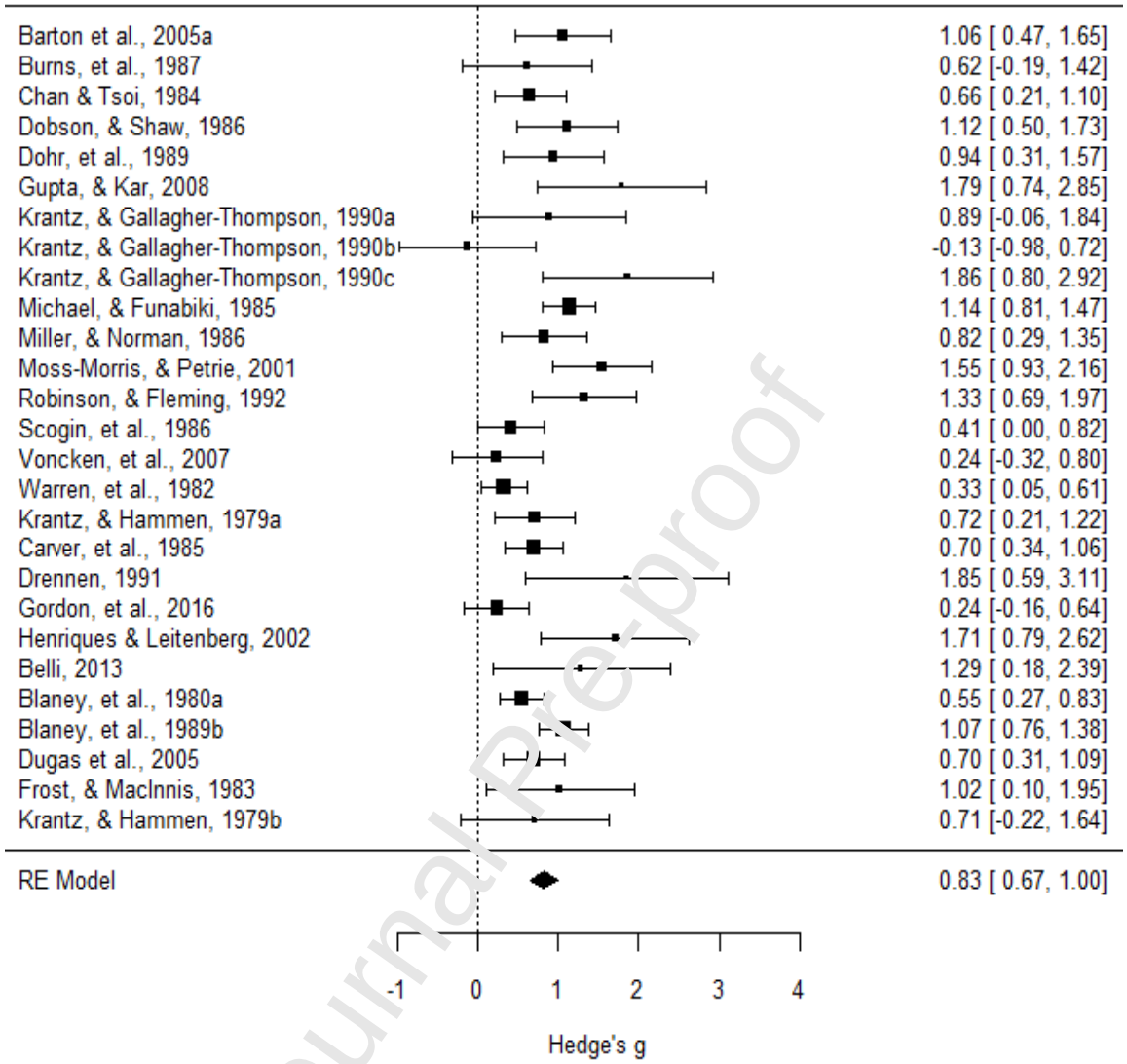
		(c) They were in a rush and could not stop
AST-D	Forming a mental image of a given situation, then they are asked to imagine that it was happening to them and rate how pleasant the image is.	*It's New Year's Eve. You think about the year ahead of you.*
AUSD	Rating the degree of concern for a given situation on a 5-point scale ranging from 'not at all concerned' to 'extremely concerned'.	*While on my way out tonight I was stopped in the street.*
CEQ	Rating the similarity of a type of biased thought (catastrophizing, overgeneralizing, personalization, or selective abstraction) to the one that they would have had in a similar circumstance on a 5-point scale ranging from "almost exactly like I would think" to "not at like I would think".	*You hand in a report to your boss that has taken you four hours to write. Your boss, however, doesn't say anything about it. You think to yourself, '(S)he must think I did a lousy job'.*
CDQ	Writing down open thoughts and feelings in response to a certain situation. Then, score each response for the total number of distortions based on the definitions of Burns (1980): all or nothing thinking, overgeneralization, selective abstraction, disqualifying the positive, mind-reading, fortune-telling, magnification or minimization, emotional reasoning, labelling and mislabeling, and personalization.	*Your spouse (or friend) is quieter than usual tonight.*
CDQ/CBQ	Choosing among four response options (depressed-distorted, depressed-non-distorted, non-depressed-distorted, or non-depressed-non-distorted) the one that would best represent their own response to a certain situation.	A character is depicted as a member of an organization who was encouraged by friends to run for the presidency of the organization. She (he) eventually lost. Participants are instructed to put themselves on the situation and think how they would think and feel. Then they are presented with this scenario: *When you first heard you'd lost, you immediately: 1) feel bad and imagine I've lost by a landslide* (depressive-distorted) *2) shrug it off as unimportant* (non-depressive distorted*), 3) feel sad and wonder what the total counts were* (depressive-non-distorted*), 4) shrug it off, feeling I tried as hard as I could (*non-depressive-non-distorted*).*
CQ	Choosing one of four options (two depressed, two non-depressed) about what they would think and feel given a certain situation.	Examples are not provided by the author(s).
IBQ	Being in a certain situation, participants are asked to write down all explanations/interpretations that came to their mind and choose the	*You call a good friend of yours and leave a message suggesting getting together later in the week. A few days pass, and you*

	most likely for them. Then, they rate the positivity and negativity of each on a Likert scale from 1 (not at all) to 5 (extremely). All responses were also coded by two independent coders as negative or positive.	*haven't heard from them. Why haven't they returned your call?*
IDQ	Different situations are followed by two options (negative self-dispositional and normalizing appraisals). Participants had to indicate (yes/no) whether they had experienced each situation in the preceding three months.	*If I felt cut off from other people, I would probably think it was because:* *-I am an insensitive person.* *-Things are difficult at the moment and I have little energy for other things.*
IEM	Imaging themselves in a certain situation, they are asked to rate their overall reaction on a 4-point Likert scale ranging from 1 = very positive to 4 = very negative.	*Your boss criticizes some aspects of your recent proposal but encourages you to keep working on what he considers to be a generally good idea.*
II	Indicating how often, on a 5-point scale ranging from "never think that way" to "always think that way", they would respond to a certain situation in a biased manner (arbitrary inference, selective abstraction, overgeneralization, magnification, minimization, personalization, and catastrophizing).	*A friend walks by and does not appear to see you. Do you think that your friend is just avoiding you? (arbitrary inference).*
IJQ-sub	Imagining themselves in a certain situation, participants rank four kinds of interpretations (profoundly negative, mildly negative, positive, and neutral) on a scale ranging from 4= most likely to 1=least likely.	*You made an appointment with an acquaintance to go to the cinema. Shortly before the appointment this person leaves a message on your answering machine that the appointment has to be cancelled. Why does this acquaintance cancel the appointment?* *a) This acquaintance doesn't like me (profoundly negative).* *b) This acquaintance has made another appointment and considers the appointment with me not important enough (mildly negative).* *c) This acquaintance likes to go to the cinema with me but couldn't cancel another, tedious, appointment (positive).* *d) This acquaintance feels sick (neutral).*
IQSD-sub	Rating, on a 5-point scale, how likely it is that different thoughts (one functional and two dysfunctional/depressogenic) would come to their	*You lie in bed awake and can't fall asleep. What thoughts occur to you?*

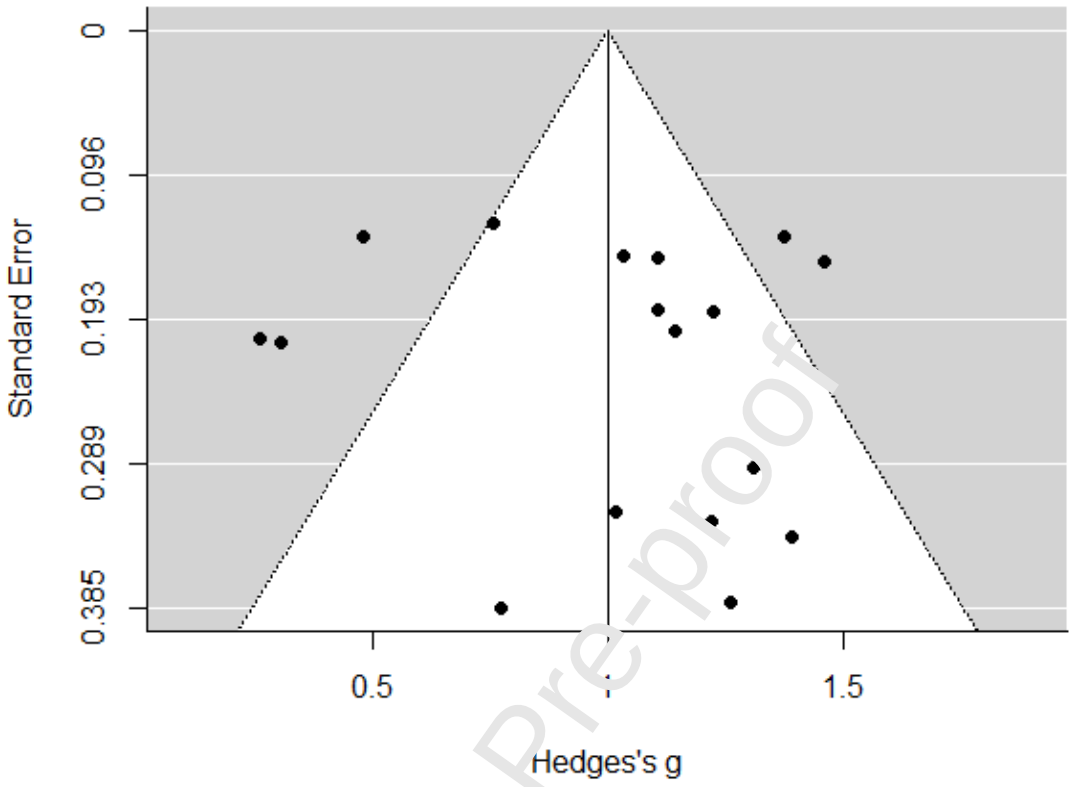
	mind in a certain situation.	*a) This is how it always starts. Now nothing makes a difference anyway.* *b) A lot happened today. My mind has to process all this.* *c) This is terrible. I can't stop ruminating.*
SCT	Completing short sentence stems (agent–verb combinations or simple nouns) based on how they had been feeling within the previous week. Independent raters coded the responses as negative, positive, or neutral.	*I think…* *The world…*
Drennen (1991)	Rating how much each of the 10 different cognitive distortions apply to them on a scale ranging from 0= not like me to 4= extremely like me.	Definitions of 10 different Cognitive Distortions (Burns, 1980, pp. 40-41)
Nunn, Mathews, & Trower (1997)	Imaging themselves in a certain situation, they must rate the similarity of each of the four options (one neutral filler, one general description of an emotional reaction, and two distorted reactions) to how they would likely think and feel.	*A friend phones you about a trip to the cinema, but is unable to suggest a mutually convenient time:* *1. People have better things to do than see you.* *2. It's your fault for not being sufficiently organized.* *3. General emotional reaction.* *4. Neutral filler.*

Measures of catastrophizing: CERQ-C= Cognitive Emotion Regulation Questionnaire-catastrophizing; CEQ-C= Cognitive Error Questionnaire- catastrophizing; FEQ-sub= Future Events Questionnaire- estimate the likelihood-negative subscale; GCQ-sub=Grief cognitions questionnaire-catastrophic misinterpretations subscale; LEM= Likelihood estimation measure- negative events-self subscale; PCQ= Probability-Cost Questionnaire-depression-probability subscale; REF= Refraining from catastrophic thinking scale; The availability test-negative future-self subscale.

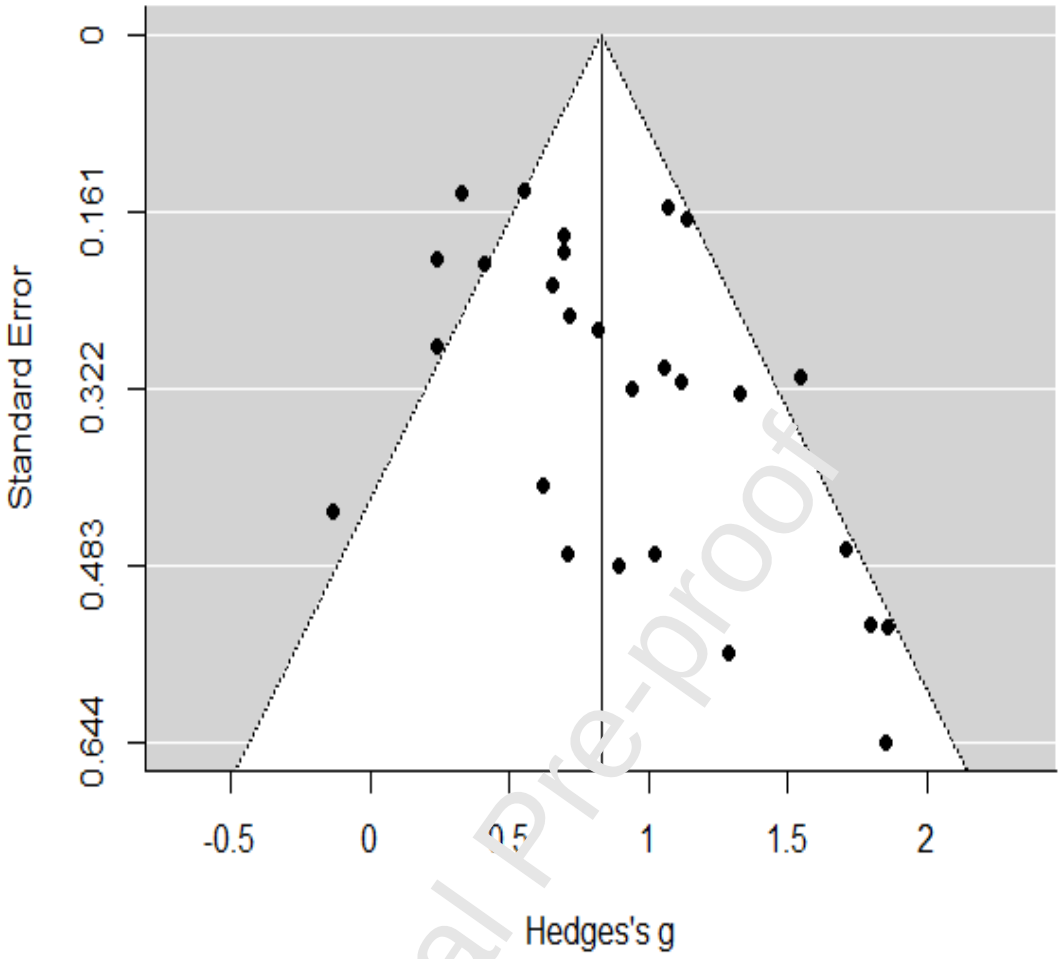
Measures of interpretation bias: ASSIQ= Ambiguous Social Situations Interpretation Questionnaire; AST-D = Ambiguous Scenarios Test for Depression-pleasantness rating; AUSD= Ambiguous/Unambiguous Situations Diary; CEQ = Cognitive Error Questionnaire; CDQ= Cognitive Distortion Questionnaire (Burns, Shaw, & Croker, 1987); CDQ/CBQ = Cognitive Distortion/Bias Questionnaire-depression distortion subscale (Krantz & Hammen, 1979); CQ= Cognitions Questionnaire; IBQ = Interpretation Bias Questionnaire; IDQ= Interpretations of Depression Questionnaire; IEM= Interpretation of Events Measure; II= Interpretation Inventory; IJQ-sub= Interpretation and Judgmental Questionnaire-multiple choice-cross situations subscale; IQSD-sub=Interpretation Questionnaire for Social Phobia and Depression- depression distortion subscale; SCT = Sentence Completion Test for Depression; and two measures without name (Drennen, 1991; Nunn, Mathews, & Trower, 1997).



Supplementary *Figure 2*. Forest plot for interpretation bias studies without outliers



Supplementary *Figure* 3. Funnel plot for catastrophizing studies excluding outliers.



Supplementary *Figure* 4. Funnel plot for interpretation bias studies excluding outliers.

Supplementary references: Primary studies

*Barton, S., Morley, S., Bloxham, G., Kitson, C., & Platts, S. (2005). Sentence completion test for depression (SCD): An idiographic measure of depressive thinking. *British Journal of Clinical Psychology*, *44*(1), 29-46.

*Beach, S. R., Nelson, G. M., & O'Leary, K. D. (1988). Cognitive and marital factors in depression. *Journal of Psychopathology and Behavioral Assessment*, *10*(2), 93-105.

*Belli, S. R. (2013). *Why bother? It's gonna hurt me": the role of interpersonal cognitive biases in the development of anxiety and depression* (Doctoral dissertation). Oxford University, UK

*Berna, C., Lang, T. J., Goodwin, G. M., & Holmes, E. A. (2011). Developing a measure of interpretation bias for depressed mood: An ambiguous scenarios test. *Personality and Individual Differences*, *51*(3), 349-354.

*Blaney, P. H., Behar, V., & Head, R. (1980). Two measures of depressive cognitions: their association with depression and with each other. *Journal of Abnormal Psychology*, *89*(5), 672-682.

*Boelen, P. A., & Reijntjes, A. (2009). Negative cognitions in emotional problems following romantic relationship break-ups. *Stress and Health: Journal of the International Society for the Investigation of Stress*, *25*(1), 11-19.

*Burns, D. D., Shaw, B. F., & Croker, W. (1987). Thinking styles and coping strategies of depressed women: An empirical investigation. *Behaviour Research and Therapy*, *25*(3), 223-225.

* Carver, C. S., Ganellen, R. J., & Behar-Mitrani, V. (1985). Depression and cognitive style: Comparisons between measures. *Journal of Personality and Social Psychology*, *49*(3), 722-728.

*Chan, C. M., & Tsoi, M. M. (1984). The BDI and stimulus determinants of cognitive-related depression among Chinese college students. *Cognitive Therapy and Research*, *8*(5), 501-507.

*Corcoran, R., Cummins, S., Rowse, G., Moore, R., Blackwood, N., Howard, R., ... & Bentall, R. P. (2006). Reasoning under uncertainty: Heuristic judgments in patients with persecutory delusions or depression. *Psychological Medicine*, *36*(8), 1109-1118.

*Dalgleish, T., Golden, A. M., Yiend, J. & Dunn, B. D. (2010). Differential predictions about future negative events in seasonal and non-seasonal depression. *Psychological Medicine*, *40*(3), 459-465.

*Dasgupta, M., & Sanyal, N. (2007). Relationship between controllability awareness and cognitive emotion regulation in selected clinical samples: A psychosocial perspective. *SIS Journal of Projective Psychology & Mental Health*, *14*(1), 64-75.

*Disner, S. G., Beevers, C. G., Haigh, E. A., & Beck, A. T. (2011). Neural mechanisms of the cognitive model of depression. *Nature Reviews Neuroscience*, *12*(8), 467-477.

*Dobson, K. S., & Shaw, B. F. (1986). Cognitive assessment with major depressive disorders. *Cognitive Therapy and Research*, *10*(1), 13-29.

*Dohr, K. B., Rush, A. J., & Bernstein, I. H. (1989). Cognitive biases and depression. *Journal of Abnormal Psychology*, *98*(3), 263-267.

*Drennen, W. T. (1991). Negative schemas and depression in normal college student volunteers. *Psychological Reports*, *68*(2), 521-522.

*Dugas, M. J., Hedayati, M., Karavidas, A., Buhr, K., Francis, K., & Phillips, N. A. (2005). Intolerance of uncertainty and information processing: Evidence of biased recall and interpretations. *Cognitive Therapy and Research*, *29*(1), 57-70.

*Fava, M., Davidson, K., Alpert, J. E., Nierenberg, A. A., Worthington, J., O'Sullivan, R., & Rosenbaum, J. F. (1996). Hostility changes following antidepressant treatment: relationship to stress and negative thinking. *Journal of Psychiatric Research*, *30*(6), 459-467.

*Frost, R. O., & MacInnis, D. J. (1985). The cognitive bias questionnaire: Further evidence. *Journal of Personality Assessment*, *47*(2), 173-177.

*Garnefski, N., & Kraaij, V. (2006). Cognitive emotion regulation questionnaire–development of a short 18-item version (CERQ-short). *Personality and Individual Differences*, *41*(6), 1045-1053.

*Gordon, N. S., Chesney, S. A., & Reiter, K. (2016). Thinking positively: Optimism and emotion regulation predict interpretation of ambiguous information. *Cogent Psychology*, *3*(1), 1195068.

*Gupta, R., & Kar, B. R. (2008). Interpretative bias: Indicators of cognitive vulnerability to depression. *German Journal of Psychiatry*, *11*(3), 98-102.

*Hähnel, A. (2008). *Interpretation processes in anxiety and depression* (Doctoral dissertation). Radboud University Nijmegen, Netherlands.

*Hammen, C. (2005). Stress and Depression. *Annual Review of Clinical Psychology*, *1*(1), 293–319.

*Henriques, G., & Leitenberg, H. (2002). An experimental analysis of the role of cognitive errors in the development of depressed mood following negative social feedback. *Cognitive Therapy and Research*, *26*(2), 245-260.

*Hu, H., Alsron, B., Xu, B., & Hao, W. (2016). Comparative analysis of results from a cognitive emotion regulation questionnaire between international students from West Asia and Xinjiang college students in China. *Shanghai Archives of Psychiatry*, *28*(6), 335-342.

*Jones, S., & Day, C. (2008). Self appraisal and behavioural activation in the prediction of hypomanic personality and depressive symptoms. *Personality and Individual Differences*, *45*(7), 643-648.

*Juang, C., & Knight, B. G. (2016). Age differences in interpreting ambiguous situations: The effects of content themes and depressed mood. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *71*(6), 1024-1033.

*Kiral, K., Yetim, Ü., Özge, A., & Aydin, A. (2015). The relationships between coping strategies, social support and depression: an investigation among Turkish care-givers of patients with dementia. *Ageing & Society*, *37*(1), 167-187.

*Krantz, S. E., & Gallagher-Thompson, D. (1990). Depression and information valence influence depressive cognition. *Cognitive Therapy and Research*, *14*(1), 95-108.

*Krantz, S., & Hammen, C. L. (1979). Assessment of cognitive bias in depression. *Journal of Abnormal Psychology*, *88*(6), 611-619.

*Lei, H., Zhang, X., Cai, L., Wang, Y., Bai, M., & Zhu, X. (2014). Cognitive emotion regulation strategies in outpatients with major depressive disorder. *Psychiatry Research*, *218*(1-2), 87-92.

*Lievaart, M., van der Heiden, C., & Geraerts, E. (2013). Associations between depressive symptoms, rumination, overgeneral autobiographical memory and interpretation bias within a clinically depressed sample. *Journal of Psychology & Psychotherapy*, *7*(2013), 1-6.

*Macedo, A., Marques, C., Quaresma, V., Soares, M. J., Amaral, A. P., Araújo, A. I., & Pereira, A. T. (2017). Are perfectionism cognitions and cognitive emotion regulation strategies mediators between perfectionism and psychological distress?. *Personality and Individual Differences*, *119*, 46-51.

*Marroquín, B., Nolen-Hoeksema, S., & Miranda, R. (2013). Escaping the future: Affective forecasting in escapist fantasy and attempted suicide. *Journal of Social and Clinical Psychology*, *32*(4), 446-463.

*Martin, R. C., & Dahlen, E. R. (2005). Cognitive emotion regulation in the prediction of depression, anxiety, stress, and anger. *Personality and Individual Differences*, *39*(7), 1249-1260.

*Michael, C. C., & Funabiki, D. (1985). Depression, distortion, and life stress: Extended findings. *Cognitive Therapy and Research*, *9*(6), 659-666.

*Miller, I. W., & Norman, W. H. (1986). Persistence of depressive cognitions within a subgroup of depressed inpatients. *Cognitive Therapy and Research*, *10*(2), 211-224.

*Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., ... & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, *4*(1), 1.

*Moss-Morris, R., & Petrie, K. J. (2001). Discriminating between chronic fatigue syndrome and depression: A cognitive analysis. *Psychological Medicine*, *31*(3), 469-479.

*Nunn, J. D., Mathews, A., & Trower, P. (1997). Selective processing of concern-related information in depression. *British Journal of Clinical Psychology*, *36*(4), 489-503.

*Pössel, P. (2011). Can Beck's theory of depression and the response style theory be integrated?. *Journal of Counseling Psychology*, *58*(4), 618-629.

*Potthoff, S., Garnefski, N., Miklósi, M., Ubbiali, A., Domínguez-Sánchez, F. J., Martins, E. C., ... & Kraaij, V. (2016). Cognitive emotion regulation and psychopathology across cultures: A comparison between six European countries. *Personality and Individual Differences*, *98*, 218-224.

*Robinson, P. J., & Fleming, S. (1992). Depressotypic cognitive patterns in major depression and conjugal bereavement. *OMEGA-Journal of Death and Dying*, *25*(4), 291-305.

*Rohrbacher, H., & Reinecke, A. (2014). Measuring change in depression-related interpretation bias: development and validation of a parallel ambiguous scenarios test. *Cognitive Behaviour Therapy*, *43*(3), 239-250.

*Rowland, J. E., Hamilton, M. K., Lino, B. J., Ly, P., Denny, K., Hwang, E. J., ... & Green, M. J. (2013). Cognitive regulation of negative affect in schizophrenia and bipolar disorder. *Psychiatry Research*, *208*(1), 21-28.

*Rudolph, S. G., Flett, G. L., & Hewitt, P. L. (2007). Perfectionism and deficits in cognitive emotion regulation. *Journal of Rational-Emotive & Cognitive-Behavior Therapy*, *25*(4), 343-357.

*Scogin, F., Hamblin, D., & Beutler, L. (1986). Validity of the Cognitive Error Questionnaire with depressed and nondepressed older adults. *Psychological Reports*, *59*(1), 267-272.

*Sugiura, T., & Sugiura, Y. (2015). Common factors of meditation, focusing, and cognitive behavioral therapy: Longitudinal relation of self-report measures to worry, depressive, and obsessive-compulsive symptoms among nonclinical students. *Mindfulness*, *6*(3), 610-623.

*Sugiura, T., & Sugiura, Y. (2016). Relationships between refraining from catastrophic thinking, repetitive negative thinking, and psychological distress. *Psychological Reports*, *119*(2), 374-394.

*Voncken, M. J., Bögels, S. M., & Peeters, F. (2007). Specificity of interpretation and judgemental biases in social phobia versus depression. *Psychology and Psychotherapy: Theory, Research and Practice*, *80*(3), 443-453.

*Warren, N. J., Stake, J. E., & McKee, D. C. (1982). Cognitive distortions, coping behavior, and depression in college students. *Journal of the American College Health Association*, *30*(6), 279-283.

*Wisco, B. E., & Nolen-Hoeksema, S. (2011). Effect of visual perspective on memory and interpretation in dysphoria. *Behaviour Research and Therapy*, *49*(6-7), 406-412.

Abstract

Despite the influence of Beck's cognitive models of depression, the presence and magnitude of the specific proposed cognitive biases have not been systematically investigated. After a systematic search in PsycInfo and PubMED, studies reporting self-reported outcomes on cognitive biases and depressive symptoms in depressed and/or healthy groups were included. From a total of 4,840 records, two different meta-analyses were conducted. 23 studies on 4,865 participants provided data about catastrophising and depression (g = 0.95, 95% CI [0.64; 1.26]) and 40 studies on 4,678 participants provided data about interpretation bias in depression (g = 0.78, 95% CI [0.43; 1.13]). Moderation analyses showed that the relationship between catastrophising and depression was higher in studies with more women, when the corresponding author was from a Western country, and when the instrument to measure depression was the DSM criteria, the SCL-90, the BDI, or the DASS. The relationship between interpretation bias and depressive symptoms was significant only in studies comparing depressed and healthy groups, and when using specific instruments to measure symptoms (DSM/RDC criteria plus a scale cut-off score) and cognitive bias (CDQ/CBQ, SCT, ACT-D, other). Some limitations are acknowledged, but risk of publication bias was found to be low, and these results support the utility of some self-reported measures of cognitive biases in depression.

Keywords: cognitive bias; depression; cognitive model; catastrophizing; interpretation bias

Author Biography

Inés Nieto is a Ph.D. student at Complutense University. She is currently doing her dissertation on cognitive modification bias in depression.

Elena Robles is a Ph.D. student at Complutense University. She is currently doing her dissertation on selective attention biases in chronic pain.

Carmelo Vazquez, Ph.D. is a Professor of Psychopathology at Complutense University. He has done extensive experimental psychopathology research and has published review papers and meta-analyses on cognition in depression in top journals (e.g., Psychological Bulletin; Clinical Psychology Review).

Conflicts of interest

The authors declare no conflict of interest

Highlights

- We examined the magnitude of self-reported cognitive biases in depression.
- Sixty-three studies published between 1979 and 2017 were included (N = 9543 participants).
- Meta-analyses included both studies of individuals grouped by a diagnostic category and studies using a dimensional measure of symptoms of depression.
- The most consistent evidence of biases was found for catastrophizing and interpretation bias.
- There were few evidences of the specificity of cognitive biases in depression.