A systematic review of the literature on interpretation bias and its physiological correlates

Collins, A. [1], Scott, R. B. [1], Hirsch, C. R.[2], Ottaviani, C.[3,4], Krahé, C.[5], Meeten, F. [1]

[1] School of Psychology, University of Sussex, Falmer, UK.

[2] Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK.

[3]Department of Psychology, Sapienza University of Rome, Rome, Italy.

[4]Functional Neuroimaging Laboratory, IRCCS Santa Lucia Foundation, Rome, Italy

[5]Department of Primary Care and Mental Health, Institute of Population Health, University of Liverpool, Liverpool, UK.

Correspondence:

Dr Frances Meeten, School of Psychology, Pevensey 1 Building, University of Sussex, Falmer,

BN1 9QH, United Kingdom

Email: frances.meeten@sussex.ac.uk

Abstract

An important, yet under-explored area of interpretation bias research concerns the examination of potential physiological correlates and sequalae of this bias. Developing a better understanding of the physiological processes that underpin interpretation biases will extend current theoretical frameworks underlying interpretation bias, as well as optimising the efficacy of cognitive bias modification for interpretation (CBM-I) interventions aimed at improving symptoms of emotional disorders. To this end, systematic searches were conducted across the Web of Science, PsycInfo and Pubmed databases to identify physiological markers of interpretation bias. In addition, grey literature database searches were conducted to compliment peer-reviewed research and to counter publication bias. From a combined initial total of 898 records, 15 studies were included in qualitative synthesis (1 of which obtained from the grey literature). Eligible studies were assessed using a quality assessment tool adapted from the Quality Checklist for Healthcare Intervention Studies. The searches revealed seven psychophysiological markers of interpretation bias, namely event-related potentials, heart rate and heart rate variability, respiratory sinus arrythmia, skin conductance response, pupillometry, and electromyography. The respective theoretical and practical implications of the research are discussed, followed by recommendations for future research.

Key words: interpretation bias, physiology, physiological markers, emotional disorders, cognitive bias modification

Introduction

In everyday life, we frequently encounter ambiguous situations that are open to a number of interpretations. For example, an email from your boss asking you to attend a meeting with them could be interpreted as positive (they want to congratulate you on your recent work), or negative (they want to inform you that you are at risk of redundancy). Negative interpretation bias is the tendency to disambiguate information in a consistently negative manner. The role of cognitive biases in the development and maintenance of emotional disorders is well-established (e.g., Butler & Mathews, 1983; Chen et al., 2020; Everaert et al., 2017; Eysenck et al. 1991; Fodor et al., 2020; Hirsch et al. 2016; Mathews et al. 1989; Menne-Lothmann et al., 2014) and a growing literature indicates that the modification of a negative bias (to a more benign or positive bias known as cognitive bias modification for interpretation; CBM-I), results in clinically meaningful reductions in self-reported symptoms of repetitive negative thinking such as anxiety and depression (e.g. Hirsch et al. 2020; 2021).

There is a large body of research from the field of affective neuroscience (see Pace-Schott et al., 2019 for a review) supporting the view that body sensations inform emotional experience. Research to date on interpretation bias provides knowledge of the cognitive processes that maintain emotional disorders and can be targeted for modification in interventions. Yet, little is known about the physiological states associated with negative interpretation bias. An improved understanding of the relationship between cognitions and physiology could permit us extend current knowledge of interpretation bias and to target physiological correlates of emotional disorders in a manner that could complement existing cognitive strategies. For example, understanding whether there is a physiological sequalae

to interpretation bias may optimise and indeed enhance CBM-I efficacy by permitting interpretations to be targeted (via for example, biofeedback) at the very early stages of interpretation generation, which may occur below the level of conscious awareness, but be observable through physiological measurement. Thus, it would be important to gain an understanding of the physiological processes that relate to the presence of a negative bias as measured by cognitive tasks, or a physiological change that occurs when we modify a bias from negative to more positive/benign. Thus, the present article seeks to systematically examine evidence relating to the association between interpretation bias and the underlying physiological processes.

Interpretation biases in emotional disorders

Intrusive and repetitive negative thoughts are a key feature of emotional disorders (Ehring & Watkins, 2008). Negative interpretation bias is known to be a key causal and maintaining factor of repetitive negative thinking (a transdiagnostic feature of anxiety and depression) whereby the tendency to make consistently negative interpretations of ambiguous information contributes to heightened perception of threat, which in turn leads to increases in repetitive negative thought (Hirsch & Mathews, 2012). Evidence suggests that interpretation bias is associated with a range of emotional disorders (Beck & Clark, 1997; Hirsch et al. 2016) such as anxiety and depression (Anderson et al., 2012; Butler & Mathews, 1983; Everaert et al., 2012; Everaert et al., 2017; Mathews & MacLeod, 2005, 2012; Hallion & Ruscio, 2011; Menne-Lothmann et al., 2014) as well as paranoia (Trotta et al., 2021; a common symptom of psychosis). Typically, the interpretation bias literature has examined three key areas. The first examines at what stage in information processing the bias occurs (e.g. is it a fast and relatively automatic process or a slower reflective process?). The second

area investigates whether interpretation bias can be experimentally manipulated (using a pre-post experimental design) and changes in bias subsequently observed on a cognitive task (near transfer). The third area examines the 'far-transfer effects' of bias training, examining, for example, how positive interpretation bias training affects performance on a subsequent behavioural worry task (e.g. Hirsch et al., 2009) or subsequent performance in stressful task (e.g. Joormann et al., 2015). These research areas will be discussed in more detail below.

At what stage in processing does interpretation bias occur?

Research has shown that interpretations can occur at different stages of information processing. 'Online' interpretations are considered to be generated spontaneously and reflexively without the opportunity to reflect and revise, whereas 'offline' interpretations capture conscious, effortful processing, whereby individuals have time to reflect upon the various likely resolutions of ambiguity (Bargh, 1994; Feng et al., 2019; Hirsch et al. 2016; Teachman et al., 2012). Cognitive and behavioural assessments of interpretation bias depend upon these respective time points, where reaction times or response latencies provide useful indices of online bias (e.g. quick, speedy responses) and Likert scaling, rank ordering or endorsement ratings provide useful indices of offline bias (e.g., slow, reflective responses). There exist a large number of paradigms in the interpretation bias literature, each of which varies in the type of stimuli or material employed, such as ambiguous scenarios (offline measures) or single-target prime words and images (online measures; see Schoth & Liossi, 2017 for a review of methods).

Cognitive assessments of interpretation bias tend to capture responses at a fairly late stage in information processing, where one could generate a number of potential

interpretations and pick from one that seems most relevant. While informative as to this type of reflective information processing in emotional disorders, these approaches do not enable an understanding of the nature of interpretations generated at very early stages in information processing, at the point of encountering ambiguity. Capturing interpretation generation at an early stage in information processing is important as it allows the examination of the automaticity of negative interpretation biases. This automaticity is arguably a key contributory factor in the maintenance of emotional disorders, as it can result in interpretations which are outside of conscious awareness, are uncontrollable, and thus are difficult to target in widely used treatment approaches such as cognitive behavioural therapy (Teachman et al., 2012; Hirsch et al. 2016).

Can interpretation bias be experientially manipulated (near-transfer) and does this manipulation affect subsequent cognitions (far-transfer)?

Robust evidence from experimental studies supports the causal role of interpretation bias in emotional disorders. Here, studies that employ CBM-I (a computerised procedure that systematically trains individuals to interpret ambiguous information in a positive/benign manner) show pre-to-post training changes in tasks measuring interpretation bias (near-transfer effects; Hirsch et al., 2018; Menne-Lothmann et al., 2014; Salemink et al., 2022) and on measures of emotional symptoms such as worry, depression, and physiological changes such as heart rate and skin conductance responses during stressor tasks (far-transfer effects; Hallion & Ruscio, 2011; Hirsch et al., 2016; Joormann et al., 2015, Menne-Lothmann et al., 2014; Yiend et al. 2005; for a comprehensive review that examines the evidence for near - and far-transfer effects of CBM-I see Hirsch et al., 2016).

Application of psychophysiological measures to the study of interpretation bias

Research shows that maladaptive cognitions associated with interpretation bias, such as perseverative thinking (e.g., worry or rumination; cf. Everaert et al., 2017; Hirsch & Mathews, 2012; Krahé et al., 2019) evoke sustained and directional alterations in autonomic arousal, one example being low resting state heart rate variability (HRV), a measure of the dynamic interplay between the parasympathetic and the sympathetic nervous system (Diamond, & Fisher, 2017; Brosschot et al., 2010; Hyde, et al. 2019; Ottaviani et al., 2016). The neurovisceral integration model (Thayer et al., 2009; Smith et al., 2017) highlights the integration of body-mind processes, linking better pre-frontal inhibitory processes to higher resting HRV. This relationship was traditionally viewed as a uni-directional 'down-stream' process where the prefrontal cortex integrates information from different systems to regulate HRV. However, more recent work (Mather & Thayer, 2018) discusses evidence in support of 'up-stream' links between increased HRV and increases in functional connectivity in brain regions associated with emotion regulation. Investigators are increasingly turning their attention to examination of brain-mind-body interactions in psychopathology, in order to understand the presentation of psychopathology across multiple axes (e.g. cognitive, autonomic, neurobiological), with the ultimate aim of elucidating targets for therapeutic intervention.

Over the past decades, there has been an increase in the implementation of physiological and neural measures for investigating brain-mind-body interactions, especially those involved in affective and cognitive processes (Tooley et al., 2017; van der Ploeg et al., 2017; Ottaviani et al., 2017). For example, commonly used methods include electroencephalography (EEG) and functional magnetic resonance imaging (fMRI). EEG offers high

temporal resolution when examining neural responses during sensory and cognitive processing (Sur & Sinha, 2009). Whereas fMRI offers greater spatial information, thus allowing researchers to better understand which brain regions are associated with, for example, the cognitions underpinning therapeutic efficacy in psychotherapy (Marwood et al. 2018). A variety of physiological measures typically assess autonomic responses (e.g., heart rate (HR) and its variability (HRV), blood pressure, skin conductance response (SCR), pupillometry, electromyography (EMG) etc. that are important for examining stress-related/arousal responses (Hyde et al., 2019).

Physiological changes during the occurrence of interpretation bias

There is a growing body of research examining the relationship between interpretation bias assessments and physiological responses to resolving ambiguity. Considering how interpretation bias has been examined in the cognitive literature is important, as it dictates the ways in which we can begin to examine physiology in relation to cognitive measures of interpretation bias. One way of examining the relationship between interpretation bias and physiology is to track physiological change as interpretation occurs. A common cognitive task used as an 'online' measure of interpretation bias is the lexical decision task (Feng et al., 2019, Hirsch & Mathews, 1997; 2000). Here, participants read a scenario that remains ambiguous until the final word is presented. They are then presented with a letter string that is either a non-word (not a word in the English dictionary e.g., surbey), or a positive word, or a negative word that disambiguates the sentence. Participants are asked to indicate (by pressing a button) as quickly and as accurately as possible whether the presented word is a word or not. The reaction times for the positive and negative words provide an index of interpretation bias – with faster reactions to negative words indicating

a more negative bias.

Event-related potentials (ERPs), which are evoked deflections in the ongoing EEG time-locked to the presentation of certain stimuli, have been measured and used to infer evidence of interpretations occurring at early stages of processing (prior to that measured by behavioural responses such as reaction times). Schick et al. (2013) propose that these ERPs can be used as an indirect assessment of interpretation bias. As noted by Moser et al. (2012) the cognitive 'online' assessments cannot determine the time course of responses. However, Moser and colleagues argue that measuring ERPs allows for a temporal assessment of interpretation bias. For example, the N400 is an ERP component measured in studies typically investigating semantic encoding (occurring 300-600ms post-stimulus), that is usually evident when an individual's expectation of an upcoming stimulus has been violated (Kornblum & Requin, 2019). As interpretations involve expectations regarding the outcome/resolutions of ambiguous situations, researchers (e.g., Feng et al., 2019; Moser et al., 2012) have proposed that the N400 is useful for examining fast-occurring cognitive processes related to interpretative biases (referred to elsewhere in the literature as online interpretation biases, cf. Hirsch et al., 2016) measured, for example, during lexical decision tasks (as described above; Feng et al. 2019, 2020; Moser et al., 2009, 2012). As the N400 captures expectancy of words, designing tasks which involve presenting terminal words which are either positive or negative allows researchers to assess the match/mismatch between expected endings and the presented words. In this way, N400 can be used to give some insight into how participants likely expected the ambiguity to be resolved in these sentences. Relatedly, the P600 component (which is proposed to reflect semantic and thematic violations) has also been used to infer neurophysiological correlates of

interpretation bias in socially anxious participants (Moser et al., 2008). Other researchers have assessed physiological responses at the time of disambiguation of ambiguous information using methods such as EMG (e.g., Lawson et al., 2002). Arguably, whilst this type of measure can capture the level of arousal when interpretation occurs, it does not provide information about the time-course of interpretation. Overall, however, measures which capture physiological indices at the time that interpretation occurs, provide evidence of an association between cognitive and physiological concomitants of interpretation bias (i.e., EEG, HR, HRV, EMG, SCR, pupillometry).

Physiological markers of near-and far-transfer of interpretation bias training

Physiological measures have also been used to identify whether manipulation of interpretation bias, via CBM-I, is associated with corresponding changes in autonomic responses. One way of doing this is to map physiological changes from before to after interpretation bias training, which can act as an indirect measure of the near transfer of interpretation bias training (cf. Meeten et al., 2017). However, the majority of research in this area has examined physiological stress responses to a task administered *after* positive interpretation bias training (e.g. far-transfer of training effects). For example, a number of studies have assessed CBM-I efficacy via assessment of post-training SCR, an index of stress-mediated arousal and particularly of sympathetic nervous system activity (Nowakowski et al., 2015; Van Bockstaele et al., 2020; Grisham et al., 2014; Joormann et al., 2015; Rozenman et al., 2020; Beadel et al., 2013). Research may begin integrating interpretation bias research within a stress-disease framework (Lang et al., 2017; Verkuil et al. 2010), allowing better understanding of the processes of change underlying CBM-I. However, it should be noted that

the association between bias change and a measure of arousal during (for example) a subsequent stressor task does not constitute a physiological correlate of the bias per se, rather an examination of how change in bias affects physiological responding.

Overall, it is clear that psychophysiological processes have potential application in interpretation bias research. Whilst previous interpretation bias research has largely focused on cognitive and behavioural tasks to assess bias, physiological measures that are measured during resolution of ambiguity offer a number of exciting opportunities to extend knowledge in this field and support our understanding of the processes that contribute to the development and determination of cognitive biases. For example, these measures can help by pinpointing the stage in information processing at which interpretation occurs, by looking at the time-course of processing effects, and by giving a more covert measure of processing. Among other important considerations for including physiological parameters, computational modelling is increasingly used to emulate, and predict, cognitive/psychological processes. Physiological measures are favourable for modelling such processes, which can contribute to the refinement of established cognitive models of interpretation bias (Eguchi et al., 2017; Van den Bergh et al., 2021).

In summary, it is well documented that interpretation bias is an important feature in the development and maintenance of psychopathology and that CBM-I provides an effective tool to improve symptoms of emotional disorders. However, an important gap in the literature concerns the role of (neuro)physiological processes in the generation and maintenance of (negative) interpretation biases. A better understanding of the physiological processes associated with interpretation bias will augment the existing theoretical

framework of interpretation bias, whilst potentially bridging the mind-body interaction in psychopathology. Thus, the current study systematically examined the physiological correlates and sequalae of interpretation bias within the adult population. To the best of the authors' knowledge, this is the first pre-registered systematic review to address the current research question.

Method

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009) were followed when conducting this systematic review.

Search strategy

The review was registered on the Open Science Framework (https://osf.io/v6q2n/). Three databases, Web of Science, PsychInfo and Pubmed were searched from inception to August 20th 2020 and again on February 9th 2022, whereby three primary search terms "interpretation bias" OR "expectancy bias" OR "inferential bias" were combined with the following separate search terms: heart (rate, variability), skin (response), galvanic, cortisol, pupil* (pupillometry, pupillary, pupil), eye track, barore* (baroreflex, baroreceptor), vagal, blood pressure, electroencephalog*(electroencephalography, electroencephalograph, electroencephalographic), EEG, event related potential, ERP, *MRI (e.g. fMRI, structural MRI), imaging, electromyog* (electromyography, electromyogram, electromyographic), EMG, respiratory sinus arrhythmia, parasympathetic, psychophys* (psychophysiological, psychophysical, psychophysiologic), phys* (physiological, physiology, physiologic), bio*

(biological, biology, biomarkers). Additionally, article reference lists were scanned. We identified grey literature following the recommendations of Rothstein and Hopewell (2009). Namely, the above search strategies were conducted on: Google search and scholar, Cochrane Library, controlled-trials.com, greylit.org, opengrey.eu, ntrl.ntis.gov, biomedcentral.com, ovid.com and World Cat.

The inclusion criteria were: written in English, adult participant population (18 years or over), sampling from either clinical (e.g. a group of participants who have received a diagnosis via a standardised assessment such as the Structured Clinical Interview for DSM-5) or sub-clinical/analogue (e.g. as having 'high' levels of worry, anxiety, or depression, for example as indicated by a cut-off on a questionnaire measure that has clinical and community norms) populations, peer reviewed and/or grey literature research, include at least one assessment of interpretation bias[1], include at least one assessment of psychophysiology, tasks present emotionally ambiguous information. The exclusion criteria were: Review/opinion papers, qualitative designs, alternative biases (e.g., attention bias), beyond the target literature of emotional ambiguity (e.g., those assessing racial or prejudicial views).

Quality Assessment

Study quality, including both peer-reviewed and grey literature studies, was assessed

[1] Hirsch and colleagues (Hirsch et al., 2016, p. 282) define interpretation as "the product of the semantic process by which ambiguity is resolved". Any task that measures interpretation must therefore provide the opportunity for an ambiguous situation to be disambiguated in a negative or benign/positive manner. If an assessment did not do this (e.g. assessment of facial expressions), then it was not included in the review.

using a quality assessment checklist (adapted from Downs & Blacks Quality Checklist for Healthcare Intervention Studies, 1998). The checklist examines key areas relating to study reporting quality (11 items, e.g., are predictions clearly articulated?) external validity (1 item, e.g., generalisability), internal validity (6 items, e.g., risk of bias, risk of confounding variables) and physiological measure and data quality (5 items based on Sörnmo & Laguna, 2005), providing a total of 23 items and a maximum quality score of 27. Each item was scored 0 or 1, with the exception of four items which could be scored from 0 to 2. Comparison of the quality scoring of studies made by the two members (i.e., inter-rater reliability) produced an agreement rate of 81% and a Cohen's kappa value of 0.78, which indicates moderate strength of agreement between the two scorers. This value of kappa is significantly different from zero ($\kappa = 0.78$, $p < 0.001$). See supplementary materials for a copy of the checklist.

Coding

A standardized data coding form was created in order to extract key information from the respective studies: 1. authors/title paper, 2. publication year, 3. hypotheses, 4. the country study conducted in, 5. language/ethnicity, 6. sample size, 7. sub-groups, 8. mean age, 9. gender, 10. study design, 11. Main independent variable (IV), 12. IV manipulation, 13. main dependent variable (DV), 14. (N) physiological measure, 15. (N) IB measure, 16. emotional construct, 17. CBM inclusion, 18. data analysis strategies, 19. effect size, 20. significance testing outcome, 21. key effect of IV on DV, 22. key outcome, 23. implication(s), 24. key limitations.

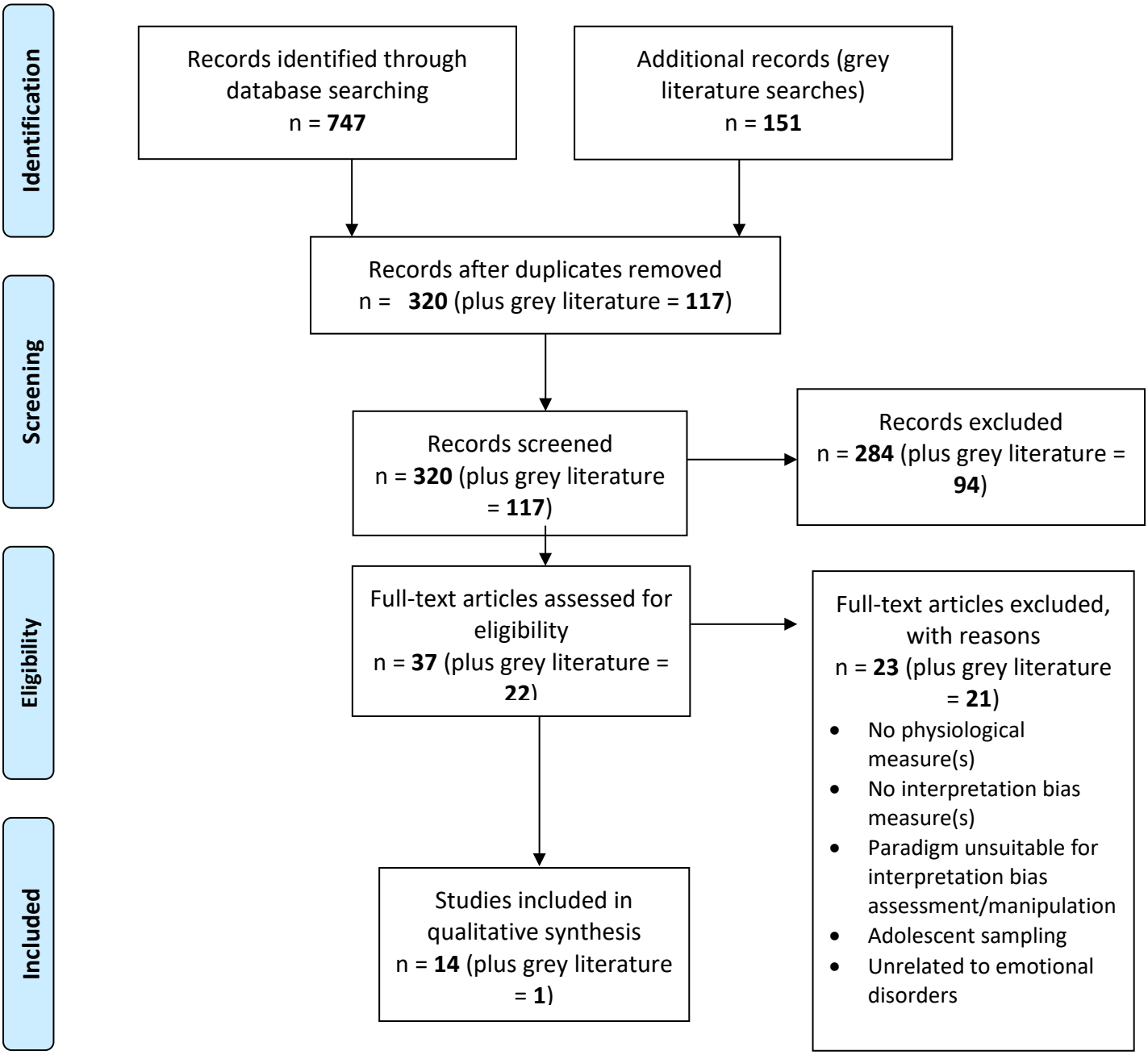

Figure 1: PRIMSA flow diagram of study selection/screening procedure

Results

Overview of studies and participant characteristics

A total of 898 studies were identified (see Figure 1.). After duplicates were extracted 437 remained. Titles and abstracts were screened resulting in removal of 378 studies. Fifty-nine full-text articles were assessed, of which 15 met criteria for inclusion in the review. *Table 1* depicts study demographics (e.g., sample sizes, country, gender, and mean age) whereas *Table 2* provides experimental results (e.g., interpretation bias (IB) measure, physiological measure, emotional construct, inclusion of CBM-I, effects of IV on DV and quality assessment score). Overall, the present review identified seven physiological measures associated with interpretation bias, namely: ERP, HR, HRV, SCR, EMG and pupillometry. From the fifteen studies, five studies examined physiological measures (typically HR, HRV and SCR) in combination (Beadel et al., 2013; Grisham et al., 2014; Nowakowski et al., 2015; Rozenman et al., 2020; Van Bockstaele et al., 2020). HR was the most frequently measured physiological variable (Beadel et al., 2013; Grisham et al., 2014; Howard et al., 2018; Joormann et al., 2015; Nowakowski et al., 2015; Rozenman et al., 2020; Van Bockstaele et al., 2020), followed by ERP (Feng et al., 2019, 2020; Moser et al., 2009, 2012), followed by SCR (Beadel et al., 2013; Howard et al., 2018; Nowakowski et al., 2015; Rozenman et al., 2020), them HRV (Grisham et al., 2014 ; Meeten, 2017 – grey literature; Van Bockstaele et al., 2020), EMG (Lawson et al. 2002; Whitton et al., 2013), RSA (Rozenman et al., 2020), and pupillometry (Cowden Hindash et al., 2021). The key findings obtained from the studies are discussed below. For further information concerning descriptions of physiological measures and/or tasks employed in the studies (e.g., interpretation bias tasks or stress-reactivity tasks), see the supplementary material.

[insert table 1 here]

[Insert table 2 here]

Quality assessment

The adopted quality assessment tool scored key aspects such as reporting quality, external and internal validity and physiological data quality. Overall, studies yielded a mean quality score of 17.31 (*SD* = 3.28), with a minimum and maximum score ranging from 11-21. Thus, we can infer a reasonable standard of study quality within the present review.

Reporting quality: Generally, studies performed consistently well within this domain, with nearly all studies achieving maximal scoring for reporting on study aims/hypotheses, demographics, experimental tasks, and outcomes in sufficient detail and clarity. Seventy five percent of studies reported estimates of the random variability in behavioural outcome data, as well as for reporting exact *p*-values underlying significance testing (reporting considered to be best practice). However, 13% of the studies failed to report associated effect sizes of key outcomes (e.g., Cohen's d, η^2, etc.). Study pre-registration is a crucial step for upholding academic integrity and the principals that support optimal scientific practice/reporting, whilst overcoming the replication crisis associated with psychological research (Cumming, 2014). Determining pre-registration status was therefore included as a key criteria within this quality domain. However, given that reviewed studies were conducted at different time points, the year of publication was taken into consideration (i.e. it was recognised that older studies would pre-date pre-registration practices). A cut-off of 2014 was applied whereby there was felt to be a reasonable expectation for studies to be pre-registered if they were published

after that year (note, we acknowledge that this is a somewhat arbitrary date based on the authors' best judgement of pre-registration practice; Cumming, 2014). Sixty percent of studies were not pre-registered. Finally, studies were assessed according to employment of data sensitivity tests concerning inferential outcomes. For example, sensitivity testing may include the use of Bayes factors (together with other strategies such as stopping rules), which, unlike orthodox null hypothesis significance testing, can help distinguish the relative strength of evidence for the hypotheses being tested i.e. results can be shown to sensitively support the alternative hypothesis (H1), sensitively support the null hypothesis (H0), or be insensitive in which case judgement must be suspended (Dienes, 2014). However, no studies scored for this criteria, indicating that significance testing is currently the dominant approach for hypothesis testing within this research field.

External validity: All studies scored maximally for the generalisability criteria (i.e., the extent to which participants were generalisable to the target population from which they were recruited).

Internal validity (risk of bias or confounds): The majority of studies failed to sufficiently render participants naïve to research goals/agenda, whilst there was also a marked degree of variability in scoring concerning attempts to blind the researchers measuring the main outcomes (46.60% fulfilled this criteria). Whilst all studies achieved maximum marks for implementing the correct/appropriate inferential tests for analysing the main outcomes, only 20% ensured appropriate validation and reliability tests for the main outcomes measured (e.g., reporting upon test-statistics concerning the reliability and/or validity of interpretation bias tasks/assessments etc). Crucially however, all studies ensured random

allocation procedures for participants to conditions.

Psychophysiological data quality: Overall, there was a marked degree of variability in scoring within this domain, although it can be argued that the wide range of physiological measures employed across the studies may play a role in generating some inconsistencies. Just over half of studies (60%) included pre-screening assessments for (potentially) confounding variables, which may otherwise affect the quality of psychophysiological data, namely medical history, caffeine intake, etc. The majority of studies (86.67%) reported baseline physiological measures, whilst 60% of studies provided estimates of the random variability in physiological outcome data; 73.33% of studies provided raw, unstandardised unites of physiological data within the write up of (descriptive) results, which is important for the purposes of external researcher assessment. Finally, 66.66% studies reported the pre-processing steps in effective detail (e.g., artifact correction methods, etc.), which is important for discerning aspects of the signal-to-noise ratio, reliability etc.

Physiological measures associated with interpretation bias

1. Online assessments of interpretation bias and physiological measures

Four studies employed an EEG method to examine ERPs associated with a cognitive assessment of interpretation bias. All studies assessed interpretation bias using the lexical decision task (LDT). Using this methodology, EEG was assessed in parallel to the assessment of bias via the LDT task. Feng et al. (2019) found that whilst N400 amplitude was highest among low worriers during negatively vs. positively disambiguated (worry related) scenarios (indicating that negative outcomes were unexpected for low worriers), high worriers

showed no differences in amplitude to negative or positive disambiguation, suggesting that high worriers lack the positive interpretation bias observed in low worriers. By contrast, Feng et al. (2020) did not find an ERP correlate associated with (imagery-enhanced) CBM-I in high worriers, suggesting that multiple training sessions may be required in order to capture training-congruent effects at the neural level.

Moser et al. (2008) found larger P600 responses in participants scoring low in social anxiety during negatively disambiguated (socially related) scenarios, whilst participants scoring high in social anxiety showed no differences in amplitude to negative or positive scenario resolution. However, findings from their later study do not concur (Moser et al., 2012); interpretation bias was not discernible from the P600 marker, although N400 was a correlate of positive interpretation bias in healthy participants but not in the emotionally vulnerable clinical groups (i.e., those with major depressive disorder (MDD), or social anxiety and co-morbid dysphoria). Interestingly, the P600 amplitude was observed to be generally smaller in the clinical groups regardless of emotional valence/disambiguation, which the authors suggest may reflect a trend of dampened baseline attention.

Lawson et al. (2002) examined EMG correlates of interpretation bias by recording the startle eye blink response (Lawson et al., 2002). Eye-blink is part of the startle response and larger blink reflex magnitudes are associated with negative as opposed to neutral imagery (Bradely et al., 1999). Lawson et al. (2002) thus proposed that eye blink magnitude can provide a measure of interpretation that (unlike cognitive measures) is not influenced by response bias and demand effects. Lawson and colleagues identified markedly larger startle reflexes (eye-blink responses) when participants, scoring highly on depressive traits, were

instructed to generate an (imagined) interpretative response following the presentation of emotionally ambiguous acoustic stimuli, as compared to those participants scoring lower on the depressive scale who indexed smaller reflexes.

Pupillometry was also employed as a method of assessing potential physiological correlates of interpretation bias. Pupil size is another physiological marker that is influenced by the autonomic system. Pupil size is determined by two pathways the *parasympathetic constriction pathway* and the *sympathetic dilation pathway*. Pupil dilation is thus associated with wakefulness, feelings of high arousal, and the fight-flight response (Mathôt, 2018). It is also well established that pupil dilation provides an index of cognitive processing or load (e.g. Beatty & Kahneman, 1966). Only one study assessed pupillometry indices of interpretation bias. Cowden Hindash et al. (2021) recruited participants with a diagnosis of major depressive disorder (MDD) and healthy controls, examining pupil dilatory responses during interpretation bias assessment (i.e., a sentence-word association task where an ambiguous sentence is presented, followed by a single word that could either be associated with a negative or a benign interpretation of the sentence with the participant required to respond as quickly as possible to say whether or not the word is associated with the sentence). They examined pupil reactivity, i.e. the extent of dilation as compared to a baseline dilation. Results showed markedly greater pupil dilation in the MDD group occurring in negatively based trials (specifically, when participants reject negative word associations) compared to healthy controls, who showed no difference in pupil reactivity when rejecting a negative interpretation. The results suggest that those in the MDD found it more effortful to reject negative word associations as doing so contradicts their trait bias. The authors suggests that pupil reactivity might be able to differentiate between trait

interpretation bias tendencies.

Finally, Howard et al. (2018) assessed HR and SCR during interpretation bias assessment between participants characterised as having type-D personality with participants having non-type-D personality. Increases in HR and in electrodermal activity from moment to moment are associated with physiological arousal, commonly associated with a stress response, but as noted above can also be associated with positive responses such as surprise. As noted by Howard et al. (2018), Type D personality is associated with social inhibition and high levels of negative affectivity. As a result, socially ambiguous scenarios were selected as an appropriate interpretation bias measure. Here, the cognitive measure of interpretation was a scenarios task where participants were presented with social scenarios that were clearly negative, clearly neutral, ambiguously negative, or ambiguously neutral. Participants were asked to read the scenarios and rate how difficult they would find it to respond verbally to that situation and how distressing and how threatening they would find the situation. Results indicated that type-Ds showed a significantly higher HR response during discernment of ambiguously neutral scenarios (i.e., when scenarios remained unresolved, probing participants' own biases). Within-group analysis revealed higher HR during clearly negative and ambiguously neutral situations in type-D versus non-type-Ds who indexed lower HR. There was no effect of ambiguity on SCR.

2. Near-transfer effects of interpretation bias training and physiological measures

Only one study from the grey literature (Meeten et al., 2017) attempted to map physiological change associated with interpretation bias training. The study examines changes in HRV on the basis that lower HRV has been associated with a number of anxiety disorders

(cf. Chalmers et al., 2014) and higher HRV has been associated with effective emotion regulation and adaptive functioning (Thayer et al., 2009). Specifically, Meeten et al. (2017) engaged participants with high trait levels of worry in either positive interpretation bias training, or in a sham training (control condition). Two cognitive measures of interpretation bias together with resting state HRV were assessed pre and post training. No change was found on the cognitive measures. However, HRV increased significantly in the positive vs. the sham training condition. The authors note that these results should be interpreted with caution as no training effect was seen on the cognitive measures.

3. Far-transfer effects of interpretation bias training and physiological markers

Six studies measured HR as a proxy marker of stress reactivity post CBM-I intervention, typically assessed during an independent stressor task. For example, Nowakowski et al. (2015) found no marked HR differences in participants who were socially anxious following positive CBM-I, as compared to control CBM-I, and an alternative cognitive restructuring group. Joormann and colleagues (2015) similarly found no change in HR after participants with MDD received positive CBM-I, although negative CBM-I resulted in an elevated pattern of HR (significant increase from baseline to anticipation phases of the stress reactivity task), yet there were no HR differences during the stress recovery phases across any conditions. By contrast, Van Bockstaele et al. (2020) found that positive CBM-I resulted in significantly lower HR during the stress recovery phase, suggesting that participants showed a faster recovery rate from induced stress, as compared to the negative CBM-I group. Interestingly, self-reported stress did not concur with HR/physiological findings; participants still reported high levels of distress. Grisham et al. (2014) identified a significant HR increase from baseline to

stress-induced phases of stress reactivity after participants (scoring highly on obsessive compulsive disorder traits) received negative CBM-I (designed to inflate responsibility bias); positive CBM-I on the other hand did not result in any HR directional changes. However, in accordance with Van Bockstaele and colleagues (2020), Grisham et al. similarly obtained inconsistencies between self-reported stress and physiological indices.

Rozenman and colleagues (2020) assessed HR pre and post four-week CBM-I intervention (12 sessions) in participants scoring highly on self-reported trait anxiety. The positive CBM-I intervention resulted in significantly reduced HR during the stress-recovery phase of the stressor task, as compared to control CBM-I. This was the only study to assess physiological indices across multi-session training, findings of which are important for discerning long-term (stable) physiological effects associated with interpretation bias modification. Beadel et al. (2013) assessed HR during CBM-I as well as during post-CBM-I stress reactivity; participants scoring highly in obsessive compulsive disorder traits were allocated to either positive CBM-I or neutral CBM-I (half negative, half positive trials). The authors found no change in HR for any condition or during the stressor tasks. It should be highlighted that this was the only study in which physiological indices were inversely related to predictions of interpretation bias. Specifically, it was hypothesised that CBM-I works by modifying cognitive contingency processes – participants learn to associate emotional ambiguity with positive resolution, as opposed to through fear exposure and desensitisation; the latter of which is argued to evoke physiological arousal. Thus, physiological reactivity was predicted to remain constant during CBM-I.

Two studies assessed HRV in relation to interpretation bias, where phasic HRV

suppression is typically associated with a physiological stress response (e.g. Park et al. 2014). For participants receiving positive CBM-I, Grisham et al. (2014) identified significantly higher HRV (and concurrently lower HR) during the post-stressor task as compared to baseline, whereas participants in the negative condition showed no change between baseline and stressor task. The authors contend that this cannot be explained by negative training containing more negative scenarios, because self-reported stress and negative affect were apparently equal for both training conditions during stress reactivity assessment. However, as seen above and noted elsewhere, there can exist marked discordance between self-report and physiological measures during emotional (or stress) reactivity assessments (Lang et al. 2017). Participants undergoing CBM-I may be unaware of the acute changes occurring at the physiological level.

Van Bockstaele et al. (2020) did not find HRV differences between positive and negative CBM-I conditions. However, positive CBM-I with the addition of cognitive load resulted in significantly increased HRV at post training. Finally, Meeten (2017) found that positive CBM-I was associated with a significant increase in HRV from baseline to post training, as compared to control CBM-I in participants with high-trait worry scores, where no such significant increase was observed.

An additional parameter of HRV is respiratory sinus arrythmia (RSA), typically used as an index of cardiac vagal tone, reflecting interactions between the sympathetic and parasympathetic nervous systems. Only one study assessed the extent to which RSA changes as a result of CBM-I; Rozenman et al. (2020) calculated the difference between minimum and maximum HR changes during respiration. However, there were no changes or difference

in RSA associated with CBM-I, which suggests that CBM-I training was not associated with a reduced stress response as measured by RSA. However, this is at odds with findings on other physiological indexes (SCR and HR) measured in the Rozenman et al. (2020) paper. One possibility is that certain physiological parameters may be less sensitive to the effects of the training as compared to others (i.e., HRV, HR).

Three studies sought to assess SCR responses related to a stressor task post CBM-I, although one study was unable to report SCR indices due to technical/equipment issues and thus results are omitted here. Again, returning to Rozenman et al. (2020), higher SCR was associated with stress reactivity after control CBM-I (four weeks of training), as compared to positive CBM-I. The data suggests that baseline SCR response did not differ between positive and control CBM-I pre-intervention, whilst SCR decreased post positive CBM-I. Taken together, these converging physiological indices reflect reduced sympathetic arousal occurring after multi-session positive CBM-I, namely SCR and HR decreases. In accordance with predictions, Beadel et al. (2013) found that SCR did not differentiate between or change across CBM-I condition nor subsequent stress reactivity.

One study assessed contractions of the levator labii, a muscle group shown to be associated with expressive facial responses of disgust (Whitton et al., 2013). Here, Whitton et al. (2013) sought to assess the malleability of disgust-based interpretational responses via CBM-I, indexed via EMG measures of levator labii muscle contractions. After participants received either positive CBM-I or negative CBM-I (designed to induce disgust-related biases), they were exposed to a number of disgust-eliciting stimuli whereupon EMG facial recording was conducted. Despite some behavioural-based training effects, the authors did not

capture EMG-related responses differentiating CBM-I conditions.

Discussion

The purpose of the current review was to systematically examine physiological correlates and sequalae of interpretation bias. Research to date has examined the association between interpretation bias and physiology using a number of different paradigms and using a range of different physiological measures. Searches revealed seven potential measures, namely ERP, HR, HRV/RSA, SCR, EMG and pupillometry. Studies examined physiological markers of the bias at the same time as the bias was being measured by a cognitive task, for example ERPs (Feng et al., 2019; 2020), EMG (Lawson et al., 2002), HR (Howard et al., 2018), and SCR (Howard et al., 2018). Only one study examined change in physiology via HRV before and after positive interpretation bias training (Meeten et al., 2017), whereas HR, HRV, and SC measures were typically assessed during an independent stressor task (Beadel et al., 2013; Grisham et al., 2014; Joormann et al., 2015; Nowakowski et al., 2015; Rozenman et al., 2020; Van Bockstaele et al., 2020). The main theoretical and practical implications will be discussed, following recommendations for future research aiming to expand upon the current literature.

According to the reviewed articles, the relationship between psychophysiology and interpretation bias can be examined both indirectly and directly. For example, a number of studies sought to (indirectly) investigate CBM-I efficacy via physiological measures of stress reactivity. Such studies employed a range of stressor tasks following CBM-I (see supplementary materials), assessing potential differences in stress arousal (e.g., anticipation, recovery, etc.), emotional reactivity, and disorder-specific symptoms between

conditions (usually between positive and negative bias training). Several authors highlight the need to include concurrent subjective measures during emotional/stress assessment to reduce demand characteristics or enhance validity (Grisham et al., 2014; Lang et al., 2017). The few studies that did include both subjective and objective measures of emotion/stress, i.e. those taking self-report and physiological indices of distress, found marked discordance between these measures (Grisham et al., 2014; Van Bockstaele et al., 2020). However, the primary goal of such studies was to investigate internalising symptom changes associated with CBM-I. Thus, caution is warranted before assuming physiological markers of interpretation bias *per se*, or even CBM-I, since such physiological responses were obtained during remote/independent stressor tasks that do not measure interpretation bias. However, such paradigms are useful for discerning far-transference effects of CBM-I, an important step in the generalisability of bias intervention for emotional disorder treatment. In addition, such findings provide a novel way for adapting CBM-I that aims to target somatic symptoms associated with anxiety and mood disorders (e.g., such as biofeedback that has proven effective for improving a range of emotional and mood disorder symptoms, cf. Lehrer et al., 2020), symptoms that play a crucial role in the maintenance of psychopathology (Hyde et al. 2019).

Are there physiological markers of interpretation at the time that the interpretation is made?

There was considerably consistent evidence for the ERP N400 markers of interpretational responses (Feng et al., 2019; Moser et al., 2009, 2012), although this effect was not found by Feng et al. (2020). ERPs may provide information about the processes occurring during the resolution of ambiguity. There was some evidence to suggest that

autonomic responses to a negative interpretation also occur at early stages in processing, for example from studies examining EMG measures (Lawson et al., 2002) and HR (Howard et al., 2018), but no evidence of electrodermal changes in the study examining SCR during an online task (Howard et al., 2018). Understanding the time course in which interpretation occurs allows consideration of how we might better intervene in bias modification. However, the findings from this review indicate that an interpretation is likely to have already occurred by this point. Future therapeutic approaches might consider exploring biofeedback techniques which support modification of the bias at these early stages in information processing. Future research in this area might also extend the experimental findings by collecting peripheral physiology data (HR, SCR, HRV etc.) at the same time as measuring ERPs to further develop our understanding of the physiological responses to negative interpretation.

Are there physiological markers of change in interpretation bias (near and far transfer effects)?

Meeten et al. (2017) reported an increase in HRV from before to after positive interpretation bias training. One interpretation of this type of physiological change is that we are observing a reduction in a generalised threat response. However (as noted above), this study did not observe pre- to post-training effects on cognitive measures of bias and thus replication of this work is warranted before any firm conclusions can be made about physiological changes associated with cognitive bias training. There is evidence to support the idea that positive interpretation bias training results in a reduced stress response on a subsequent stress task as measured by HR (Grisham et al., 2014; Howard et al., 2018;

Joormann et al., 2015; Rozenman et al., 2020), HRV (Grisham et al. 2014; Van Bockstaele et al., 2020), and SCR (Rozenman et al., 2020). However, this was not uniformly observed with some studies not finding the predicted association between positive CBM-I training and a blunted stress response in a subsequent stressor task on SCR (Beadel et al., 2013; Nowakowski et al., 2015), on HR (Van Bockstaele et al., 2020) and on EMG (Whitton et al., 2013).

Methodological considerations in interpreting these research findings

Several factors are likely to have influenced the research findings examined in this review. For example, the emotional saliency of ambiguity (discussed below), type of interpretation bias examined or tasks employed (note that studies assessed a variety of bias types, such as threat, disgust, anxiety, worry, etc.), CBM-I dose and, in some cases, the measurement of concurrent physiological parameters. The psychophysiological literature suggests that robust autonomic modulations, associated with behavioural modification paradigms, usually unfold gradually, over a longer time period. Thus, repeated training or dose may be required to observe effects at the physiological level (Ottaviani et al., 2017). For example, only one study (including concomitant measures) detected SC changes pre- and post- multi-session CBM-I during stress reactivity (Rozenman et al., 2020), yet SCR alone reveals minimal information of autonomic behaviour unless combined with directional patterns of HR or HRV (Hyde, Ryan & Waters, 2019). In addition, specific biases may play a more prominent role in eliciting autonomic arousal as compared to others. For example, it is more likely that anxiety or threat-orientated bias possess a clear physiological basis (Bockstaele et al., 2020; Rozenman et al., 2020) than disgust or responsibility biases (Beadel et al., 2013; Whitton et al., 2013). This suggests that physiological measures may be better

placed to detect affective processes underlying interpretational responses. Indeed, it is well known that affect and physiological regulation are closely tied across autonomic systems (Moors, 2009). However, this brings another important point to attention concerning general valence effects, as discussed next.

Alongside rigid cognitive patterns (such as trait negative interpretation bias), emotional disorders also present a range of somatic symptoms that can result in inflexible physiological reactivity (Brosschot et al., 2007), as well as prolonged activation of sympathetic nervous system activity or blunted parasympathetic nervous system activity (Brosschot, 2010; Diamond, & Fisher, 2017; Hyde et al., 2019; Lang et al., 2017; Ottaviani et al., 2016). Thus, resting state activity in those with emotional disorders may produce exaggerated or weakened biological reactions during interpretative responses, particularly if interpretation assessment involves emotional, or disorder-specific, ambiguity. This could potentially lead to an increase of false-positives (or negatives) when comparing CBM-I conditions or participant groups (e.g., emotionally vulnerable vs. healthy). For example, Joorman and colleagues (2015) found that negative CBM-I resulted in significantly elevated HR from baseline to anticipation phases of the post stress reactivity task, whereas the non-bias training condition (nor positive CBM-I), induced no such HR changes, yet the authors point out that this may be because the depression diagnosed participants within the non-training condition demonstrated overall elevated baseline HR. Thus, it is important that studies acknowledge potential psychophysiological baseline differences to avoid such ceiling effects, depending upon the participant population or disorder-specific emotional vulnerability or severity.

Given that the present review focuses on an emerging (and expanding) research field, researchers may continue developing and adapting interpretation bias tasks to meet the constraints (or requirements) of specific physiological measures. As already noted, bias measures can capture both implicit and explicit interpretational processing, each of which possess distinct theoretical implications, advantages and disadvantages (Hirsch et al., 2016; Schoth & Liossi, 2017). As compared to ERP studies examining fast, initial processing (Feng et al., 2019, 2020; Moser et al., 2009, 2012), the review is largely lacking studies assessing physiological indices associated with explicit interpretation bias measures. One study sought to identify averaged HR and SCR responses measured concurrently during the ambiguous scenario task, according to type of emotional ambiguity (e.g., clearly negative; clearly neutral; ambiguously negative; ambiguously neutral) alongside self-reports of anticipated distress, perceived threat and difficulty of providing a verbal response (Howard et al., 2018). McCleod and colleagues (2002) measured the magnitude of the eye-blink reflex, via EMG, during participants' resolution of emotional ambiguity in imagination (following the presentation of acoustically presented ambiguous stimuli). Meeten et al., (2017) found that higher HRV was associated with participants' rejection of negatively disambiguated scenarios in the recognition task. Given that such explicit bias measures require a marked degree of contemplative thinking on the part of the participant, it is important for future researchers to identify suitable physiological markers that best capture reflective interpretative responses.

Limitations

There are a number of limitations to the evidence discussed in this review. The quality

assessment process highlighted a number of weaknesses that could be addressed in future research and while those addressed in the quality assessment tool are outlined above, it is worth highlighting some key areas. For example, very few studies examined reliability and validity of key outcomes measures. Of note is that many studies did not present a rational for sample size decisions. Most studies measured a number of dependent variables, both cognitive and physiological. Future work exploring the physiological bases of interpretation bias should ensure that research studies are sufficiently powered to examine the physiological variable of interest. The majority of studies in the review reported effect sizes relevant to the dependent variables, which will support future researchers in their sample size calculations. Other important limitations relate to a lack of diversity in research participant samples. All studies except one (Van Bockstaele et al., 2020) were conducted in countries where English is the main language and all studies were in American, Europe, or Australia, with majority female samples. Future research should examine both cognitive and physiological markers of interpretation bias in more diverse samples.

Clinical Implications

There are key important clinical implications of this work. Our understanding of the dynamic interaction between mind and body is acknowledged in a number of existing therapeutic interventions. For example, the applied relaxation approach to the treatment of anxiety (Hayes-Skelton et al., 2013) as well as transcutaneous vagus nerve stimulation and HRV biofeedback interventions for mental health (Lehrer et al., 2020; Vanderhasselt & Ottaviani, 2022). Developing our understanding of the physiological bases of interpretation bias can help us understand more about the process of interpretation bias. As discussed by

Hirsch et al. (2016), we need to understand more about the stage in information processing at which interpretation occurs in order to pinpoint targets for intervention. EEG studies (Feng et al., 2019, Moser et al., 2008, 2012) suggest that expectancy violations (indicative of a cognitive bias) can be observed at earlier stages in processing at the point of encountering and initially resolving ambiguity. This suggests that bias modification could usefully target early stages of processing, thus challenging the automaticity of biased interpretation. However, not all studies using EEG showed the predicted physiological signature (e.g. Feng et al. 2020; although this study assessed the presence of bias and was not an interpretation bias training study) and further research in this area is warranted to consolidate knowledge in this area. Many studies examined in the present review examine the effect of CBM-I on physiological responses to post-training stress assessment task. These methodologies do not provide direct information about the association between bias change and change in physiology. To further develop the clinical implications of this area, it will be important to look at the direct association between change in bias via CBM-I and pre-post change in physiology. Once we have better understanding of how interpretation bias correlates with physiological measures, we can begin to think about which physiological markers can be used to support bias change. For example, we could seek to exploit those physiological markers to create a training environment in which individuals attempt to shape positive expectations at the earliest possible stage. While doing this they could receive real-time feedback on their effectiveness as reflected in their physiological responses, and would ultimately be presented a training scenario outcome consistent with those expectations. It is anticipated that both the real-time feedback and outcome dependency would provide a direct form of reward, thus potentially enhancing the learning process in bias training.

In summary, the present review identified seven psychophysiological measures associated with interpretation bias: ERP, HR, HRV, RSA, SCR, EMG and pupillometry. Whilst there is a clear need, and application of, physiological measures in interpretation bias research, there are also important practical and theoretical implications to consider. For example, it is important to distinguish between studies assessing direct associations between interpretative biases and physiological indices and those assessing indirect associations between CBM-I efficacy and physiology (via stress reactivity). However, future CBM-I intervention studies may wish to incorporate physiological markers to examine processes of *bias change*, bridging the gap between mind and body and in turn providing novel treatment strategies for emotional disorders. It is clear from the review that differing physiological measures impose differing demands or requirements, which must be factored into interpretation bias research accordingly (as discussed above). The neuroscientific literature of interpretation bias is newly emerging but holds potential for developing existing theoretical models and treatment approaches in new and exciting ways.

Acknowledgments

Preparation of this manuscript was supported by a Medical Research Council (MRC) New Investigator Grant (MR/W005077/1) awarded to FM, CO, and CH. CH receives salary support from the National Institute for Health Research (NIHR), Mental Health Biomedical Research Centre at South London and Maudsley NHS Foundation Trust, and King's College London. The views expressed are those of the authors and not necessarily those of King's College London, NHS, National Institute for Health Research (NIHR), or the MRC.

References

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition: Basic processes; Applications* (pp. 1–40). Lawrence Erlbaum Associates, Inc

Beadel, J. R., Smyth, F. L., & Teachman, B. A. (2014). Change processes during cognitive bias modification for obsessive compulsive beliefs. *Cognitive Therapy and Research*, *38*(2), 103-119.

Beatty, J., & Kahneman, D. (1966). Pupillary changes in two memory tasks. *Psychonomic Science, 5*(10), 371–372. https://doi.org/10.3758/BF03328444

Beck AT, Clark DA. 1997. An information processing model of anxiety: automatic and strategic processes. *Behavior Research & Therapy. 35*, 49–58.

Beck, A. T., & Haigh, E. A. (2014). Advances in cognitive theory and therapy: The generic cognitive model. *Annual Review of Clinical Psychology*, *10*, 1-24.

Brosschot, J. F. (2010). Markers of chronic stress: Prolonged physiological activation and (un) conscious perseverative cognition. *Neuroscience and Biobehavioral Reviews, 35*, 46-50.

Brosschot, J. F., Gerin, W., & Thayer, J. F. (2006). The perseverative cognition hypothesis: A review of worry, prolonged stress-related physiological activation, and health. *Journal of Psychosomatic Research, 60*, 113-124.

Brosschot, J. F., Van Dijk, E., & Thayer, J. F. (2007). Daily worry is related to low heart rate variability during waking and the subsequent nocturnal sleep period. *International Journal of Psychophysiology, 63*(1), 39-47.

Brosschot, J. F., Verkuil, B., & Thayer, J. F. (2010). Conscious and unconscious perseverative cognition: Is a large part of prolonged physiological activity due to unconscious stress?. *Journal of psychosomatic research, 69*(4), 407-416.

Butler, G., & Mathews, A. (1983). Cognitive processes in anxiety. *Advances in Behaviour Research and Therapy, 5*(1), 51-62.

Chalmers, J. A., Heathers, J. A., Abbott, M. J., Kemp, A. H., & Quintana, D. S. (2016). Worry is associated with robust reductions in heart rate variability: a transdiagnostic study of anxiety psychopathology. *BMC Psychology, 4*(1), 1-9.

Cowden Hindash, A. H. C. (2018). *An Experimental Examination of Automatic Interpretation Biases in Major Depression*. University of South Florida. Unpublished thesis.

Davey, G. C., & Meeten, F. (2016). The perseverative worry bout: a review of cognitive,

affective and motivational factors that contribute to worry perseveration. *Biological Psychology, 121*, 233-243.

Diamond, A. E., & Fisher, A. J. (2017). Comparative autonomic responses to diagnostic interviewing between individuals with GAD, MDD, SAD and healthy controls. *Frontiers in Human Neuroscience, 10*, 677.

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 randomised and non-randomised studies of health care interventions. *Journal of Epidemiology & Community Health*, *52*(6), 377-384.

Eguchi, A., Walters, D., Peerenboom, N., Dury, H., Fox, E., & Stringer, S. (2017). Understanding the neural basis of cognitive bias modification as a clinical treatment for depression. *Journal of consulting and clinical psychology*, *85*(3), 200.

Ehring, T., & Watkins, E. R. (2008). Repetitive negative thinking as a transdiagnostic process. *International Journal of Cognitive Therapy, 1*(3), 192-205.

Everaert, J., Koster, E. H., & Derakshan, N. (2012). The combined cognitive bias hypothesis in depression. *Clinical psychology review*, *32*(5), 413–424. https://doi.org/10.1016/j.cpr.2012.04.003

Everaert, J., Podina, I. R., & Koster, E. H. W. (2017). A comprehensive meta-analysis of interpretation biases in depression. *Clinical Psychology Review*. doi:https://doi.org/10.1016/j.cpr.2017.09.005

Eysenck, M. W., Mogg, K., May, J., Richards, A., & Mathews, A. (1991). Bias in interpretation of ambiguous sentences related to threat in anxiety. *Journal of Abnormal Psychology, 100*(2), 144-150.

Feng, Y. C., Krahé, C., Meeten, F., Sumich, A., Mok, C. M., & Hirsch, C. R. (2020). Impact of imagery-enhanced interpretation training on offline and online interpretations in worry. *Behaviour Research and Therapy*, *124*, 103497.

Feng, Y. C., Krahé, C., Sumich, A., Meeten, F., Lau, J. Y., & Hirsch, C. R. (2019). Using event-related potential and behavioural evidence to understand interpretation bias in relation to worry. *Biological Psychology*, *148*, 107746.

Frewen, P. A., Dozois, D. J., & Lanius, R. A. (2008). Neuroimaging studies of psychological interventions for mood and anxiety disorders: empirical and methodological review. Clinical Psychology Review, 28(2), 228-246.

Grey, S., & Mathews, A. (2000). Effects of training on interpretation of emotional ambiguity. The Quarterly Journal of Experimental Psychology: Section A, 53(4), 1143-1162.

Grisham, J. R., Becker, L., Williams, A. D., Whitton, A. E., & Makkar, S. R. (2014). Using cognitive bias modification to deflate responsibility in compulsive checkers. *Cognitive Therapy and*

Research, *38*(5), 505-517.

Hayes-Skelton, S. A., Roemer, L., Orsillo, S. M., & Borkovec, T. D. (2013). A contemporary view of applied relaxation for generalized anxiety disorder. *Cognitive Behaviour Therapy*, *42*(4), 292–302. https://doi.org/10.1080/16506073.2013.777106

Hallion, L. S., & Ruscio, A. M. (2011). A meta-analysis of the effect of cognitive bias modification on anxiety and depression. *Psychological Bulletin*, *137*(6), 940–958. https://doi.org/10.1037/a0024355

Hirsch, C. R., & Mathews, A. (2012). A cognitive model of pathological worry. *Behaviour Research and Therapy*, *50*(10), 636-646.

Hirsch, C. R., Krahé, C., Whyte, J., Bridge, L., Loizou, S., Norton, S., & Mathews, A. (2020). Effects of modifying interpretation bias on transdiagnostic repetitive negative thinking. *Journal of Consulting and Clinical Psychology, 88*(3), 226.

Hirsch, C. R., Krahé, C., Whyte, J., Krzyzanowski, H., Meeten, F., Norton, S., & Mathews, A. (2021). Internet-delivered interpretation training reduces worry and anxiety in individuals with generalized anxiety disorder: A randomized controlled experiment. *Journal of Consulting and Clinical Psychology*, *89*(7), 575.

Hirsch, C. R., Krahé, C., Whyte, J., Loizou, S., Bridge, L., Norton, S., & Mathews, A. (2018). Interpretation training to target repetitive negative thinking in generalized anxiety disorder and depression. *Journal of Consulting and Clinical Psychology, 86*(12), 1017.

Hirsch, C. R., & Mathews, A. (2000). Impaired positive inferential bias in social phobia. *Journal of Abnormal Psychology, 109*(4), 705–712. https://doi.org/10.1037/0021-843X.109.4.705

Hirsch, C. R., Meeten, F., Krahé, C., & Reeder, C. (2016). Resolving ambiguity in emotional disorders: The nature and role of interpretation biases. *Annual Review of Clinical Psychology, 12,* 281-305.

Howard, S., O'Riordan, A., & Nolan, M. (2018). Cognitive bias of interpretation in Type D personality: associations with physiological indices of arousal. *Applied Psychophysiology and Biofeedback*, *43*(3), 193-201.

Hyde, J., Ryan, K. M., & Waters, A. M. (2019). Psychophysiological markers of fear and anxiety. *Current Psychiatry Reports, 21*(7), 1-10.

Joormann, J., Waugh, C. E., & Gotlib, I. H. (2015). Cognitive bias modification for interpretation in major depression: Effects on memory and stress reactivity. *Clinical Psychological Science*, *3*(1), 126-139.

Kornblum, S., & Requin, J. (Eds.). (2019). *Preparatory states and processes*. Psychology Press.

Krahé, C., Whyte, J., Bridge, L., Loizou, S., & Hirsch, C. R. (2019). Are Different Forms of Repetitive Negative Thinking Associated With Interpretation Bias in Generalized Anxiety

Disorder and Depression? *Clinical Psychological Science*, *7*(5), 969–981. https://doi.org/10.1177/2167702619851808

Kreifelts, B., Brück, C., Ethofer, T., Ritter, J., Weigel, L., Erb, M., & Wildgruber, D. (2017). Prefrontal mediation of emotion regulation in social anxiety disorder during laughter perception. *Neuropsychologia, 96,* 175-183.

Kreifelts, B., Weigel, L., Ethofer, T., Brück, C., Erb, M., & Wildgruber, D. (2019). Cerebral resting state markers of biased perception in social anxiety. *Brain Structure and Function, 224*(2), 759-777.

Laborde, S., Allen, M. S., Borges, U., Hosang, T. J., Furley, P., Mosley, E., & Dosseville, F. (2021). The influence of slow-paced breathing on executive function. *Journal of Psychophysiology, 36*(1), 13–27.

Lang, P. J., McTeague, L. M., & Bradley, M. M. (2017). The psychophysiology of anxiety and mood disorders: The RDoC challenge. *Zeitschrift für Psychologie, 225*(3), 175–188. https://doi.org/10.1027/2151-2604/a000302

Lawson, C., MacLeod, C., & Hammond, G. (2002). Interpretation revealed in the blink of an eye: Depressive bias in the resolution of ambiguity. *Journal of Abnormal Psychology*, *111*(2), 321.

Lehrer, P., Kaur, K., Sharma, A., Shah, K., Huseby, R., Bhavsar, J., ... & Zhang, Y. (2020). Heart rate variability biofeedback improves emotional and physical health and performance: a

systematic review and meta analysis. *Applied Psychophysiology and Biofeedback*, *45*(3), 109-129.

Llera, S. J., & Newman, M. G. (2010). Effects of worry on physiological and subjective reactivity to emotional stimuli in generalized anxiety disorder and nonanxious control participants. *Emotion, 10*(5), 640.

Marwood, L., Wise, T., Perkins, A. M., & Cleare, A. J. (2018). Meta-analyses of the neural mechanisms and predictors of response to psychotherapy in depression and anxiety. *Neuroscience & Biobehavioral Reviews, 95*, 61-72.

Mather, M., & Thayer, J. (2018). How heart rate variability affects emotion regulation brain networks. *Current Opinion in Behavioral Sciences*, *19*, 98–104. https://doi.org/10.1016/j.cobeha.2017.12.017

Mathews, A., & MacLeod, C. (2005). Cognitive vulnerability to emotional disorders. *Annual. Review of Clinical. Psycholology, 1*, 167-195.

Mathews, A., Richards, A., & Eysenck, M. (1989). Interpretation of homophones related to threat in anxiety states. *Journal of Abnormal Psychology, 98*(1), 31-34.

Mathôt S. (2018). Pupillometry: Psychology, Physiology, and Function. *Journal of Cognition*, *1*(1), 16. https://doi.org/10.5334/joc.18

Meeten, F. (2017). *Modification of Interpretation Biases in Worry: An Examination of Cognitive and Physiological Responses.* Kings College London. Unpublished thesis.

Menne-Lothmann C, Viechtbauer W, Höhn P, Kasanova Z, Haller SP, Drukker M, et al. (2014) How to Boost Positive Interpretations? A Meta-Analysis of the Effectiveness of Cognitive Bias Modification for Interpretation. *PLoS ONE 9*(6): e100925. https://doi.org/10.1371/journal.pone.0100925

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine*, *6*(7), e1000097.

Moors, A. (2009). Theories of emotion causation: A review. *Cognition and Emotion, 23*(4), 625-662.

Moser, J. S., Hajcak, G., Huppert, J. D., Foa, E. B., & Simons, R. F. (2008). Interpretation bias in social anxiety as detected by event-related brain potentials. *Emotion*, *8*(5), 693.

Moser, J. S., Huppert, J. D., Foa, E. B., & Simons, R. F. (2012). Interpretation of ambiguous social scenarios in social phobia and depression: evidence from event-related brain potentials. *Biological Psychology*, *89*(2), 387-397.

Nowakowski, M. E., Antony, M. M., & Koerner, N. (2015). Modifying interpretation biases: Effects on symptomatology, behavior, and physiological reactivity in social anxiety. *Journal*

of Behavior Therapy and Experimental Psychiatry, *49*, 44-52.

Ottaviani, C., Thayer, J. F., Verkuil, B., Critchley, H. D., & Brosschot, J. F. (2017). Can't Get You Out of My Head: Brain-Body Interactions in Perseverative Cognition. *Frontiers in Human Neuroscience, 11*, 634.

Ottaviani, C., Thayer, J. F., Verkuil, B., Lonigro, A., Medea, B., Couyoumdjian, A., & Brosschot, J. F. (2016). Physiological concomitants of perseverative cognition: A systematic review and meta-analysis. *Psychological Bulletin, 142*(3), 231.

Poppelaars, E. S., Harrewijn, A., Westenberg, P. M., & van der Molen, M. J. (2018). Frontal delta-beta cross-frequency coupling in high and low social anxiety: An index of stress regulation?. *Cognitive, Affective, & Behavioral Neuroscience*, *18*(4), 764-777.

Qi, Y., Gu, R., Cao, J., Bi, X., Wu, H., & Liu, X. (2017). Response bias-related impairment of early subjective face discrimination in social anxiety disorders: an event-related potential study. *Journal of Anxiety Disorders, 47*, 10-20.

Rothstein, H. R., & Hopewell, S. (2009). Grey literature. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 103–125). Russell Sage Foundation.

Rozenman, M., Gonzalez, A., Logan, C., & Goger, P. (2020). Cognitive bias modification for threat interpretations: Impact on anxiety symptoms and stress reactivity. *Depression and Anxiety*, *37*(5), 438-448.

Schoth, D. E., & Liossi, C. (2017). A systematic review of experimental paradigms for exploring biased interpretation of ambiguous information with emotional and neutral associations. *Frontiers in Psychology, 8*, 171.

Schick et al. (2013

Sibrava, N. J., & Borkovec, T. D. (2006). The cognitive avoidance theory of worry. Worry and its psychological disorders: Theory, assessment and treatment, 1, 239-256.

Smith, R., Thayer, J. F., Khalsa, S. S., & Lane, R. D. (2017). The hierarchical basis of neurovisceral integration. *Neuroscience & Biobehavioral Reviews*, *75*, 274-296.

Sörnmo, L., & Laguna, P. (2005). *Bioelectrical signal processing in cardiac and neurological applications* (Vol. 8). Academic Press.

Standage, H., Harris, J., & Fox, E. (2014). The influence of social comparison on cognitive bias modification and emotional vulnerability. *Emotion*, *14*(1), 170.

Steinman, S. A., Portnow, S., Billingsley, A. L., Zhang, D., & Teachman, B. A. (2020). Threat and benign interpretation bias might not be a unidimensional construct. *Cognition and Emotion, 34*(4), 783-792.

Stuijfzand, S., Creswell, C., Field, A. P., Pearcey, S., & Dodd, H. (2018). Research Review: Is anxiety associated with negative interpretations of ambiguity in children and adolescents? A systematic review and meta-analysis. *Journal of Child Psychology and Psychiatry, 59*(11), 1127-1142.

Sur, S., & Sinha, V. K. (2009). Event-related potential: An overview. *Industrial Psychiatry Journal*, *18*(1), 70.

Teachman, B. A., Joormann, J., Steinman, S. A., & Gotlib, I. H. (2012). Automaticity in anxiety disorders and major depressive disorder. *Clinical psychology review*, *32*(6), 575–603. https://doi.org/10.1016/j.cpr.2012.06.004

Thayer, J. F., Hansen, A. L., Saus-Rose, E., & Johnsen, B. H. (2009). Heart rate variability, prefrontal neural function, and cognitive performance: the neurovisceral integration perspective on self-regulation, adaptation, and health. *Annals of Behavioral Medicine*, *37*(2), 141-153.

Tooley, M. D., Carmel, D., Chapman, A., & Grimshaw, G. M. (2017). Dissociating the physiological components of unconscious emotional responses. *Neuroscience of Consciousness*, *2017*(1), nix021.

Van Bockstaele, B., Clarke, P. J., Notebaert, L., MacLeod, C., & Salemink, E. (2020). Effects of cognitive load during interpretation bias modification on interpretation bias and stress reactivity. *Journal of Bbehavior Therapy and Experimental Psychiatry*, *68*, 101561.

Van den Bergh, O., Brosschot, J., Critchley, H., Thayer, J. F., & Ottaviani, C. (2021). Better safe than sorry: A common signature of general vulnerability for psychopathology. *Perspectives on Psychological Science*, *16*(2), 225-246.

van der Molen, M. J., Harrewijn, A., & Westenberg, P. M. (2018). Will they like me? Neural

and behavioral responses to social-evaluative peer feedback in socially and non-socially anxious females. *Biological Psychology, 135*, 18-28.

van der Ploeg, M. M., Brosschot, J. F., Versluis, A., & Verkuil, B. (2017). Peripheral physiological responses to subliminally presented negative affective stimuli: A systematic review. *Biological Psychology, 129*, 131-153.

Vanderhasselt, M. A., & Ottaviani, C. (2022, January 1). Combining top-down and bottom-up interventions targeting the vagus nerve to increase resilience. *Neuroscience and Biobehavioral Reviews*. Elsevier Ltd. https://doi.org/10.1016/j.neubiorev.2021.11.018.

Verkuil, B., Brosschot, J. F., Gebhardt, W. A., & Thayer, J. F. (2010). When worries make you sick: a review of perseverative cognition, the default stress response and somatic health. *Journal of Experimental Psychopathology*, *1*(1), jep-009110.

Whitton, A. E., Grisham, J. R., Henry, J. D., & Palada, H. D. (2013). Interpretive bias modification for disgust. *Journal of Experimental Psychopathology*, *4*(4), 341-359.

Yiend, J., Mackintosh, B., & Mathews, A. (2005). Enduring consequences of experimentally induced biases in interpretation. *Behaviour Research and Therapy, 43*(6), 779–797. https://doi.org/10.1016/j.brat.2004.06.007

Table 1: Study demographics with studies presented in alphabetical order

Study	***N* sample**	**Country**	**Gender**	***Mean* Age (*SD*)**
Beadel et al., (2013)	*N* = 75 Conditions: CBM-I positive: *N* = 36 CBM-I control: *N* = 39	USA	57 female: 18 male	*M* = 18 Conditions: CBM-I positive: *M* = 19 (1) CBM- control: *M* = 19 (1)
Cowden Hindash et al. (2021)	*N = 53* Participant groups: MDD: *N* = 25 Controls: *N* =28	USA	38 female: 15 male	Participant groups: MDD: *M* = 21 (7) Control: *M* = 21 (5)
Feng et al. (2019)	*N* = 55 Participant groups: High Worry: *N* = 28 Low Worry: *N* = 27	UK	33 female: 22 male	*M* = 27.69 (9.8) Participant groups: High worry: *M* = 26 (9) Low worry: *M* = 29 (10)
Feng et al. (2020) (Study 2)	*N* = 66 Conditions: CBM-I positive: *N* = 35 CBM-I control: *N* = 31	UK	30 female: 5 male	Conditions: CBM-I positive: *M* = 28(8) CBM-I control: *M* = 26(8)

Grisham et al. (2014)	*N* = 95 Conditions: CBM-I positive: *N* = 47 CBM-I negative: *N* = 48	Australia	71 female: 24 male	*M* = 20 (3)
Howard et al. (2018)	*N = 80*	Ireland	80 female: 0 male	*M = 21 (2)*
Joormann et al. (2015)	*N* = 76 (48 MDD, 28 healthy) Conditions: CBM-I positive: *N* = 14 (healthy) *N* = 16 (MDD) CBM-I negative: *N* = 14 (healthy) *N* = 15 (MDD) No CBM-I: *N* = 17 (MDD)	USA	26 female: 48 male	Conditions: CBM-I positive: *M* = 39 (11) (MDD) *M* = 41 (11) (healthy) CBM-I negative: *M* = 30 (12) (MDD) *M* = 42 (13) (healthy) No CBM-I: *M* = 34 (11) (MDD)
Lawson et al. (2002)	*N = 54* Participant groups: High depression: *N* = 27 Low depression: *N* = 27	Australia	40 female: 14 male	Participant groups: High depression: *M* = 19 *(*2) Low depression: *M* = 20 (3)
Meeten et al. (2017)*	*N* = 49 Conditions: CBM-I positive: *N* = 25 CBM-I control: *N* = 24	UK	42 female: 7 male	Conditions: CBM-I positive: *M* = 24 (7) CBM- control: *M* = 25 (5)

Moser et al., (2008)	*N = 34* Participant groups: High social anxiety: *N* = 16 Low social anxiety: *N* = 18	USA	21 female: 13 male Participant groups: High social anxiety: 10 female, 6 male Low social anxiety: 11 female, 7 male	 Participant groups: High social anxiety: *M* = 26 (7) Low social anxiety: *M* = 25 (7)
Moser et al. (2012)	*N = 63* Participant groups: Social anxiety: *N* = 17 MDD: *N* = 10 Social anxiety + MDD: *N* = 17 Controls: *N* = 19	USA	45 female: 18 male Participant groups: Social anxiety: 12 female: 5 male MDD: 9 female: 1 male Social anxiety + MDD: 10 female: 7 male Controls: 14 female: 5 male	 Participant groups: Social anxiety: *M* = 25 (5) MDD: *M* = 26 (6) Society anxiety + MDD: *M* = 24 (6) Controls: *M* = 24 (3)
Nowakowski et al., (2015)	*N* = 72 Conditions: CBM-I positive: *N* = 24 CBM-I control: *N* = 24 Cognitive restructuring: *N* = 24	USA	51 female: 21 male Conditions: CBM-I positive: 16 female: 8 male CBM-I control: 17 female: 7 male Cognitive restructuring: 18 female: 6 male	 Conditions: CBM-I positive: *M* = 24 (4) CBM-I control: *M* = 28 (10) Cognitive restructuring: *M* = 24 (6)
Rozenman et al., (2020)	*N = 24* Conditions: CBM-I: *N = 12* *CBM-I* Control: *N = 12*	USA	13 female: 11 male	*M* = 20 (2)

Van Bockstaele et al., (2020)	*N = 74* Conditions: CBM-I positive (cognitive load): *N* = 19 CBM-I positive: *N* = 16 CBM-I negative (cognitive load): *N* = 18 CBM-I negative: *N* = 18	Holland	50 female: 21 male	Conditions: CBM-I positive (cognitive load): *M* = 20 (1) CBM-I positive: *M* = 26 (12) CBM-I negative (cognitive load): *M* = 20 (2) CBM-I negative: *M* = 20 (3)
Whitton et al., (2013)	*N = 60* *Participant groups:* *CBM-I positive: N = 30* *CBM-I negative: N = 30*	*Australia*	*37 female: 23 male*	*M = 33 (17)*

General notation: IB = Interpretation Bias; CBM-I = Cognitive Bias Modification for Interpretations (positive, negative, control conditions)
Disorder abbreviations: PTSD = Post Traumatic Stress Disorder; MDD = Major Depressive Disorder; OCD = Obsessive Compulsive Disorder
*Asterisk: * = grey literature study*

Table 2: Experimental results

Study	***Physiological marker***	***IB measure(s)***	***Point in study that physiological measurement was taken***	***Emotional construct***	***CBM-I***	***Effect of IV on interpretation bias***	***Effect of IV on Physiological marker***	***Quality score (M)***
Measures taken at the time that interpretation occurs								
Cowden Hindash et al. (2021)	Pupilometry	- Word-sentence association paradigm for depression (adapted)	During IB assessment: - Word-sentence association paradigm	Major depression disorder: - BDI-II - STAI-T - STAI-S - PANAS (trait positive affect) - PANAS (trait negative affect)	✗	✓	✓	**16**
Feng et al. (2019)	ERP	- Lexical decision task - Scenario task - Recognition task - Sentence word association task	During IB assessment: - Lexical decision task - Sentence word association task	Pathological worry: - PSWQ - WDQ - GAD-7 - PHQ-9	✗	✓	✓	**20**
Feng et al. (2020) (Study 2)	ERP	- Recognition task - lexical decision task	During IB assessment: - Lexical decision task (post CBM-I)	Pathological worry: - PSWQ - GAD-7 - PHQ-9	✓	✓	✗	**19**

Howard et al. (2018)	HR & SCR	- (Social) scenario task	During IB assessment: - Scenario task	Type D personality (Negative affectivity, social inhibition): - DS14 (Denollet 2005)	✕	✓	✓ HR ✕ SCR	**14.5**
Lawson et al. (2002) (Experiment 2)	EMG	- Acoustic blends of words	During IB assessment: - Acoustic blends of words	Depressive traits - BDI-II - STAI-T	✕	✓	✓	**11**
Moser et al. (2008)	ERP	- Lexical decision task	During IB assessment: - Lexical decision task	Social anxiety: - SPIN - BFNE - LSAS, - SIAS - DASS	✕	✓	✓	**20**
Moser et al. (2012)	ERP	- Lexical decision task	During IB assessment: - Lexical decision task	Social anxiety: - SPIN - QIDS - LSAS - SIAS - BDI-II - DASS-D	✕	✓	✓	**20**
Measures taken pre and post interpretation bias training (near-transfer)								
Meeten et al. (2017)*	HRV	- Recognition task	During CBM-I	Pathological worry: - PSWQ - STAI-T	✓	✕	✓	**19**

		- Sentence completion task						
Measures taken after interpretation bias training during a subsequent stressor task (far-transfer)								
Beadel et al. (2013)	HR & SCR	- Recognition task	During CBM-I and post stressor task: - OC Trash task - OC Thought task (adapted from Magee and Teachman 2007 - OC Anagram task (adapted from Macleod et al., 2002)	OCD traits: - SUDS - STAI-T - DASS-7	✓	✓	✕	**21**
Grisham et al. (2014)	HR & HRV	- Recognition task	During post CBM-I stressor task: - Sweet-sorting task (adapted from Ladouceur et al., 1995; Reeves et al., 2010)	Trait & state anxiety (VAS based on Lothmann et al., 2011) OCD: - PI-R-CHCK - RAS	✓	✓	✓	**19**
Joormann et al. (2015)	HR	- Ambiguous Scenario task	During post CBM-I stressor task (Waugh et al., 2010)	Major depression disorder: - BDI-II - DSM-IV	✓	✓	✓	**16**

Nowakowski et al. (2015)	HR & SCR	-Recognition task -Ambiguous social situations interpretation questionnaire	During post CBM-I stressor task: - Perception of Speech Performance (Rapee & Lim, 1992)	Social anxiety: SPIN SCID-IV	✓	✓	✕	**11**
Rozenman et al. (2020)	HR, SCR, & RSA	- Word-sentence association task	During pre- and post- CBM-I stressor task: - Impromptu speech task (Beidel et al., 2013)	Trait anxiety: -STAI-T	✓	✓	✓ HR ✓HRV ✕ RSA	**16.5**
Van Bockstaele et al. (2020)	HR & HRV	- Recognition task	During post CBM-I stressor task: - Anagram stress task (Macleod et al., 2002)	Trait & state anxiety: -STAI-S -STAI-T	✓	✓	✕ HR ✓HRV	**17**
Whitton et al. (2013)	EMG	- Recognition task	During an imagery disgust eliciting task	Trait disgust DS-R OCI-R DASS-21	✓	✕	✕	**16**

General notation: IV = Independent variable IB = Interpretation Bias; CBM-I = Cognitive Bias Modification for Interpretations (positive, negative, control conditions); M (Quality Score) = mean score
Disorder abbreviations: PTSD = Post Traumatic Stress Disorder; MDD = Major Depressive Disorder; OCD = Obsessive Compulsive Disorder
Scale abbreviations: PSWQ = Penn State Worry Questionnaire; WDQ = Worry Domains Questionnaire; GAD-7 = Generalised Anxiety Disorder 7-item scale; PHQ-9 = Patient Health Questionnaire 9-item scale; DSM-IV = Diagnostic Statistical Manual for Mental Disorder; SPIN = Social Phobia Inventory; QIDS = Quick Inventory for Depression Symptoms; LSAS = Liebowitz Social Anxiety Scale; SIAS = Social Interaction Anxiety Scale; BDI-II = Beck Depression Inventory – 2nd version; DASS-D = Depression-Anxiety-Stress Scales – Depression subscale; SCID-IV = Structured Clinical Interview for DSM-IV; STAI-I/STAI-S = Trait-State Anxiety Inventory; OCI-R = Obsessive–Compulsive Inventory-Revised-Obsessions Subscale; RAS = Responsibility Attitudes Scale; ASSIQ = Ambiguous social situations interpretation questionnaire; DS-R = Disgust Scale Response; SUDS = Subjective Units of Distress Scale; PANAS = Positive & Negative Affective Schedule

*Asterisk: * = grey literature*