**ORIGINAL ARTICLE**

# Interpretation Bias Across Body Dysmorphic, Social Anxiety and Generalized Anxiety Disorder—A Multilevel, Diffusion Model Account

Fanny Alexandra Dietel[1] · Anne Möllmann[1] · Paul-Christian Bürkner[2] · Sabine Wilhelm[3] · Ulrike Buhlmann[1]

## Abstract

**Background** Interpretation biases are suggested to be transdiagnostic phenomena, but have rarely been compared across different disorders and current concerns.

**Methods** We investigated explicit, decision-based, and more implicit, reaction time-based interpretation bias in individuals with body dysmorphic disorder (BDD; $N = 29$), social anxiety disorder (SAD; $N = 36$), generalized anxiety disorder (GAD; $N = 22$), and non-clinical controls (NC; $N = 32$), using an adapted Word Sentence Association Paradigm (WSAP).

**Results** Results indicated that interpretation bias occurred transdiagnostically, while content-specific bias patterns varied meaningfully across groups. BDD and SAD shared explicit and, more inconsistently, implicit interpretation biases for appearance-related and social situations. The GAD group exhibited an explicit and implicit negative interpretation bias for general situations, and an additional implicit lack of positive bias. Mechanistic Wiener diffusion model analyses revealed that interpretation bias patterns were mainly driven by speeded information uptake, potentially mirroring disorder-specific associative memory organization.

**Conclusions** These findings have important implications for understanding interpretation biases as both etiological and treatment factors.

## Introduction

The tendency to misinterpret ambiguous situations, termed interpretation bias, is proposed to maintain a range of emotional disorders (e.g., Mathews and MacLeod 2005).

✉ Ulrike Buhlmann
ulrike.buhlmann@uni-muenster.de

1 Department of Psychology, Clinical Psychology and Psychotherapy, Westfälische Wilhelms-Universität, Fliednerstrasse 21, 48149 Münster, Germany

2 Cluster of Excellence Simtech, University of Stuttgart, Stuttgart, Germany

3 Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

Consistently, interpretation biases are considered to be transdiagnostic phenomena (Craske et al. 2009; Harvey et al. 2004; Mansell et al. 2008), emerging from aberrant information processing (see Gladwin and Figner 2014; Ouimet et al. 2009; Strack and Deutsch 2004) that is guided by underlying dysfunctional beliefs and current concerns. Such current concerns are linked to pursuing a goal (e.g., achieving attractiveness) and prompt the preferential processing of goal-relevant stimuli (e.g., others' reactions to one's appearance; Klinger and Cox 2011). Over time, current concerns may thereby shape associative memory networks through learning processes (Anderson 1983; Collins and Loftus 1975; Foa et al. 2006) and produce disorder-specific cognitive profiles—including interpretation biases—that elicit emotional states (e.g., Beck and Clark 1988; Beck and Haigh 2014).

To date, these disorder-specific cognitive profiles have mainly been investigated by contrasting self-reported cognitive content. A classic meta-analysis in this field yielded mixed results, demonstrating that anxiogenic and

depressogenic cognitions were related to symptoms of both disorders, with only the latter showing a specific association with depression (Beck and Perkins 2001). However, more recent research supported the assumption of cognitive profiles, indicating that some emotional disorders are linked to distinct underlying core beliefs (Cooper et al. 2006; Dozois et al. 2008; Schulz et al. 2008).

Considering interpretation biases, prior studies have mainly focused on their assessment in *single* disorders, e.g., depression (e.g., Hindash and Amir 2012; Wisco and Nolen-Hoeksema 2010), social anxiety disorder (SAD; e.g., BeVonard and Amir 2009; Hirsch and Clark 2004; Huppert et al. 2003; Voncken et al. 2003), generalized anxiety disorder (GAD; e.g., Hazlett-Stevens and Borkovec 2004), and eating disorders (e.g., Rosser et al. 2010). Conversely, few studies have systematically compared interpretation bias patterns (e.g., Buhlmann et al. 2002; McManus et al. 2000; Voncken et al. 2007). Exemplarily, Buhlmann et al. (2002) showed that individuals with body dysmorphic disorder (BDD) favored more threatening and less non-threatenineg interpretations for appearance-related (e.g., "While talking to some colleagues, you notice that some people take special notice of you.") and social scenarios (e.g., "You are having a conversation with some friends. You say something and the conversation stops.") than individuals with obsessive–compulsive disorder and mentally healthy controls. However, both clinical groups exhibited a negative interpretation bias for generally ambiguous scenarios (e.g., "A letter marked "URGENT" arrives."), potentially reflecting an overarching vulnerability factor for psychopathology (Buhlmann et al. 2002). Similarly, Voncken et al. (2007) demonstrated that individuals with SAD (vs. individuals with depression and mentally healthy controls) displayed a distinct negative interpretation bias for social scenarios. Further, individuals with depression, compared to non-clinical controls, exhibited a more global negative bias encompassing various situation categories.

In sum, the studies cited above indicate that maladaptive, concern-specific interpretation biases characterize a range of emotional disorders. Nevertheless, prior research has mainly relied on restrictive self-report formats, such as forced-choice, which allow unlimited time for evaluation and may be prone to confounds (e.g., response selection bias and demand effects; Mathews and MacLeod 2005). Further, such measures preclude a clear differentiation between tendencies to endorse negative and reject positive interpretations (i.e., pronounced negative bias and a lack of positive bias). However, these tendencies have been shown to represent two distinct, equally pertinent factors in bias phenomenology, being tied to different behavioral implications that warrant further investigation (Huppert et al. 2003; Steinman et al. 2020). Specifically, while negative interpretation bias has been linked to behavioral avoidance, a lack of positive bias has been suggested to dampen positive affect during behavioral approach (Amir et al. 2012; Kuckertz and Amir 2017).

Additionally, most cognitive theories posit that interpretation bias is characterized by a more implicit, reaction time-based (RT) component (Hirsch and Clark 2004), requiring assessment via opaque, time-limited task designs (e.g., Schoth and Liossi 2017). Within such implicit designs, RT is conceived to index the speed by which situational interpretations are accessible and activated within semantic memory (see de Houwer et al. 2009). Hence, when resolving ambiguity, RT theoretically quantifies the associative strength between a situation and its valent interpretation. However, it has not been studied yet whether observed RT differences indeed emerge from relatively faster information uptake, or confounds, such as low response thresholds or general response slowing (Voss et al. 2015). Overall, examining the aforementioned bias indices and dissecting their underlying cognitive components within associative memory organization appears critical to further characterize interpretation biases phenomenologically.

The Word Sentence Association Paradigm (WSAP; Beard and Amir 2009) was designed to yield different explicit and implicit interpretation bias indices. In this task, participants are asked to judge as fast as possible if an ambiguous sentence and a positive or negative interpretation are related. Importantly, both decision rates (i.e., the explicit component) and congruent RT (i.e., the more implicit component) may be recorded. Using the WSAP, Beard and Amir (2009) demonstrated that participants high (vs. low) in social anxiety exhibited a pronounced negative and a lack of positive interpretation bias. Further, individuals high (vs. low) in social anxiety also showed this bias implicitly, as they were slower in endorsing positive and rejecting negative, and faster in endorsing negative and rejecting positive interpretations. Since this study, interpretation biases have been investigated using the WSAP in various emotional disorders (see Gonsalves et al. 2019, for a review). However, to our knowledge, the WSAP has never been used to compare interpretation biases across disorders and current concerns. Such comparative assessments are paramount to determining common and disorder-specific bias features that may be addressed within interventions, such as Cognitive Bias Modification for Interpretation (CBM-I; see Cristea et al. 2015; Hallion and Ruscio 2011; Jones and Sharpe 2017; Menne-Lothmann et al. 2014, for meta-analyses).

The present study investigated interpretation biases for different current concerns across three disorders: BDD, SAD, and GAD. The key characteristics of these disorders involve preoccupation about subjectively perceived bodily flaws in BDD, fear and avoidance of social situations in SAD, as well as anxiety and worry about various domains in GAD (American Psychiatric Association 2013). BDD, SAD and GAD can be represented on a continuum

of phenomenological proximity. In this respect, BDD and SAD can be considered phenomenologically *proximal*. Both disorders are characterized by social anxiety and avoidance, similar onset and trajectories, and high mutual comorbidity (Fang and Hofmann 2010; Pinto and Phillips 2005). BDD and SAD might further relate on a cognitive level as they potentially share maladaptive social and appearance-related interpretation patterns, given their overlap in anxiety and appearance-related concerns during social situations (Fang and Hofmann 2010). Indeed, cognitive-behavioral models of BDD and SAD propose that interpretation biases maintain symptoms (e.g., Hofmann 2007; Wilhelm et al. 2013). Relatedly, SAD-specific cognitive-behavioral interventions have been found to improve BDD symptoms (Fang et al. 2013), suggesting common underlying factors (e.g., fear of negative evaluation or self-focused attention; Fang and Hofmann 2010). Thus, investigating interpretation bias profiles in BDD and SAD would further elucidate their role in symptom maintenance, which remains unstudied at present.

GAD can be viewed as phenomenologically *distal* to BDD and SAD. Despite some commonalities, such as trait anxiety, excessive worry, avoidance, and safety behaviors to reduce anxiety (American Psychiatric Association 2013; Craske et al. 2009; Turk et al. 2005), GAD appears distinct on a cognitive level as it is associated with various current concerns (i.e., different worry domains, primarily concerning a potentially aimless future, relationships, work incompetence, and physical threat; Dugas et al. 1998). The differences between GAD, BDD and SAD are further reflected in their low mutual comorbidity rates and diverging age of onset (e.g., Gunstad and Phillips 2003). Moreover, cognitive-behavioral models of GAD identify intolerance of ambiguity as a catalyst for habitual negative interpretations and worry (see Hirsch et al. 2016, for an overview). However, it remains unclear how individuals with GAD respond to concerns present in other disorders.

Addressing these questions, we explored interpretation bias patterns in individuals fulfilling self-report DSM-5 criteria for BDD, SAD, GAD, and non-clinical controls (NC). Using an adapted version of the WSAP (Hindash and Amir 2012), we assessed decision rates and RT for positive and negative interpretations in three categories central to these disorders: appearance-related, social, and generally threatening situations. We used a multilevel approach (MLM) and Wiener diffusion models (Ratcliff and McKoon 2008; Voss et al. 2013) to determine cognitive parameters within RT. In reference to classic cognitive models, we explored the contributive value of decisional processes in implicit interpretation bias, which may be interpreted as indices of associative memory underlying interpretations (McKoon and Ratcliff 2012; White et al. 2010).

Consistent with prior evidence, we hypothesized that (1) clinical groups would overall exhibit maladaptive interpretation biases across situation categories, endorsing more negative and fewer positive interpretations than NC (transdiagnostic hypothesis). Within clinical groups, we predicted that (2) individuals with BDD would exhibit an appearance-related interpretation bias (vs. SAD and GAD) and a social interpretation bias (vs. GAD). Further, we expected that (3) SAD (vs. GAD) would show an appearance-related and social interpretation bias, and last, (4) individuals with GAD would show a general interpretation bias compared to NC only (current concern hypotheses). We further assumed that these disorder-specific patterns would be reflected in concurrent RT, i.e., faster endorsement and slower rejection of negative interpretations, and the reverse pattern for positive interpretations, as compared to NC.

## Methods

### Participants

Participants for clinical groups were recruited via online advertisements in disorder-specific Internet fora and social networks. Non-psychological fora and social networks were used for the recruitment of controls. Participation was not reimbursed. However, participants could enter a lottery comprising six Amazon vouchers worth 150€ in total.

Inclusion criteria were: (1) aged between 18 and 65 years, (2) no previous or current self-reported diagnosis of psychotic disorder, bipolar disorder or substance abuse/dependency, (3) no acute suicidality [as indicated by Patient Health Questionnaire, Depression Module (PHQ-9; Gräfe et al. 2004)], item 9 ≤ 3), and (4) fluent in German. Clinical groups had to fulfill inclusion criteria and the self-report DSM-5 criteria of BDD, SAD, or GAD (excluding mutual comorbidity). NC had to meet inclusion criteria with scores below the clinical cut-offs of 14 on the Body Dysmorphic Symptoms Inventory ('Fragebogen körperdysmorpher Symptome,' FKS; Buhlmann et al. 2009), 5.7 on the Generalized Anxiety Disorder Questionnaire-IV (GAD-Q-IV; Newman et al. 2002), 19 on the Social Phobia Inventory (SPIN; Sosic et al. 2008), 11 on the PHQ-9 (Gräfe et al. 2004), while not fulfilling clinical diagnoses of BDD, SAD or GAD.[1]

One thousand one hundred sixty-two individuals accessed the study's landing page, and 600 participants (51.64%) started the experiment. Some data sets had to be discarded due to: missing data related to participant dropout or technical difficulties ($N = 161$), non-fulfillment of inclusion

---

[1] .DSM-5 self-report items and self-conceptualized items used in this study (e.g., SWAP stimuli, state affect scales) can be obtained from the Online Appendices in the OSF data repository.

criterion 1 ($N = 2$), and inclusion criterion 3 ($N = 43$). We also excluded comorbid and subclinical cases as per the aforementioned criteria ($N = 275$). The final sample ($N = 119$) comprised the following groups: BDD ($N = 29$), SAD ($N = 36$), GAD ($N = 22$), and NC ($N = 32$).

## Measures and Materials

Unless otherwise specified, higher scores on the following scales indicate higher symptom severity.

### Diagnostic Criteria (DSM-5)

We rephrased the DSM-5 criteria of BDD, SAD and GAD (American Psychiatric Association 2013) into questions, using a dichotomous response format ("Yes"/"No"). Some criteria were broken down into two questions to improve comprehensibility. Accordingly, we administered six questions for BDD, ten questions for SAD, and eight questions for GAD. Internal consistencies for these self-report items were acceptable to high (BDD: *KR-20* = 0.76; SAD: *KR-20* = 0.96; GAD: *KR-20* = 0.94).

### Body Dysmorphic Symptoms Inventory (Fragebogen körperdysmorpher Symptome; FKS)

To measure BDD symptom severity, we administered the FKS (Buhlmann et al. 2009). Sum scores on this scale range between 0 and 64. Internal consistency in this sample was high (Cronbach's α = 0.89).

### Social Phobia Inventory (SPIN)

We used the SPIN (Connor et al. 2000; Sosic et al. 2008) to assess SAD symptom severity. This scale yields sum scores between 0 and 68. Internal consistency in the present sample was high (α = 0.90).

### Generalized Anxiety Disorder Questionnaire-IV (GAD-Q-IV)

We administered the GAD-Q-IV (Newman et al. 2002) to assess GAD symptom severity. Sum scores on this scale range between 0 and 12. Internal consistencies for the dichotomous (*KR-20* = 0.78) and interval items (α = 0.86) were high in this sample.

### Patient Health Questionnaire, Depression Module (PHQ-9)

To assess depression and suicidality, we administered the PHQ-9 (Gräfe et al. 2004; Kroenke and Spitzer 2002). Sum scores of this scale range between 0 and 27. Internal consistency was high in the present sample (α = 0.92).

### Spielberger Trait Anxiety Inventory (STAI-T, Form Y)

We assessed trait anxiety as a possible covariate of interpretation bias using the trait version of the STAI-T, Form Y (Spielberger et al. 1983). Sum scores on this inventory range between 20 and 80. Internal consistency in the present sample was high (α = 0.92).

### State Scales

To rule out unintended affect changes induced by the SWAP, which could influence subsequent symptom measures, we assessed state distress, self-esteem, and body dissatisfaction before and after the SWAP.

**Affect Scales** We assessed state anxiety, shame, sadness, disgust, and frustration each on single-item, ten cm-visual analog scales ranging from 0 ("not at all") to 100 ("extremely"). Affect scales were administered before and after the SWAP. Single-item scores were summed per time point (see Dietel et al. 2018).

**State Body Dissatisfaction and Self-Esteem (BISS and RSES-S)** We administered the 6-item Body Image State Scale (BISS; Cash et al. 2002) and 4-item Rosenberg Self-Esteem Scale, State Version (RSES-S; e.g., Nezlek and Kuppens 2008). Higher scores on the RSES-S reflect higher state self-esteem. Internal consistencies were high in the present sample at pre-assessment (α = 0.83 for the affect scales, α = 0.85 for the BISS, and α = 0.90 for the RSES-S).

**Sentence Word Association Paradigm (SWAP)** To assess interpretation bias, we used a version of the WSAP (Hindash and Amir 2012) presenting the ambiguous sentence before the valent word (named Sentence Word Association Paradigm, SWAP; see Dietel et al. 2018).

The situation set consisted of 240 sentence-word-combinations extracted from a pre-validated stimulus pool (see Appendix for examples). This stimulus pool was compiled from pre-existing stimulus sets of interpretation bias assessment studies in BDD (Buhlmann et al. 2002), SAD (Beard et al. 2011; Beard and Amir 2008) and GAD (Ogniewicz et al. 2014). Stimulus sets were kindly provided by the authors and translated. Further stimuli were generated based on expert consensus. Stimulus pools for appearance-related, social and general situations were then pre-validated in separate online surveys.[2] Negative and positive interpretations in

---

[2] Unpublished validation data of stimulus pools in online surveys (participants for appearance-related situation set: $N = 544$, social: $N = 427$, general: $N = 415$) indicated that higher symptom scores on respective measures [FKS, SPIN, Penn State Worry Questionnaire (PSWQ)] were positively correlated with endorsed negative interpretations ($r = .69$, $r = .69$, $r = .51$). The high-symptom groups (FKS > 33, SPIN > 24, PSWQ > 51) rated situations to be significantly

the present situation sets did not differ in word length for all situation categories (all $ps > 0.17$).

Participants initially saw ten practice trials, followed by 3 (appearance-related, social, general threat) × 80 sentence-word combinations. Each sentence was presented once. 50% of all trials contained a positive word. Participants thus received 240 trials in two blocks of 120 trials, separated by a 45 s-pause with an on-screen countdown to enhance concentration. Participants could skip this pause.

Trials began with a central black fixation cross displayed against a white background for 500 ms. An ambiguous sentence then appeared in the center of the screen for 3500 ms. It was subsequently replaced by a positive or negative word. Participants were requested to indicate as fast as possible whether the sentence and word were related (i.e., pressing "L" for "Yes"; "S" for "No"). The next trial was initiated after each decision. Decisions and RT were recorded. Internal consistencies per situation category were acceptable to excellent for decision rates ($\alpha = 0.83$–$0.92$) and RT ($\alpha = 0.70$–$0.95$).

## Procedure

The experiment was written in Inquisit Web Version 4 (Millisecond Software 2015) and accessible via desktop browsers. Upon informed consent, participants received screening questions for inclusion criteria and the PHQ-9. In case of a study exclusion, participants were directed to a webpage providing contact information of the principal investigator and additional mental health services. Upon meeting the inclusion criteria, participants completed the state scales (affect scales, BISS, RSES-S), SWAP, and state scales once again. Participants then received the STAI-T, FKS, DSM-5 BDD criteria, SPIN, DSM-5 SAD criteria, GAD-Q-IV, and DSM-5 GAD criteria. Upon completion, participants were redirected to a separate webpage to enter their e-mail address for the lottery. The experiment lasted 36.27 min on average.

## Design and Statistical Analyses

Data analyses were conducted using (1) SPSS Statistics Version 24 and (2) R (R Core Team 2018) with the RStudio interface (RStudio Team 2018). SPSS was used to initially investigate between-group demographic and psychometric

---

Footnote 2 (continued)

less pleasant (all $ps < .001$, Cohen's $ds = -0.45$ to $-1.04$) and more arousing (all $ps < .001$, $ds = 0.68$–$1.30$). Situations in stimulus pools were judged as at least moderately vivid ($M = 5.26$–$5.59$, $SD = 0.78$–$0.83$).

---

differences via analyses of variance (ANOVA), Welch's $t$-tests and $\chi^2$ tests, as well as frequentist analyses. Non-parametric alternatives were employed when basic assumptions of ANOVA were violated. R was used for Bayesian MLM-based analyses, treating individual responses as nested within both participants and items assuming a maximal multilevel structure (see OSF for the exact specification of the multilevel structure).

### Bayesian Multilevel Models (MLM) and Wiener Diffusion Models

We employed MLM as they offer greater flexibility and test power when treating interdependent, incomplete data sets that do not meet sphericity assumptions (e.g., Hoffman and Rovine 2007; Quené and Van den Bergh 2004). Inter alia, this approach is advantageous regarding the previously discussed problem of interdependency in SWAP-based RT data (see Dietel et al. 2018; Möbius et al. 2015).

Within this framework, we assumed the binary responses to be Bernoulli distributed (i.e., applying multilevel logistic regression) and the corresponding response times to be distributed according to an exponentially modified Gaussian distribution, which is a common choice for modeling RT in cognitive tasks (e.g., Balota and Yap 2011; Heathcote et al. 1991). We applied several tidyverse packages (Wickham 2017) for data preparation and plotting as well as the *brms* package (Bürkner 2017, 2018), which is based on the probabilistic programming language Stan (Carpenter et al. 2017). Overall, we employed a joint MLM of both binary decisions and RT, including a supplementary analysis via Wiener diffusion modeling to disentangle cognitive components underlying group-wise RT differences (e.g., Link and Heath 1975; Ratcliff 1978; Ratcliff and McKoon 2008; Wagenmakers 2009).

Wiener diffusion models allow for an in-depth analysis of information contained within RT distribution, i.e., information processing components (Wagenmakers 2009). These models conceptualize stimulus processing as a noisy accumulation of evidence over time, where one of the responses is given when the evidence reaches a set boundary (see Fig. 1). This accumulation may be represented within a bivariate probability density function, incorporating both RT and response options (i.e., decision rates) comprehensively (Vandekerckhove et al. 2011). The probability density function ultimately yields four parameters that may be mapped onto cognitive processes: drift rate (i.e., speed of information uptake), boundary separation (i.e., amount of information considered before a decision), starting point or initial bias (i.e., a priori biases in decision thresholds), and non-decision time (i.e., encoding and execution processes unrelated to decisions; Voss et al. 2013). Importantly, diffusion models assume parameters to vary from trial to
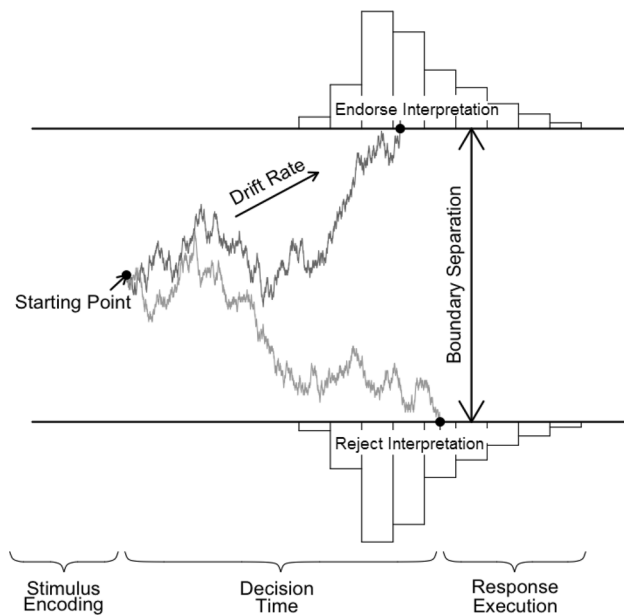
**Fig. 1** Depiction of the Wiener diffusion model and its parameters. Figure adapted from van Ravenzwaaij et al. (2012). Optimal decision making in neural inhibition models. Psychological Review, 119, 201–215. Copyright 2012 by American Psychological Association. Adapted with permission

trial, allowing for trial-based analyses and thereby avoiding information loss. They have been previously applied in the mechanistic analysis of binary choice tasks, e.g., the Implicit Association Test (Klauer et al. 2007; van Ravenzwaaij et al. 2011), and other cognitive bias assessments, e.g., the dot-probe paradigm (Price et al. 2019).

In this analysis, we predicted all four parameters of the Wiener diffusion model using a maximal multilevel structure. For drift rate, non-decision time, and boundary separation, the intercept, as well as the effects of item type and valence, were modeled as varying across participants. The intercept was also set as varying across items. Since the initial bias is, by definition, not affected by item properties, we only assumed a varying intercept across participants. All varying effects were allowed to correlate across participants and items. We included RT between 100 and 5000 ms (Luce 1986; Whelan 2008).

### Frequentist Analyses

To improve consistency with prior studies, we performed mixed repeated-measures analysis of variance (ANOVAs) on endorsement rates, RT, and state scales (Dietel et al. 2018; Möbius et al. 2015), using two-tailed testing at an alpha-level of 0.05. For the affect induction analysis, log-transformed state measures were submitted to a 3 (group: BDD, SAD, GAD, NC) way ANCOVA on post-assessment state measures, setting pre-assessment measures as a covariate

(Dietel et al. 2018). Percentual endorsement rates were submitted to a 4 (group: BDD, SAD, GAD, NC) × 3 (category: appearance-related, social, generally ambiguous) × 2 (valence: positive, negative) mixed ANOVA. We explored critical interactions using mixed ANOVAs and multiple comparison posthoc tests. Given sample size imbalance and heterogeneous variance, we report Games-Howell posthoc test outcomes (see Jaccard et al. 1984). For RT analysis, we eliminated the first and last percentile of observed RT, i.e., below 468.61 ms and above 5108.56 ms (Dietel et al. 2018; Möbius et al. 2015). Medians were subjected to a 4 (group: BDD, SAD, GAD, NC) × 2 (valence: positive, negative) × 2 (decision: endorse, reject) mixed repeated measures ANOVA per situation category. We followed up significant interactions using Bonferroni-corrected *t*-tests.[3]

### A Priori Power Analysis

The a priori power analysis using G*Power (Faul et al. 2007) was based on a WSAP-based interpretation bias assessment study with a similar setup, comparing mentally healthy controls and individuals with clinical SAD (Amir et al. 2012). This study reported large between-group effect sizes for negative (Cohen's $d = 1.23$) and positive interpretations ($d = 1.19$). Assuming this large between-group (i.e., mentally healthy vs. clinical individuals) effect size for positive and negative interpretations (analysis parameters: $1 - \beta = 0.95$, $\alpha = 0.05$) yielded a sample size of $n = 20$ per group (hence: $N = 80$ in total) to observe these effects. Given the lack of WSAP-based studies comparing interpretation patterns between clinical groups, effect sizes for these comparisons could not be estimated a priori.

## Results

Graphical and numerical summaries (including 95% credibility intervals of the effects) illustrating estimates for all MLM-based analyses can be found in the OSF repository.

### Baseline Measures

Groups did not differ in gender, age, and years of education (all $ps > 0.08$). Further, there was a significant difference in psychotherapy status, $\chi^2(3) = 11.36$, $p < 0.001$, $V = 0.31$. On a numerical level, clinical groups received psychotherapy more frequently than non-clinical controls, with a significant difference in psychotherapy status between GAD and NC. As expected, clinical groups additionally differed from

---

[3] The model codes and supplementary results of all analyses are available via Online Appendix A in the OSF repository.

**Table 1** Demographics and questionnaire measures for all groups

|  | BDD ($n=29$) | SAD ($n=36$) | GAD ($n=22$) | NC ($n=32$) |
|---|---|---|---|---|
|  | N (%) | N (%) | N (%) | N (%) |
| Gender (% female) | 25 (86.2)[a] | 23 (63.9)[a] | 13 (59.1)[a] | 21 (65.6)[a] |
| Medication (%) | 4 (13.8)[a] | 4 (11.1)[a] | 2 (9.1)[a] | 0 (0)[a] |
| Psychotherapy (%) | 4 (13.8)[a] | 5 (13.9)[a] | 7 (31.8)[b] | 0 (0)[a] |
|  | M (SD) | M (SD) | M (SD) | M (SD) |
| Age (years) | 24.31 (4.24)[a] | 29.06 (11.18)[a] | 27.27 (7.13)[a] | 30.69 (13.02)[a] |
| Education (years) | 15.93 (3.61)[a] | 16.83 (3.61)[a] | 16.32 (3.52)[a] | 15.97 (3.14)[a] |
| PHQ-9 | 8.21 (3.63)[a] | 6.75 (3.50)[a] | 11.09 (5.43)[b] | 4.63 (2.24)[c] |
| STAI-T | 45.93 (9.52)[a] | 44.47 (9.42)[a] | 51.73 (9.93)[a] | 32.50 (5.75)[b] |
| SPIN | 19.52 (8.77)[a] | 24.56 (10.78)[b] | 17.59 (10.25)[a] | 9.44 (4.70)[c] |
| FKS | 23.72 (8.89)[a] | 11.22 (6.13)[b] | 9.50 (6.27)[b] | 5.91 (4.02)[c] |
| GAD-Q-IV | 4.55 (2.60)[a] | 4.36 (2.25)[a] | 8.02 (2.99)[b] | 2.38 (1.42)[c] |

*BDD* body dysmorphic disorder, *SAD* social anxiety disorder, *GAD* generalized anxiety disorder, *NC* non-clinical controls, *PHQ-9* Patient Health Questionnaire, Depression Module, *STAI-T* Spielberger State-Trait Anxiety Inventory, Trait Version, Form Y, *FKS* Body Dysmorphic Symptoms Inventory, *GAD-Q-IV* Generalized Anxiety Disorder Questionnaire-IV

Means sharing superscripts ([a, b, c]) do not differ ($p > .05$, on Bonferroni-corrected $t$-tests or non-parametric tests, where applicable)

**Table 2** Mean state ratings (SD) for all assessment points

|  | Pre-assessment | | | | Post-assessment | | | |
|---|---|---|---|---|---|---|---|---|
|  | BDD ($n=29$) | SAD ($n=36$) | GAD ($n=22$) | NC ($n=32$) | BDD ($n=29$) | SAD ($n=36$) | GAD ($n=22$) | NC ($n=32$) |
| State distress (sum scores) | 110.59[a] (94.56) | 75.78[a] (74.08) | 138.82[a] (102.00) | 48.06[b] (51.11) | 113.28[a] (105.99) | 72.47[a] (80.65) | 116.27[a] (109.43) | 38.28[b] (56.05) |
| RSES-S (sum scores) | 25.83[a] (6.90) | 27.44[a] (8.49) | 24.86[a] (8.63) | 33.25[b] (4.21) | 25.72[a] (6.71) | 27.94[a] (8.21) | 25.59[a] (9.82) | 33.97[b] (4.51) |
| BISS (mean scores) | 5.37[a] (0.66) | 4.81[b] (0.72) | 4.89[b] (0.66) | 4.45[b] (0.64) | 5.21[a] (0.80) | 4.82[b] (0.71) | 4.78[b] (0.84) | 4.39[b] (0.79) |

*BDD* body dysmorphic disorder, *SAD* social anxiety disorder, *GAD* generalized anxiety disorder, *NC* non-clinical controls, *RSES-S* Rosenberg Self-Esteem Scale, State Version, *BISS* Body Image State Scale

Means sharing superscripts ([a, b]) do not differ within time point ($p > .05$, on Bonferroni-corrected $t$-tests or non-parametric tests, where applicable)

non-clinical controls regarding scores on the PHQ-9, STAI-T, SPIN, FKS, GAD-Q-IV, affect scales, BISS, and RSES-S (see Tables 1, 2).

## Unintended Affect Induction

We conducted multilevel linear regression models on the mean BISS, sum distress, and sum RSES state scales to investigate unintended SWAP-based affect induction. Overall, there were no substantial pre-post differences across all groups and state scales, indicating no unintended change in state distress, body dissatisfaction, or self-esteem (see Table 2).[4]

---

[4] The pattern of results using an ANCOVA on post-state scale values, setting pre-state scale values as a covariate (Dietel et al. 2018; Premo et al. 2016), did not differ (see Online Appendix B in the OSF data repository).

## Decision Rates

### Appearance-related Situations

As shown in Fig. 2, for positive appearance-related interpretations, the BDD and SAD group showed substantially lower endorsement rates (vs. GAD and NC). Conversely, for negative appearance-related interpretations, the BDD group exhibited substantially higher endorsement rates than all other groups. Similarly, the SAD and GAD group (vs. NC) demonstrated substantially higher endorsement rates for this category. All other differences remained non-substantial within a 95% credibility interval.

**Fig. 2** Mean endorsement rates (%) per situation category. Error bars represent 95% credibility intervals of the mean. *BDD* body dysmorphic disorder, *SAD* social anxiety disorder, *GAD* generalized anxiety disorder, *NC* non-clinical controls
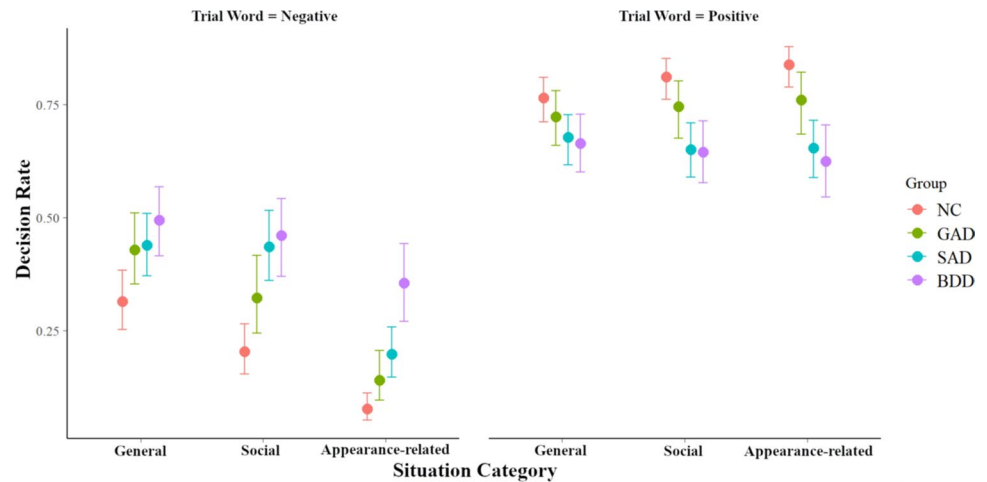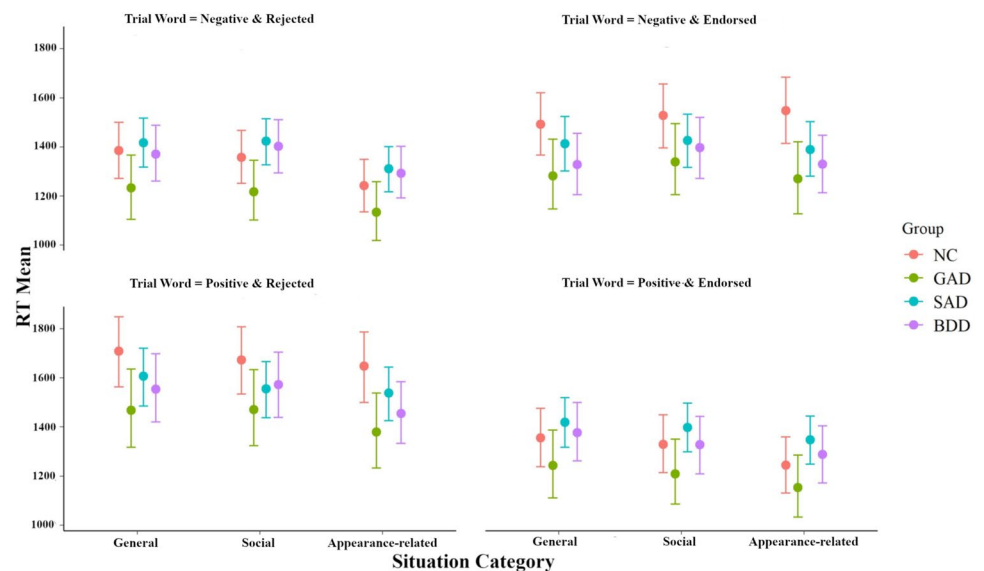
**Fig. 3** Mean reaction times (in ms) per situation category and decision. Error bars represent 95% credibility intervals of the mean. *BDD* body dysmorphic disorder, *SAD* social anxiety disorder, *GAD* generalized anxiety disorder, *NC* non-clinical controls

### Social Situations

For positive social interpretations, endorsement rates were substantially lower in the BDD and SAD group, compared to GAD and NC. For negative social interpretations, both the BDD and SAD group showed substantially higher endorsement rates than all other groups. Further, the GAD group endorsed more negative social interpretations than the NC group. All other differences were non-substantial within a 95% credibility interval.

### General Situations

For positive general situations, endorsement rates were substantially lower in the BDD and SAD as compared to the NC group. For negative general interpretations, the BDD, SAD, and GAD group exhibited substantially higher endorsement

rates than NC. All other differences remained non-substantial within a 95% credibility interval.[5]

### Reaction Times: Multilevel Analysis

#### Appearance-related Situations

As evident in Fig. 3, the BDD and SAD group endorsed positive appearance-related interpretations substantially slower than the GAD group. Conversely, the GAD and the

---

[5] The pattern of results using mixed ANOVA was overall similar. However, comparably small effects (e.g., negative interpretation endorsements in GAD vs. NC) did not reach significance within these analyses. The results should be interpreted within the limitations of frequentist analyses in small samples demonstrating between-group sample size imbalance (see Online Appendix B on the OSF data repository).
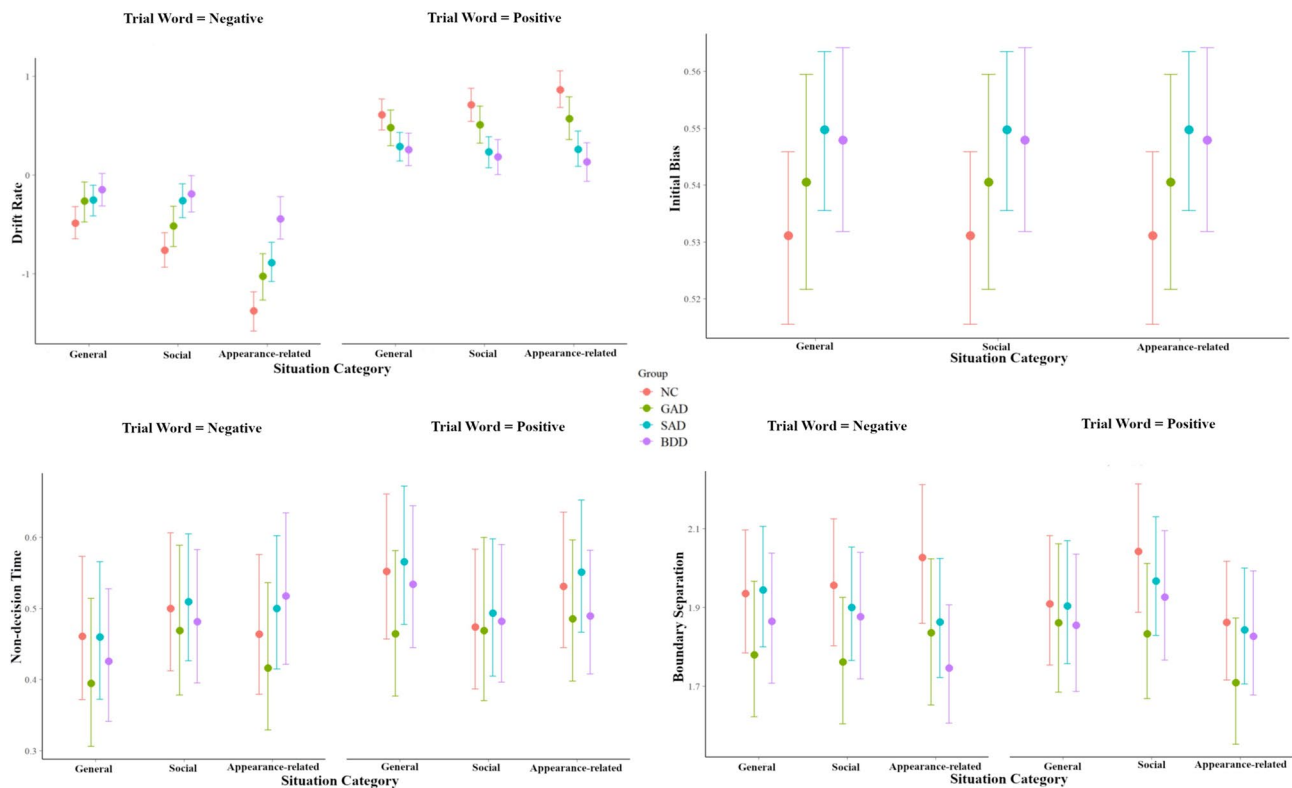
**Fig. 4** SWAP-effects across participants on drift rate, initial bias, non-decision time and boundary separation. Error bars represent 95% credibility intervals of the mean. *BDD* body dysmorphic disorder, *SAD* social anxiety disorder, *GAD* generalized anxiety disorder, *NC* non-clinical controls

BDD group showed substantially faster rejection of positive appearance-related interpretations than the NC. For negative appearance-related interpretations, the BDD, SAD and GAD group demonstrated faster endorsement (vs. NC), while the BDD and SAD group exhibited slower rejection RT (vs. GAD). All other differences remained non-substantial within a 95% credibility interval.

### Social Situations

For positive social interpretations, the SAD group showed substantially slower endorsements than the GAD group. Upon rejecting negative social interpretations, the BDD and SAD group were substantially slower than the GAD group. All other differences remained non-substantial within a 95% credibility interval.

### General Situations

The GAD group rejected positive general interpretations substantially faster than the NC group. Conversely, the GAD and the BDD group endorsed negative general interpretations substantially faster than the NC group. All other

differences remained non-substantial within a 95% credibility interval.[6]

## Reaction Times: Wiener Diffusion Models

Figure 4 displays mean SWAP effects on drift rate, boundary separation, initial bias, and non-decision time, including 95% credibility intervals of the effects.

### Drift Rates

As shown in Fig. 4 (upper left panel), there were non-zero SWAP effects for drift rates across all categories and groups. More positive values indicated relatively faster endorsement, and more negative values reflected faster rejection of the

---

[6] Within the aforementioned limitation of interdependent data, we only detected group-wise RT differences for the rejection of negative interpretations (appearance-related and social situations) and the endorsement of negative interpretations (general situations) using mixed ANOVAs. All results remained unchanged when setting age and gender as covariates to investigate potential influences of group-wise demographic differences (see Online Appendix B in the OSF data repository).

valent word. For negative appearance-related interpretations, drift rate values were overall negative (i.e., reflecting faster rejection), with the BDD group showing less negative values than all other groups. Further, the SAD and GAD group exhibited less negative values than the NC group in this situation category. For positive appearance-related interpretations, drift rates were overall positive (i.e., reflecting faster endorsement), with the BDD and SAD group showing less positive values than the other groups.

Similarly, for negative social interpretations, drift rates for the BDD and SAD group were less negative than GAD and NC drift rates. Again, for positive social interpretations, the BDD and SAD group exhibited less positive values than GAD and NC. The latter pattern was identically evident for positive general interpretations. All other between-group and between-category differences were non-substantial.

### Non-decision Time, Boundary Separation and Initial Bias

Figure 4 shows minimal SWAP effects across categories and groups, with no substantial between-group or between-category differences for all other Wiener diffusion model indices (i.e., non-decision time, boundary separation, and initial bias). Individuals thus did not differ across groups and situation categories with regard to non-decisional components, such as response caution or general response slowing (Voss et al. 2013).

## Discussion

This study investigated the phenomenology of interpretation bias across different clinical disorders (i.e., BDD, SAD, GAD, vs. NC) and current concerns. Using the SWAP paradigm, we examined explicit and more implicit, RT-based bias components for positive and negative interpretations. Further, we proposed a multilevel, diffusion model-based approach in analyzing SWAP-based RT bias indices to examine the relative contribution of underlying cognitive processes.

Explicit interpretation bias (i.e., based on decision rates) was present in BDD, SAD, and GAD, as consistent with the transdiagnostic hypothesis. However, bias patterns were shaped by content-specific differences between these groups, in line with the current concern hypothesis. As expected, both individuals with BDD and SAD, compared to GAD and NC, showed diminished positive appearance-related interpretation bias. However, BDD participants displayed a pronounced negative appearance-related interpretation bias, relative to all other groups. Similarly, both the BDD and SAD groups, vs. GAD and NC, exhibited reduced positive and enhanced negative social interpretation bias patterns. Only the BDD group endorsed fewer positive general

interpretations than the GAD and NC group, while all clinical groups demonstrated a more pronounced negative general interpretation bias than NC. These results are largely in line with the findings of Buhlmann et al. (2002), showing that BDD is associated with reduced positive and enhanced negative interpretation bias for appearance-related, social, and general situations. Importantly, the similarities of interpretation profiles further support the postulate that BDD and SAD are cognitively proximal (Fang and Hofmann 2010; Fang and Wilhelm 2015). Future research might aim to identify contributive factors driving these commonalities, such as fear of negative evaluation (Fang and Hofmann 2010). Examining dimensional relationships between these factors, bias patterns, behavioral implications (e.g., approach and avoidance), and symptom severity represent critical next steps to empirically explore the etiological role of interpretation biases within emotional disorders.

Relatedly, all clinical groups endorsed more negative general interpretations, suggesting a common vulnerability factor reflected in this pattern (Buhlmann et al. 2002). Nonetheless, it should be noted that general situations encompass relatively heterogeneous current concerns (e.g., romantic relationships, finances, health concerns), which might, in sum, be of broader relevance *across* mental disorders as compared to more circumscribed concerns. In this respect, future studies might disentangle response patterns to specific current concerns contained in general situation sets.

In line with the current concern hypothesis, our results indicate a pronounced negative but no lack of positive interpretation bias for general situations in GAD compared to NC. One WSAP-based study had previously demonstrated this pattern in an unselected sample, showing that explicit threat bias, and not reduced positive bias, was predictive of GAD symptoms (Ogniewicz et al. 2014). The present study extends these findings by demonstrating the identical relationship in a sample meeting self-report DSM-5 GAD criteria.

Concerning the implicit, RT-related bias component, MLM analyses revealed a relatively inhomogeneous outcome. For appearance-related and social situations, the BDD and SAD (vs. GAD) group exhibited differences in RT, with slower endorsement of positive and slower rejection of negative interpretations being the most consistent findings. RT differences for these groups (vs. GAD and NC) were non-substantial for general situations. Moreover, findings for the GAD group reflected a differentially faster endorsement of negative and faster rejection of positive general interpretations compared to NC. We were further able to identify differential RT to both positive and negative general interpretations in GAD, while there was only a negative explicit interpretation bias. Overall, patterns of decision rates and RT partially diverged, which, within methodological limitations, might be due to several reasons,

such as quick interpretation changes or conflicting evaluation of individuated behaviors (see Rydell et al. 2008). In sum, these discrepancies highlight the benefit of RT-based indices as a supplement in understanding the architecture of interpretation bias. Nevertheless, more implicit RT results only partially supported the current concern hypothesis, as clinical groups showed alterations that were specific for some, but not all current concerns. Hence, further research is needed to reinvestigate the relationship of implicit and explicit interpretation bias, i.e., within clinically diagnosed, larger samples, and lab-based settings.

The examination of cognitive components via Wiener diffusion models revealed that between-group and between-content RT differences originated mainly from alterations in drift rate; that is, faster information uptake. As all other parameters did not vary substantially across groups and concerns, RT differences cannot be attributed to non-decisional factors, e.g., differential response caution (i.e., boundary separation) or response bias (i.e., initial bias). Drift rate relates to recognition and classification speed of stimuli, and thus, inter alia, associative strength in memory (McKoon and Ratcliff 2012; White et al. 2010). Hence, the current findings illustrate that faster reaction to an interpretation is a product of closely associated sentence-word-combinations being processed more easily than other distal combinations. In sum, as content-specific drift rate differences and endorsement rates were meaningfully associated with clinical group status, these results suggest that associative memory networks might indeed be organized as per disorder-relevant current concerns. Future studies should validate this novel application and interpretation of Wiener diffusion model indices for this task within larger replication studies.

This study has some limitations. Aiming to reach a heterogeneous range of participants, it was entirely web-based, thus conducted in a less controlled environment, with clinical diagnoses based on self-report DSM-5 criteria. Previous research indicates no validity differences between in-lab and web-based settings for most experimental tasks (e.g., Hilbig 2016; Ramsey et al. 2016; Semmelmann and Weigelt 2017), which is in line with the good psychometric properties found for all measures used in this study. Nonetheless, regarding self-report diagnoses, it remains unclear how well participants were able to accurately respond to diagnostic criteria, which might have affected validity. Overall, replication studies using clinician-administered interviews appear warranted.

Relatedly, the present study design did not allow for an assessment of comorbidities, which are prevalent in the disorders tested (Kessler et al. 2005). Depression, in particular, has been shown to be associated with a medium-sized effect on interpretation bias patterns (Everaert et al. 2014). Hence, future research should investigate the influence of comorbidities determined through standardized clinical interviews. In this respect, diagnostic procedures might especially aid in differentiating GAD from depression. The considerable overlap between these disorders (e.g., concerning repetitive negative thinking) has been empirically discussed (e.g., Mennin et al. 2008), highlighting problems in effectively distinguishing them through self-report. Notably, depression scores for the GAD group in this study were significantly higher than for all other clinical groups and might have influenced interpretation bias patterns.

Further, it is noteworthy that SWAP interpretation options might be somewhat heterogeneous, potentially affecting the association between interpretation bias and symptoms. For instance, while some interpretations reflect attributions (e.g., "funny" vs. "stupid"), others more indirectly refer to behavioral responses (e.g., "help" vs. "avoid"; see Supplementary Table A1 in the Appendix). This heterogeneity is an inherent characteristic of most interpretation bias stimulus sets and potentially yields a more comprehensive picture of bias features. Nonetheless, future studies might focus on specific aspects of interpretation bias (e.g., attribution) to test their association with symptom severity. Last, although our study was powered to detect large between-group differences in mentally healthy vs. clinical populations, statistical power might not have been sufficient to observe small to medium-sized effects (e.g., between clinical groups). The investigation of these effect sizes prospectively requires replication in a larger sample.

In sum, this study is the first to provide a comparative account of interpretation bias across BDD, SAD, and GAD, showing its transdiagnostic presence and specific modulation by current concerns, most coherently for explicit bias patterns. Our findings also demonstrate a shared propensity of BDD and SAD to misinterpret ambiguous appearance-related and social situations. Further, GAD appears to be characterized by both an explicit and implicit negative, as well as a lack of positive general interpretation bias. Additionally, mechanistic insights from RT-based results indicate a differential structural organization of associative memory within different disorders and current concerns, which is consistent with pertinent models of information processing (e.g., Beck and Haigh 2014). They further underscore the importance of identifying and targeting disorder-relevant interpretation bias profiles in cognitive-behavioral therapy, for example, via functional analyses, idiosyncratically tailored cognitive interventions, and CBM-I. In this respect, clinicians should attend to potential cognitive overlap, e.g., for appearance-related and social scenarios in BDD and SAD. Prospective CBM-I programs should equally incorporate such overlap and flexibly address different current

concerns within training rationales to enhance intervention efficacy.

## Compliance with Ethical Standards

**Conflict of Interest** The authors declare that they have no conflicts of interest.

**Informed Consent** Informed consent was obtained from all individual participants included in the study.

**Animal Rights Statement** This article does not contain any studies with animals performed by any of the authors.

## References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. https://doi.org/https://doi.org/10.1176/appi.books.9780890425596

Amir, N., Prouvost, C., & Kuckertz, J. M. (2012). Lack of a benign interpretation bias in social anxiety disorder. *Cognitive Behaviour Therapy, 41*(2), 119–129. https://doi.org/10.1080/16506073.2012.662655.

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior, 22*(3), 261–295. https://doi.org/10.1016/S0022-5371(83)90201-3.

Balota, D. A., & Yap, M. J. (2011). Moving beyond the mean in studies of mental chronometry: the power of response time distributional analyses. *Current Directions in Psychological Science, 20*(3), 160–166. https://doi.org/10.1177/0963721411408885.

Beard, C., & Amir, N. (2008). A multi-session interpretation modification program: changes in interpretation and social anxiety symptoms. *Behaviour Research and Therapy, 46*(10), 1135–1141. https://doi.org/10.1016/j.brat.2008.05.012.

Beard, C., & Amir, N. (2009). Interpretation in social anxiety: when meaning precedes ambiguity. *Cognitive Therapy and Research, 33*(4), 406–415. https://doi.org/10.1007/s10608-009-9235-0.

Beard, C., Weisberg, R. B., & Amir, N. (2011). Combined cognitive bias modification treatment for social anxiety disorder: a pilot trial. *Depression and Anxiety, 28*(11), 981–988. https://doi.org/10.1002/da.20873.

Beck, A. T., & Clark, D. A. (1988). Anxiety and depression: an information processing perspective. *Anxiety Research, 1*(1), 23–36. https://doi.org/10.1080/10615808808248218.

Beck, A. T., & Haigh, E. A. P. (2014). Advances in cognitive theory and therapy: the generic cognitive model. *Annual Review of Clinical Psychology, 10*, 1–24. https://doi.org/10.1146/annurev-clinpsy-032813-153734.

Beck, R., & Perkins, T. S. (2001). Cognitive content-specificity for anxiety and depression: a meta-analysis. *Cognitive Therapy and Research, 25*(6), 651–663. https://doi.org/10.1023/A:1012911104891.

Buhlmann, U., Wilhelm, S., Glaesmer, H., Brähler, E., & Rief, W. (2009). Fragebogen körperdysmorpher Symptome (FKS): Ein screening-instrument [Body dysmorphic symptoms inventory: A screening instrument]. *Verhaltenstherapie, 19*, 237–242. https://doi.org/10.1159/000246278.

Buhlmann, U., Wilhelm, S., McNally, R. J., Tuschen-Caffier, B., Baer, L., & Jenike, M. A. (2002). Interpretive biases for ambiguous information in body dysmorphic disorder. *CNS Spectrums, 7*(6), 435–443. https://doi.org/10.1017/S1092852900017946.

Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal, 10*(1), 395–411. https://doi.org/https://doi.org/10.32614/RJ-2018-017

Bürkner, P.-C. (2017). brms: an R package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 80*(1), 1–28. https://doi.org/https://doi.org/10.18637/jss.v080.i01

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: a probabilistic programming language. *Journal of Statistical Software, 76*(1). https://doi.org/https://doi.org/10.18637/jss.v076.i01

Cash, T. F., Fleming, E. C., Alindogan, J., Steadman, L., & Whitehead, A. (2002). Beyond body image as a trait: the development and validation of the Body Image States Scale. *Eating Disorders, 10*(2), 103–113. https://doi.org/10.1080/10640260290081678.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407–428. https://doi.org/10.1037/0033-295X.82.6.407.

Connor, K. M., Davidson, J. R., Churchill, L. E., Sherwood, A., Weisler, R. H., & Foa, E. (2000). Psychometric properties of the Social Phobia Inventory (SPIN): new self-rating scale. *The British Journal of Psychiatry, 176*(4), 379–386. https://doi.org/10.1192/bjp.176.4.379.

Cooper, M. J., Rose, K. S., & Turner, H. (2006). The specific content of core beliefs and schema in adolescent girls high and low in eating disorder symptoms. *Eating Behaviors, 7*(1), 27–35. https://doi.org/10.1016/j.eatbeh.2005.05.007.

Craske, M. G., Rauch, S. L., Ursano, R., Prenoveau, J., Pine, D. S., & Zinbarg, R. E. (2009). What is an anxiety disorder? *Depression and Anxiety, 26*(12), 1066–1085. https://doi.org/10.1002/da.20633.

Cristea, I. A., Kok, R. N., & Cuijpers, P. (2015). Efficacy of cognitive bias modification interventions in anxiety and depression: meta-analysis. *The British Journal of Psychiatry, 206*(1), 7–16. https://doi.org/10.1192/bjp.bp.114.146761.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: a normative analysis and review. *Psychological Bulletin, 135*(3), 347–368. https://doi.org/10.1037/a0014211.

Dietel, F. A., Möbius, M., Steinbach, L., Dusend, C., Wilhelm, S., & Buhlmann, U. (2018). Effects of induced appearance-related interpretation bias: a test of the cognitive-behavioral model of body dysmorphic disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 61*, 180–187. https://doi.org/10.1016/j.jbtep.2018.07.003.

Dozois, D. J. A., Martin, R. A., & Bieling, P. J. (2008). Early maladaptive schemas and adaptive/maladaptive styles of humor. *Cognitive*

*Therapy and Research, 33*(6), 585. https://doi.org/10.1007/s10608-008-9223-9.

Dugas, M. J., Freeston, M. H., Ladouceur, R., Rhéaume, J., Provencher, M., & Boisvert, J.-M. (1998). Worry themes in primary GAD, secondary GAD, and other anxiety disorders. *Journal of Anxiety Disorders, 12*(3), 253–261. https://doi.org/10.1016/S0887-6185(98)00013-9.

Everaert, J., Duyck, W., & Koster, E. H. W. (2014). Attention, interpretation, and memory biases in subclinical depression: a proof-of-principle test of the combined cognitive biases hypothesis. *Emotion, 14*(2), 331–340. https://doi.org/10.1037/a0035250.

Fang, A., & Hofmann, S. G. (2010). Relationship between social anxiety disorder and body dysmorphic disorder. *Clinical Psychology Review, 30*(8), 1040–1048. https://doi.org/10.1016/j.cpr.2010.08.001.

Fang, A., Sawyer, A. T., Aderka, I. M., & Hofmann, S. G. (2013). Psychological treatment of social anxiety disorder improves body dysmorphic concerns. *Journal of Anxiety Disorders, 27*(7), 684–691. https://doi.org/10.1016/j.janxdis.2013.07.005.

Fang, A., & Wilhelm, S. (2015). Clinical features, cognitive biases, and treatment of body dysmorphic disorder. *Annual Review of Clinical Psychology, 11,* 187–212. https://doi.org/10.1146/annurev-clinpsy-032814-112849.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/BF03193146.

Foa, E. B., Huppert, J. D., & Cahill, S. P. (2006). Emotional processing theory: an update. In B. O. Rothbaum (Ed.), *Pathological anxiety: emotional processing in etiology and treatment* (pp. 3–24). New York: The Guilford Press.

Gladwin, T. E., & Figner, B. (2014). "Hot" cognition and dual systems: introduction, criticisms, and ways forward. In E. Wilhelms & V. F. Reyna (Eds.), *Frontiers of cognitive psychology. Neuroeconomics, judgment, and decision making* (pp. 157–180). London: Psychology Press.

Gonsalves, M., Whittles, R. L., Weisberg, R. B., & Beard, C. (2019). A systematic review of the word sentence association paradigm (WSAP). *Journal of Behavior Therapy and Experimental Psychiatry, 64,* 133–148. https://doi.org/10.1016/j.jbtep.2019.04.003.

Gräfe, K., Zipfel, S., Herzog, W., & Löwe, B. (2004). Screening psychischer Störungen mit dem "Gesundheitsfragebogen für Patienten (PHQ-D)" [Screening for psychiatric disorders with the Patient Health Questionnaire (PHQ): Results from the German validation study]. *Diagnostica, 50*(4), 171–181. https://doi.org/10.1026/0012-1924.50.4.171.

Gunstad, J., & Phillips, K. A. (2003). Axis I comorbidity in body dysmorphic disorder. *Comprehensive Psychiatry, 44*(4), 270–276. https://doi.org/10.1016/S0010-440X(03)00088-9.

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.

Harvey, A. G., Watkins, E., & Mansell, W. (2004). *Cognitive behavioural processes across psychological disorders: A transdiagnostic approach to research and treatment*. Oxford University Press.

Hazlett-Stevens, H., & Borkovec, T. D. (2004). Interpretive cues and ambiguity in generalized anxiety disorder. *Behaviour Research and Therapy, 42*(8), 881–892. https://doi.org/10.1016/S0005-7967(03)00204-3.

Heathcote, A., Popiel, S. J., & Mewhort, D. J. (1991). Analysis of response time distributions: an example using the Stroop task. *Psychological Bulletin, 109*(2), 340–347. https://doi.org/10.1037/0033-2909.109.2.340.

Hilbig, B. E. (2016). Reaction time effects in lab-versus web-based research: experimental evidence. *Behavior Research Methods, 48*(4), 1718–1724. https://doi.org/10.3758/s13428-015-0678-9.

Hindash, A. H. C., & Amir, N. (2012). Negative interpretation bias in individuals with depressive symptoms. *Cognitive Therapy and Research, 36*(5), 502–511. https://doi.org/10.1007/s10608-011-9397-4.

Hirsch, C. R., & Clark, D. M. (2004). Information-processing bias in social phobia. *Clinical Psychology Review, 24*(7), 799–825. https://doi.org/10.1016/j.cpr.2004.07.005.

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*(1), 281–305. https://doi.org/10.1146/annurev-clinpsy-021815-093436.

Hoffman, L., & Rovine, M. J. (2007). Multilevel models for the experimental psychologist: foundations and illustrative examples. *Behavior Research Methods, 39*(1), 101–117. https://doi.org/10.3758/BF03192848.

Hofmann, S. G. (2007). Cognitive factors that maintain social anxiety disorder: a comprehensive model and its treatment implications. *Cognitive Behaviour Therapy, 36*(4), 193–209. https://doi.org/10.1080/16506070701421313.

Huppert, J. D., Foa, E. B., Furr, J. M., Filip, J. C., & Mathews, A. (2003). Interpretation bias in social anxiety: a dimensional perspective. *Cognitive Therapy and Research, 27*(5), 569–577. https://doi.org/10.1023/A:1026359105456.

Jaccard, J., Becker, M. A., & Wood, G. (1984). Pairwise multiple comparison procedures: a review. *Psychological Bulletin, 96*(3), 589–596. https://doi.org/10.1037/0033-2909.96.3.589.

Jones, E. B., & Sharpe, L. (2017). Cognitive bias modification: a review of meta-analyses. *Journal of Affective Disorders, 223,* 175–183. https://doi.org/10.1016/j.jad.2017.07.034.

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

Klauer, K. C., Voss, A., Schmitz, F., & Teige-Mocigemba, S. (2007). Process components of the implicit association test: a diffusion-model analysis. *Journal of Personality and Social Psychology, 93*(3), 353–368. https://doi.org/10.1037/0022-3514.93.3.353.

Klinger, E., & Cox, W. M. (2011). Motivation and the goal theory of current concerns. In W. M. Cox & E. Klinger (Eds.), *Handbook of motivational counseling: goal-based approaches to assessment and intervention with addiction and other problems* (pp. 1–47). New York: John Wiley and Sons, Ltd. https://doi.org/https://doi.org/10.1002/9780470979952.ch1

Kroenke, K., & Spitzer, R. L. (2002). The PHQ-9: a new depression diagnostic and severity measure. *Psychiatric Annals, 32*(9), 509–515. https://doi.org/10.3928/0048-5713-20020901-06.

Kuckertz, J. M., & Amir, N. (2017). Cognitive bias modification. In S. G. Hofmann & G. J. G. Asmundson (Eds.), *The science of cognitive behavioral therapy* (pp. 463–491). Cambridge: Elsevier Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-803457-6.00019-2

Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika, 40*(1), 77–105. https://doi.org/10.1007/BF02291481.

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.

Mansell, W., Harvey, A., Watkins, E. R., & Shafran, R. (2008). Cognitive behavioral processes across psychological disorders: a review of the utility and validity of the transdiagnostic approach. *International Journal of Cognitive Therapy, 1*(3), 181–191. https://doi.org/10.1521/ijct.2008.1.3.181.

Mathews, A., & MacLeod, C. (2005). Cognitive vulnerability to emotional disorders. *Annual Review of Clinical Psychology, 1,* 167–195. https://doi.org/10.1146/annurev.clinpsy.1.102803.143916.

McKoon, G., & Ratcliff, R. (2012). Aging and IQ effects on associative recognition and priming in item recognition. *Journal of Memory and Language, 66*(3), 416–437. https://doi.org/10.1016/j.jml.2011.12.001.

McManus, F., Clark, D. M., & Hackmann, A. (2000). Specificity of cognitive biases in social phobia and their role in recovery. *Behavioural and Cognitive Psychotherapy, 28*(3), 201–209. https://doi.org/10.1017/S1352465800003015.

Menne-Lothmann, C., Viechtbauer, W., Höhn, P., Kasanova, Z., Haller, S. P., 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.

Mennin, D. S., Heimberg, R. G., Fresco, D. M., & Ritter, M. R. (2008). Is generalized anxiety disorder an anxiety or mood disorder? Considering multiple factors as we ponder the fate of GAD. *Depression and Anxiety, 25*(4), 289–299. https://doi.org/10.1002/da.20493.

Millisecond Software. (2015). *Inquisit 4* [Computer program]. https://www.millisecond.com

Möbius, M., Tendolkar, I., Lohner, V., Baltussen, M., & Becker, E. S. (2015). Refilling the half-empty glass—investigating the potential role of the Interpretation Modification Paradigm for Depression (IMP-D). *Journal of Behavior Therapy and Experimental Psychiatry, 49*, 37–43. https://doi.org/10.1016/j.jbtep.2015.03.002.

Newman, M. G., Zuellig, A. R., Kachin, K. E., Constantino, M. J., Przeworski, A., Erickson, T., & Cashman-McGrath, L. (2002). Preliminary reliability and validity of the Generalized Anxiety Disorder Questionnaire-IV: a revised self-report diagnostic measure of generalized anxiety disorder. *Behavior Therapy, 33*(2), 215–233. https://doi.org/10.1016/S0005-7894(02)80026-0.

Nezlek, J. B., & Kuppens, P. (2008). Regulating positive and negative emotions in daily life. *Journal of Personality, 76*(3), 561–580. https://doi.org/10.1111/j.1467-6494.2008.00496.x.

Ogniewicz, A. S., Dugas, M. J., Langlois, F., Gosselin, P., & Koerner, N. (2014). An adapted word-sentence association paradigm for generalized anxiety and worry: assessing interpretation bias. *Journal of Experimental Psychopathology, 5*(4), 473–476. https://doi.org/10.5127/jep.00.

Ouimet, A. J., Gawronski, B., & Dozois, D. J. A. (2009). Cognitive vulnerability to anxiety: a review and an integrative model. *Clinical Psychology Review, 29*(6), 459–470. https://doi.org/10.1016/j.cpr.2009.05.004.

Pinto, A., & Phillips, K. A. (2005). Social anxiety in body dysmorphic disorder. *Body Image, 2*(4), 401–405. https://doi.org/10.1016/j.bodyim.2005.10.003.

Premo, J. E., Sarfan, L. D., & Clerkin, E. M. (2016). Training interpretation biases among individuals with body dysmorphic disorder symptoms. *Body Image, 16,* 54–62. https://doi.org/10.1016/j.bodyim.2015.11.004.

Price, R. B., Brown, V., & Siegle, G. J. (2019). Computational modeling applied to the dot-probe task yields improved reliability and mechanistic insights. *Biological Psychiatry, 85*(7), 606–612. https://doi.org/10.1016/j.biopsych.2018.09.022.

Quené, H., & Van den Bergh, H. (2004). On multilevel modeling of data from repeated measures designs: a tutorial. *Speech Communication, 43*(1–2), 103–121. https://doi.org/10.1016/j.specom.2004.02.004.

R Core Team. (2018). *R: a language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Ramsey, S. R., Thompson, K. L., McKenzie, M., & Rosenbaum, A. (2016). Psychological research in the internet age: the quality of web-based data. *Computers in Human Behavior, 58,* 354–360. https://doi.org/10.1016/j.chb.2015.12.049.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*(2), 59–108. https://doi.org/10.1037/0033-295X.85.2.59.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Computation, 20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420.

Rosser, B. A., Moss, T., & Rumsey, N. (2010). Attentional and interpretative biases in appearance concern: an investigation of biases in appearance-related information processing. *Body Image, 7*(3), 251–254. https://doi.org/10.1016/j.bodyim.2010.02.007.

RStudio Team. (2018). *RStudio: integrated development for R*. RStudio, Inc. http://www.rstudio.com

Rydell, R. J., McConnell, A. R., & Mackie, D. M. (2008). Consequences of discrepant explicit and implicit attitudes: Cognitive dissonance and increased information processing. *Journal of Experimental Social Psychology, 44*(6), 1526–1532. https://doi.org/10.1016/j.jesp.2008.07.006.

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. https://doi.org/10.3389/fpsyg.2017.00171.

Schulz, S. M., Alpers, G. W., & Hofmann, S. G. (2008). Negative self-focused cognitions mediate the effect of trait social anxiety on state anxiety. *Behaviour Research and Therapy, 46*(4), 438–449. https://doi.org/10.1016/j.brat.2008.01.008.

Semmelmann, K., & Weigelt, S. (2017). Online psychophysics: reaction time effects in cognitive experiments. *Behavior Research Methods, 49*(4), 1241–1260. https://doi.org/10.3758/s13428-016-0783-4.

Sosic, Z., Gieler, U., & Stangier, U. (2008). Screening for social phobia in medical in- and outpatients with the German version of the Social Phobia Inventory (SPIN). *Journal of Anxiety Disorders, 22*(5), 849–859. https://doi.org/10.1016/j.janxdis.2007.08.011.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *State-trait anxiety inventory (form Y)* (p. 77). Redwood City, CA: Mind Garden.

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. https://doi.org/10.1080/02699931.2019.1682973.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8*(3), 220–247. https://doi.org/10.1207/s15327957pspr0803_1.

Turk, C. L., Heimberg, R. G., Luterek, J. A., Mennin, D. S., & Fresco, D. M. (2005). Emotion dysregulation in generalized anxiety disorder: a comparison with social anxiety disorder. *Cognitive Therapy and Research, 29*(1), 89–106. https://doi.org/10.1007/s10608-005-1651-1.

van Ravenzwaaij, D., van der Maas, H. L. J., & Wagenmakers, E.-J. (2012). Optimal decision making in neural inhibition models. *Psychological Review, 119*(1), 201–215. https://doi.org/10.1037/a0026275.

van Ravenzwaaij, D., van der Maas, H. L., & Wagenmakers, E.-J. (2011). Does the name-race Implicit Association Test measure racial prejudice? *Experimental Psychology, 58*(4), 271–277. https://doi.org/10.1027/1618-3169/a000093.

Vandekerckhove, J., Tuerlinckx, F., & Lee, M. D. (2011). Hierarchical diffusion models for two-choice response times. *Psychological Methods, 16*(1), 44–62. https://doi.org/10.1037/a0021765.

Voncken, M. J., Bögels, S. M., & de Vries, K. (2003). Interpretation and judgmental biases in social phobia. *Behaviour Research and Therapy, 41*(12), 1481–1488. https://doi.org/10.1016/S0005-7967(03)00143-8.

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. https://doi.org/10.1348/147608306X161890.

Voss, A., Nagler, M., & Lerche, V. (2013). Diffusion models in experimental psychology: a practical introduction. *Experimental Psychology, 60*(6), 385–402. https://doi.org/10.1027/1618-3169/a000218.

Voss, A., Voss, J., & Lerche, V. (2015). Assessing cognitive processes with diffusion model analyses: a tutorial based on fast-dm-30. *Frontiers in Psychology, 6,* 336. https://doi.org/10.3389/fpsyg.2015.00336.

Wagenmakers, E.-J. (2009). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive Psychology, 21*(5), 641–671. https://doi.org/10.1080/09541440802205067.

Whelan, R. (2008). Effective analysis of reaction time data. *The Psychological Record, 58*(3), 475–482. https://doi.org/10.1007/BF03395630.

White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010). Using diffusion models to understand clinical disorders. *Journal of Mathematical Psychology, 54*(1), 39–52. https://doi.org/10.1016/j.jmp.2010.01.004.

Wilhelm, S., Phillips, K. A., & Steketee, G. (2013). *Cognitive behavioral therapy for body dysmorphic disorder—A treatment manual*. New York, NY: The Guilford Press.

Wickham, H. (2017). *Tidyverse: Easily install and load the 'tidyverse'*. R package version 1.2.1. Vienna, Austria: R Core Team.

Wisco, B. E., & Nolen-Hoeksema, S. (2010). Interpretation bias and depressive symptoms: the role of self-relevance. *Behaviour Research and Therapy, 48*(11), 1113–1122. https://doi.org/10.1016/j.brat.2010.08.004.