Measurement of Target Engagement and Network Analysis of Change Mechanisms in Web-Based Interpretation Bias Training for Anxiety

Jeremy W. Eberle

BS, BA, Tulane University, 2008

MA, University of Virginia, 2019

A dissertation presented to the

graduate faculty of the University of Virginia in

candidacy for the degree of Doctor of Philosophy

Department of Psychology

University of Virginia

August 2024

Committee:

Bethany A. Teachman, PhD (Chair)

Steven M. Boker, PhD

Teague R. Henry, PhD

Laura E. Barnes, PhD

Acknowledgments

I am grateful for the many people who supported me in this project and in my graduate education. I would like to thank my fellow lab members of the Program for Anxiety, Cognition, and Treatment at the University of Virginia for their comments on earlier versions of this dissertation and for their collaboration, friendship, and support. I would also like to thank our engineering and computer science colleagues on the MindTrails Project research team for their contributions to the MindTrails website, through which the present data were collected as part of the Managing Anxiety study. Special thanks to Dan Funk, who led software development; Sonia Baee, who led data cleaning; Diheng Zhang, who coordinated the study; and Bethany Teachman, Laura Barnes, and Mehdi Boukhechba, who oversaw the study. This work was funded by grants awarded to Teachman and Barnes from the National Institute of Mental Health (R34MH106770, R01MH113752) and the John Templeton Foundation (Science of Prospection Research Award). I greatly appreciate the impactful discussions I had in Brian Nosek's methodological reading group, the input I received from Sacha Epskamp and Terrence Jorgensen about their statistical packages, and the modeling conversations I shared with Whitney Ringwald. I am very thankful to Steve Boker and Teague Henry for their foundational training in structural equation modeling and network analysis and to my entire dissertation committee for their feedback and insights. I am particularly indebted to Bethany Teachman for her scientific rigor, supportive mentorship, enthusiasm, and humor throughout every step of graduate school. To all my mentors, colleagues, and collaborators over the years, and to my friends and family—thank you all.

Table of Contents

General Abstract

Anxiety disorders are prevalent but undertreated. Cognitive bias modification (CBM) programs target disorder-relevant processing biases without needing a therapist, providing a potential way to improve access to treatment (e.g., via web- or app-based delivery). Although CBM results have been mixed, overall the evidence suggests that CBM for interpretation biases (CBM-I) may reduce anxiety (and comorbid depression). CBM-I provides repeated practice resolving ambiguous threat-relevant situations in a benign way, aiming to foster more flexible thinking and shift the rigidly negative interpretation style associated with emotional disorders.

However, the psychometric properties of common measures of interpretation biases, the purported mechanism of change in CBM-I, are unacceptable in some studies and unreported in most. Unreliable or invalid measures make it difficult to evaluate target engagement in clinical trials and to assess the mechanism's effects on symptoms. To advance research on cognitive mechanisms in anxiety, **Study 1** evaluated the structural validity (factor structure and internal consistency) of a Recognition Ratings measure used as a primary measure of interpretation biases in CBM-I trials. Using baseline data from a trial run with anxious community adults (N = 749) on a public research website, initial confirmatory factor analysis (CFA) models inferred from the literature were unsupported. Exploratory factor analysis (EFA) suggested three factors (positive threat, negative threat, nonthreat) and potential items to exclude, and exploratory CFA models suggested the need to correlate the errors of items from the same scenario. Exploratory CFA models with (a) the three factors based on a subset of 28 threat and nonthreat items or (b) two factors (positive threat, negative threat) based on all 18 threat items had well-defined factors with generally acceptable internal consistency and construct reliability. Factor determinacy was acceptable for the negative threat and nonthreat factors but mixed for positive threat. Overall

CFA model fit was mixed. Exploratory CFA models of the threat items supported the positive threat and negative threat factors as two distinct constructs (i.e., not method artifacts). The study provides initial evidence for the measure's use in assessing positive and negative biases, although ongoing construct validation will further optimize its use. The two-factor model was selected for Study 2 given its balance of 9 positive threat and 9 negative threat items (unlike the three-factor model with 5 positive threat and 8 negative threat items).

Moreover, most research on interpretation biases and anxiety has conceptualized anxiety symptoms as caused by an underlying disorder, whereas network theory posits that disorder is constituted by causal relations among symptoms themselves, allowing for analysis of anxiety as a complex system. Adopting this perspective, **Study 2** uses data from the same randomized trial to test the effects of positive CBM-I (vs. 50-50 CBM-I and no-training comparison conditions) on networks of interpretation biases, individual anxiety symptoms, and related impairment. Cross-sectional network models testing positive CBM-I's effects on the mean levels of nodes at each of three time points (baseline, Session 3, Session 6) revealed significant direct effects (causal) and indirect effects (possibly causal), suggesting potential pathways via which positive CBM-I deactivates the network. In addition, analyses of each condition's within-person relations in temporal and contemporaneous networks across the time points revealed descriptively lower connectivity in positive CBM-I than in 50-50 CBM-I and no-training. Although these condition differences in connectivity need to be tested statistically, this tentatively suggests that positive CBM-I may also destabilize the network's structure, which may reduce the vulnerability of the network to a stable state of high activation. Together, these studies strengthen inferences about interpretation biases and advance understanding of cognitive mechanisms of change in anxiety.

Keywords: interpretation bias, anxiety, factor analysis, network analysis, mechanisms

General Introduction

Cognitive models of anxiety posit that selective interpretation of ambiguous situations as threatening, or *interpretation bias,* plays a causal role in the onset, maintenance, and reduction of anxiety (MacLeod & Mathews, 2012). Cognitive bias modification of interpretations (CBM-I) aims to shift this bias by providing practice assigning benign meanings to ambiguous situations (Teachman, 2014). Web- or app-based CBM-I programs hold promise as scalable treatments for the many people with anxiety (Baxter et al., 2013) who lack access to treatment (Chisholm et al., 2016) because they can be delivered digitally without requiring a therapist. For instance, the widely used CBM-I *ambiguous scenarios paradigm* (Mathews & Mackintosh, 2000) presents brief scenarios that involve a potential threat and remain ambiguous until a final word fragment. Completing the fragment resolves the ambiguity by assigning a positive interpretation (e.g., "You decide to have a barbeque, as the weather is so nice. As your friends arrive, you realize many of them do not know each other well. Seeing everyone there, you realize that your friends probably think that, as a host, you are succ_ssful."). The largest meta-analysis of CBM trials to date found that, despite mixed findings, CBM-I was superior overall to waitlist and active control groups in reducing anxiety (Fodor et al., 2020). This dissertation aims to strengthen research on CBM-I and cognitive mechanisms of anxiety by (a) evaluating the psychometric properties of a primary measure of interpretation biases and (b) using network modeling to test how CBM-I improves these biases and anxiety symptoms (i.e., to clarify its *mechanisms of change*).

Despite the clinical potential of CBM-I and its utility for testing cognitive theories of anxiety by directly targeting interpretation bias (MacLeod & Mathews, 2012), the psychometric properties of common measures of interpretation bias (e.g., Recognition Ratings, RR; Mathews & Mackintosh, 2000) are not reported in most studies. Examples abound of measures long assumed to be valid lacking evidence of validity once this is tested (Hussey & Hughes, 2020). The need for testing validity also applies to measures of cognitive biases more broadly. For example, after decades of research in the subfield of attention bias, the low reliability of its measures has emerged as a serious limitation (McNally, 2019). In addition, the reliability and validity of some interpretation bias measures (including a version of RR) are sometimes poor, which may contribute to mixed findings on relations between interpretation biases and anxiety (Duken et al., 2024). To detect and address any unrecognized measurement issues in the field of interpretation bias, the properties of its measures must be thoroughly evaluated. Using baseline data from a hybrid efficacy-effectiveness CBM-I trial (Ji et al., 2021) conducted in a large sample of anxious community adults on our team's public research website, **Study 1** tests and optimizes the structural validity (e.g., factor structure, internal consistency) of our team's RR for use in Study 2 and in anxious samples more generally.

In addition to evaluating the validity of this measure of target engagement in CBM-I trials, **Study 2** of this dissertation aims to clarify how CBM-I affects interpretation biases, individual anxiety symptoms, and anxiety-related impairment (and their relations) when these constructs are conceptualized as a network of interacting elements (Borsboom & Cramer, 2013; Hoffart & Johnson, 2020a). This network perspective of psychopathology holds promise for studying mechanisms of change because it focuses on causal relations among elements in the network, rather than assuming that the elements do not affect one another and instead co-occur due to an underlying disorder (Borsboom & Cramer, 2013; Hofmann et al., 2020). The study uses data from the same trial above, in which participants were randomly assigned to a positive CBM-I condition, a 50-50 CBM-I condition, or a no-training control condition. In positive CBM-I, ambiguous scenarios resolve positively (e.g., with "succ_ssful" in the scenario above) 90% of the time and negatively (e.g., with “faili_g”) 10% of the time. In 50/50 CBM-I, scenarios resolve positively 50% of the time and negatively 50% of the time, so that no reliable contingency is learned about whether ambiguous situations are likely to resolve in a positive or negative way.

Using this trial’s data, first, network intervention analyses (Blanken et al., 2019) will test the effects of CBM-I on *mean levels* of interpretation biases, anxiety symptoms, and related impairment in cross-sectional networks at different time points. Consistent with cognitive models (e.g., Smits et al., 2012; Steinman & Teachman, 2014), we expect positive CBM-I (vs. 50/50 CBM-I and no-training) to have early, direct effects on biases (causal due to randomization to condition) and indirect effects on symptoms and impairment via biases (possibly causal due to lack of temporal precedence for the biases’ relations with other network elements). Second, temporal network models (Epskamp, 2020) will estimate the lagged and contemporaneous within-person *relations* among these elements across time in each condition. A network that has stronger relations among its elements is thought to constitute greater pathology because when an external stressor activates certain elements, those elements may in turn activate other elements given the network’s dense connections (Borsboom, 2017). Adopting this perspective, we expect positive CBM-I’s networks to have less overall connectivity than the other conditions’ networks; that is, CBM-I may weaken dense relations that confer vulnerability to spreading activation (e.g., activation of one anxiety symptom will be less likely to in turn activate other anxiety symptoms).

Taken together, these studies harness the large transdiagnostic sample and experimental design of a trial of web-based CBM-I to advance research on interpretation biases and anxiety. By testing and optimizing the structural validity of a primary measure of interpretation biases, this dissertation will shed light on and improve the measurement of such biases, which is key to showing target engagement in CBM-I trials and to drawing inferences about the biases in other mechanistic studies (see Rodebaugh et al., 2016). Further, modeling interpretation biases, anxiety symptoms, and impairment as a network will clarify CBM-I's specific effects on these elements and their relations, advancing understanding of cognitive mechanisms of change in anxiety.

Study 1: Measurement of Target Engagement

Recognizing the limitations of developing a treatment package for each *DSM* disorder (e.g., we end up with many different lengthy and complex treatments, instead of investigating specific mechanisms and targeted intervention techniques), researchers are increasingly focused on identifying processes that cut across disorders and methods for targeting them (Hofmann & Hayes, 2019), such as CBM-I programs that target a transdiagnostic tendency to (over)interpret threat in anxiety disorders (Beard & Peckham, 2020). However, the psychometric properties of common measures of interpretation bias are unacceptable in some studies or, in most studies, unreported. Because claims about constructs depend on the reliability and validity of their measures (Flake et al., 2017), unreliable or invalid measures of interpretation bias could undermine claims about not only target engagement in CBM-I trials, but also causal effects of this bias on anxiety, which is the key rationale for the use of CBM-I as a treatment.

Untested measurement runs the risk of hidden invalidity. For example, an analysis of commonly used measures in social psychology showed that 88% seemed to have good structural validity based on internal consistency, the modal reported metric, but that only 4% had good validity when evaluated on factor structure, measurement invariance, and test-retest reliability. In other words, 96% of the measures failed validity tests on at least one of these four metrics (Hussey & Hughes, 2020). Such structural invalidity can undermine claims of external validity (e.g., presumed group differences, relations between constructs, and predictive effects may be invalid; Hussey & Hughes) and other claims of substantive interest (e.g., that change over time represents true change and not merely measurement error; Duken et al., 2024). Invalidity has also been a longstanding concern for measures of attentional bias (McNally, 2019) and recently emerged as a concern for some measures of interpretation bias (Duken et al.), highlighting the need for thorough evaluation of interpretation bias measures. Notably, it is important to evaluate more than simply internal consistency. For example, measures of internal consistency assume unidimensionality; factor analysis is required to assess the number of dimensions and how they are related (Dunn et al., 2014). In the present study, we test and optimize the structural validity of a primary measure of interpretation biases using data from a CBM-I trial conducted in anxious community adults on our team's public research website *MindTrails*.

Validity of Recognition Ratings

One common measure of interpretation bias is Recognition Ratings (RR; Mathews & Mackintosh, 2000), in which participants read a set of ambiguous scenarios (originally 10 with potential social threat and 10 with potential nonsocial threat) and are given each scenario's title and four possible final sentences for the scenario, two representing positive or negative interpretations of the potential threat raised in the scenario and two representing positive or negative interpretations unrelated to the threat. Participants are asked to rate how similar each sentence's meaning is to the final sentence of the scenario they previously read on a 4-point Likert scale ranging from 1 (*very different in meaning*) to 4 (*very similar in meaning*). CBM-I trials have used the mean negative and positive threat-related interpretation ratings as primary measures of negative and positive interpretation bias (e.g., Ji et al., 2021), though the number and type of scenarios and response options vary by study, and the specific materials used in a particular study are often not disclosed (Duken et al., 2024).

Despite the widespread use of RR, especially in studies of CBM-I, most studies using RR report none of the psychometric properties above (or only internal consistency). Some studies using RR scenarios from the original set used in Mathews and Mackintosh (2000) have reported questionable internal consistency (e.g., range of Cronbach's αs = .60-.69; Edwards et al., 2018; where .70 is deemed acceptable, though this cutoff is arbitrary; Dunn et al., 2014). Many studies have used modified RR scenarios to reflect the particular threat domain(s) targeted in CBM-I, with internal consistencies that are at times acceptable (αs = .73-.85; Ji et al., 2021) and other times very low (e.g., αs = .12-.22; Reuland & Teachman, 2014). We know of one study of RR that reported McDonald's omega, which avoids many issues with Cronbach's alpha (Dunn et al., 2014), finding acceptable internal consistency (ω_ts = .74-.79; Larrazabal et al., 2024). To our knowledge, no studies using RR in anxious samples have tested its factor structure or invariance. We know of one study (using Dutch materials) that evaluated its test-retest reliability, which was low over 1 week, and that also found poor internal consistency and poor convergent and concurrent validity (Duken et al., 2024). Some evidence suggests that RR differentiates people high and low in neuroticism (supporting its known groups validity) and is insensitive to mood state (supporting its ability to measure change in bias alongside change in anxiety during CBM-I; Salemink & van den Hout, 2010). However, RR's structural validity is largely unknown.

Overview and Hypotheses

In this study, we will test the structural validity of RR using data from a randomized controlled trial (RCT) of web-based CBM-I run in a large transdiagnostic sample of anxious adults on a public research website called *MindTrails* (https://mindtrails.virginia.edu). The trial (Ji et al., 2021) compared positive CBM-I, 50/50 CBM-I, and a no-training condition. With respect to the psychometric properties of RR in the trial, only Cronbach's alphas for baseline positive and negative interpretation bias (.73-.85) have been computed. We will test the factor structure and internal consistency of RR given that an interpretable factor structure is needed before invariance can be tested (Rosellini & Brown, 2021), and we will not assess test-retest reliability given that no condition was purely observational (i.e., the no-training condition still involved imagery prime tasks). We will use a staged approach, starting with confirmatory tests using baseline data (preregistration: https://doi.org/mt9r).[1] In these tests, we will assess a correlated factor model inferred from the RR development paper (Mathews & Mackintosh, 2000) and associated bifactor and higher-order models assumed in the practice of combining items across factors in scoring (e.g., Baee et al., 2024). If these models are unsupported (based on model fit, factor loading pattern, internal consistency, factor determinacy, or construct reliability; see guidelines below), we will run exploratory analyses to discover the factor structure revealed by the data and propose improved models aligned with theory and the data (e.g., exploratory factor analysis [EFA] to identify improved correlated factor models; details below).

Method for Study 1

Participants and Procedure

The present study analyzed existing baseline data from the Managing Anxiety trial (Ji et al., 2021), which was run in community adults with at least moderate anxiety symptoms (≥ 10 on doubled DASS-21-Anxiety total score, vs. DASS-42-Anxiety norms; Lovibond & Lovibond, 1995). We focused on intent-to-treat (ITT) participants (807 overall) with at least some item-level RR baseline data (N = 749). Participants were not compensated.

Most analyzed participants were female (72.4%), White (77.7%), not Hispanic or Latino (83.0%) adults (M = 33.82 years, SD = 13.56) from the United States (76.8%) who had finished at least some college (53.7%) or at least some graduate school (36.7%). Most were working full

[1] The structural validity preregistration outlines four sets of analyses. This dissertation includes analyses of RR data from Managing Anxiety that are part of the first set of preregistered analyses (Study 1 of preregistration).

(42.1%) or part (11.3%) time or were students (29.1%). Most annually earned less than $50,000 (36.9%) or between $50,000 and $100,000 (22.6%); some earned more than $100,000 (17.7%). Most were in a relationship (59.0%) or single (32.6%). For full demographics, see Table 1.

Measure

Recognition Ratings

The trial used RR consisting of 9 scenarios (3 social threat, 3 physical/health threat, 3 other threat) developed by our team to resemble those used in Mathews and Mackintosh (2000). These 9 scenarios were selected out of a set of 18 scenarios (6 for each threat type) that were piloted in a sample of 274 participants on Amazon Mechanical Turk in October 2015. The 3 scenarios whose negative threat-related ratings most highly correlated with anxiety (total score on Overall Anxiety Severity and Impairment Scale, OASIS; Norman et al., 2006) within each threat type were selected (B. Teachman, personal communication, June 9, 2022). The types of the four possible final sentences for each scenario were not changed from the original. That is, two represented positive or negative threat-relevant interpretations (called *targets* because they are focused on the anxiety-relevant threat), and two represented positive or negative threat-irrelevant interpretations (called *foils* because they are valenced but not related to the anxiety-relevant threat). However, the comprehension questions and items for two scenarios used in Mathews and Mackintosh were modified, and the other seven scenarios and their comprehension questions and items were newly created by our team. The same anchors of the 4-point Likert scale used by Mathews and Mackintosh, which ranged from 1 (*very different*) to 4 (*very similar*), were used in the present trial, but on a scale ranging from 0 (*very different*) to 3 (*very similar*; numbers were shown alongside anchors). See Appendix A for the measure used in the present trial and differences from that used in Mathews and Mackintosh. See Table SA1 in Supplement

A (at https://osf.io/ebn25/) for each item's valence and threat relevance.

Statistical Analysis

All significance tests in this study are two-tailed with an alpha level of .05. For data and analysis code, see https://osf.io/ebn25/. For preregistration deviations, see Section SA2.

Initial Confirmatory Factor Analyses

All 36 Threat and Nonthreat Items.

Correlated Factor Model. First, we used confirmatory factor analysis (CFA) to test the assumed factor structure of all 36 items. We tested four dimensions with a correlated four-factor model (9 positive threat items, 9 positive nonthreat items, 9 negative threat items, 9 negative nonthreat items; Model 1; Figure SA1) given that most studies have scored positive and negative threat items separately (as measures of target engagement; e.g., Ji et al., 2021; Eberle et al., in press; Hohensee et al., 2020; Larrazabal et al., 2024; Schmitt et al., 2023; Steinman et al., 2020; Silverman et al., 2024; Vela de la Garza Evia et al., 2024). Studies that have used foil nonthreat items have also distinguished between positive and negative items (which have been used to compute alternative measures of interpretation biases; e.g., Ji et al., 2024).

Bifactor and Higher-Order Models. We also tested two models that could support the practice of combining items across factors in scoring. For example, Baee et al. (2024) combined the threat and nonthreat items for each valence (positive, negative) when scoring RR. These composites assume a bifactor model with two *general factors* (positive, negative), each of which influences its items independently from two *specific* (or *group*) *factors* (threat, nonthreat; Model 2; Figure SA2). In this case, the general factors reflect the valence of the response options, and the specific factors reflect the threat relevance of the response options. Alternatively, a higher-order model with two *second-order factors* (positive, negative), each of which influences its items via two *first-order factors* (threat, nonthreat; Model 3; Figure SA3), could support use of the composites despite multidimensionality. In the bifactor model, we fixed covariances among the specific factors and between the specific and general factors to 0 and allowed the general factors to covary (Morin et al., 2020, p. 1055; for examples, see Figure 1 of Cai, 2010, p. 583; Model 4 of Gäde et al., 2017, p. 3; Figure 1 of Jimenez et al., 2023, p. 3). In the higher-order model, we allowed the higher-order factors to covary (e.g., Figure 8.1 of Brown, 2015, p. 289).

Model Specification. We fit the CFA models using the `cfa` function of the `lavaan` package (ver. 0.6-17; Rosseel, 2012) in R (ver. 4.3.2; R Core Team, 2023). We scaled all latent variables by fixing their variances to 1 (using the `std.lv` argument; which also fixes the residual variances of the first-order factors in the higher-order model to 1) and interpreted the completely standardized parameter estimates. Per Flora (2020), because each ordinal RR item has only four response options on the Likert scale for rating how similar the item's meaning is to one's initial interpretation of the corresponding scenario, we treated the items as categorical. We fit models to polychoric correlations using diagonally weighted least squares (DWLS) estimation with robust standard errors and a mean- and variance-adjusted test statistic (`lavaan`'s `WLSMV` estimator; see Finney & DiStefano, 2013; Rosellini & Brown, 2021) and pairwise deletion. For deviations from our preregistered estimation method, see Section SA2.1.

Model Evaluation. We evaluated model fit based on overall goodness of fit, interpreting absolute fit indices (standardized root mean square residual, SRMR; adjusted chi-square, though given our large sample we de-emphasized this, as negligible misfit can yield a significant chi-square, which was significant in every factor analysis in this study; Tables 2-5), a parsimony correction index (root mean square error of approximation, RMSEA), and comparative fit indices (comparative fit index, CFI; Tucker-Lewis Index, TLI). We interpreted `lavaan`'s robust versions of the RMSEA, CFI, and TLI corrected for DWLS estimation with categorical data, given that the original versions were created for maximum likelihood estimation with continuous data and show better fit when used with categorical data (Savalei, 2021).

Traditional guidelines for “relatively good” fit (SRMR near or < .08; RMSEA near or < 0.06; CFI and TLI near or > .95; Hu & Bentler, 1999) are based on simulated data from one specific model (3 factors, 15 items, 3 cross-loadings) and may not generalize to other models (McNeish & Wolf, 2024). Thus, we also used the `catHB` function of the `dynamic` package (per GitHub commit `bf9a430`) to compute dynamic fit index (DFI) cutoffs (for SRMR and nonrobust versions of the RMSEA and CFI under the `WLSMV` estimator). DFI cutoffs aim to identify what Hu and Bentler’s guidelines would have been for a given model at hand (McNeish, 2023). However, the DFI cutoffs were unavailable for most of the models in which we attempted to compute them (the cutoffs are unavailable when they are unable to detect hypothetically misspecified models with sensitivity ≥ 50%). As a result, we relied on the traditional guidelines for a consistent approach across models. We focused on differences in fit between models, but because no models fully met the traditional guidelines, we did not test for fit differences (given that the chi-square difference test is not appropriate when neither model in a pair of nested models shows acceptable fit; Brown, 2015, p. 145).

We also examined whether the magnitude and direction of each model’s parameter estimates (factor loadings, focusing on those ≥ 0.30, and any factor correlations) supported its structure (Bornovalova et al., 2020), especially given that bifactor models tend to overfit data (Bonifay et al., 2017; see Section SA2.2). We attempted to identify one optimal CFA model for a given set of items, based primarily on its parameter estimates and secondarily on its fit, while recognizing that poor-fitting models can yield biased estimates (Brown, 2015, p. 97).

Initial Exploratory Factor Analyses

After CFA did not support the initial correlated factor, bifactor, or higher-order models (see below), we used exploratory factor analysis (EFA) to discover the factor structure and propose a new correlated factor model. We identified the potential numbers of factors by (a) inspecting a scree plot of eigenvalues of the correlation matrix and (b) using "parallel analysis" based on principal components ("PA-PCA-m" in Lim & Jahng, 2019) using the `fa.parallel` function of the `psych` package (ver. 2.4.1; Revelle, 2023). For each potential number of factors, we used `lavaan`'s `efa` function to run EFA with an oblique rotation (oblimin; used by Flora & Flake, 2017) and `WLSMV` estimation with pairwise deletion. As sensitivity analyses, we also used other oblique rotations (geomin, promax) given rotational indeterminacy (Flora & Flake, 2017).

We then inspected the results for each potential number of factors and decided which number of factors to retain, considering model fit (adjusted chi-square; robust RMSEA and CFI) and theory-informed interpretation of the factor loadings (focusing on those ≥ 0.30) and factor correlations. We also decided whether to exclude item(s) to improve the structure, considering theory, consistently small loadings, and cross-loadings onto multiple factors (Rosellini & Brown, 2021). We also explored whether a given scenario's items were poorly behaved in general by computing the mean uniqueness of each scenario's items and analyzing whether any scenario's mean uniqueness was outlying (i.e., > 2.5 median absolute deviations above the median; see Leys et al., 2013). Although one scenario (Elevator) had an outlying mean uniqueness for some factor solutions, only one of its items (Item 1d; Table SA1) had low loadings. After excluding this item (see below), two scenarios (Elevator, Shopping) had outlying mean uniquenesses for a few solutions, but because their items loaded at least fairly strongly (i.e., > 0.45) on expected factors with no cross-loadings, these scenarios' items were retained.

After excluding each item, we re-evaluated the number of potential factors (i.e., via a scree plot and parallel analysis), reran EFA for each number of potential factors, and re-inspected the results to decide which factors to retain and whether to exclude another item.

Revised Confirmatory and Exploratory Factor Analyses

28 Threat and Nonthreat Items. After EFA suggested a correlated factor model with 28 items and three factors (positive threat, negative threat, nonthreat), we fit a series of exploratory models in the CFA framework based on these 28 items. One model had these three correlated factors (Model 5; Figure SA6), and another had two correlated factors (negative threat vs. all other items; Model 6; Figure SA7). Cross-loadings and error covariances were fixed to 0 in these models. In addition to evaluating the overall model fit and the factor loading pattern of each exploratory CFA model, we inspected significant standardized residuals (> 1.96) to evaluate localized areas of misfit and significant modification indices (> 3.84) to determine whether freely estimating certain fixed parameters would meaningfully improve the model (see Brown, 2015, pp. 99 and 102). We also ran an EFA with three correlated factors in the CFA framework (Model 4; Figure SA4; Brown, 2015, p. 168) to explore the potential need to specify correlated errors (e.g., due to method effects consistent with theory). This model freely estimated all loadings and cross-loadings, except for the cross-loadings of one item per factor that was chosen as the factor's anchor item (based on showing a high loading on the factor and low cross-loadings across rotations in the prior EFA analyses). We then fit CFA models with three correlated factors and correlated errors among items stemming from the same scenario to control for potential *item context effects*, which is covariation among items due to their placement or grouping on a survey (Models 7-8; Figures 8-9; see Podsakoff et al., 2024). We also attempted to fit CFA models with two correlated substantive factors (threat, nonthreat) and method factors (or correlated errors) among items with positive versus negative wording to control for *item characteristic effects*, which is covariation among items due to their wording or other properties (Podsakoff et al.; for details, see Section SA1.1). *Method factors* (latent variables capturing method covariance among certain indicators) and *correlated errors* (allowing the uniquenesses of certain items to covary to capture their method covariance) are two ways of accounting for *method effects*, or covariation among items attributable to a shared measurement approach (Brown, 2015, pp. 186-205).

All 18 Threat Items. After most of the 28-item CFA models with method effects yielded improper solutions, we focused remaining analyses on the 18 threat items given researchers' typical focus on scoring RR's positive and negative threat items and the importance of assessing for potential method effects due to positive versus negative wording, especially for the threat items, which for some scenarios (e.g., Blood Test) appear to directly negate each other (e.g., Item 9d is "And you know that you can handle your anxiety while you wait," and Item 9c is "And you think that you will not be able to stand your anxiety while you wait"). When ignored, such effects can suppress correlations between factors with positively versus negatively worded indicators and lead to false inferences of two substantive factors instead of one (e.g., see Green et al., 1993, and Marsh, 1996).[2] We included all threat items in subsequent analyses given the importance of a balanced number of positively and negatively worded items when modeling method effects (Marsh, 1996, p. 817).

First, we conducted a scree test, a parallel analysis, and EFAs (using the three rotations above). After EFA suggested a correlated factor model with 18 items and two factors (positive threat, negative threat), we fit a series of exploratory models in the CFA framework based on these 18 items, including a CFA model with these two correlated factors (Model 10; Figure

[2] Alternatively, unmodeled method effects can inflate factor correlations (see Brown, 2015, pp. 202-203).

SA11), an EFA in the CFA framework with these factors (Model 9; Figure SA9), and a CFA model with these factors and correlated errors among items from the same scenario (Model 11; Figure 2). We also fit CFA models with one substantive threat factor and correlated errors to control for scenario (Model 12; Figure SA12) and, further, with two correlated method factors (orthogonal to the substantive factor) to control for positive versus negative wording (Model 13; Figure SA13; Morin et al., 2020, p. 1055).[3] We then fit a bifactor model with one substantive general threat factor and two substantive specific factors (positive threat, negative threat; orthogonal to each other and to the general factor; Model 16; Figure SA16; e.g., Figure 8.2 of Brown, 2015, p. 302).[4] Finally, we tried to fit a higher-order model with one substantive second-order threat factor and two substantive first-order factors (positive threat, negative threat; orthogonal to each other; with the second-order factor loadings set equal to yield a just-identified solution given only two first-order factors; Model 17, Figure SA17; see Brown, 2015, p. 292).

Power

Our large baseline sample nearly meets or exceeds a conservative sample size guideline for CFA (e.g., an *N* participants/*q* free parameters ratio of 10; Kyriazos, 2018) for the initial correlated factor (9.60) and higher-order (9.73) models using all 36 items (Models 1 and 3).

---

[3] Although Model 13, a single-trait correlated-methods model, was identified, such models tend to yield improper solutions because both the trait factor and the correlated methods factors attempt to explain variance shared among all items, and this competition often leads to nonconvergence (Morin et al., 2020). As sensitivity analyses, we also fit Models 14-15 (Figures SA14-15), in which the positive or negative method factor was removed. In such models (recommended by Morin et al.), the trait factor is "anchored" in a *referent method*, whose method factor is removed. As a result, the trait factor explains both (a) construct-relevant variance shared among all items and (b) method variance shared among the items that correspond to the referent method (i.e., the items whose method factor was removed). We also tried to fit similar models with correlated errors for such item wording (see Section SA1.2).

[4] The bifactor model differs from the single-trait correlated-methods model only by fixing the correlation between the positive/negative (specific) factors to 0. In the single-trait correlated-methods model, the correlation between the positive/negative (method) factors is freely estimated. In the bifactor model, the specific factors are considered substantive constructs, equal in importance to the substantive general factor; making the specific factors orthogonal enables partitioning of variance explained by each of the specific and general constructs. By contrast, in the single-trait correlated-methods model, the method factors are not considered substantive constructs; freely estimating their correlation enables the trait factor to explain all construct-relevant variance (Morin et al., 2020, pp. 1055-1056).

Although we do not meet the ratio for the initial bifactor model (6.87; Model 2), the ratio is not strict and others have proposed a ratio of 5 (Kyriazos, 2018), which we exceed. We also exceed conservative guidelines based on *N* participants (> 500 is “very good”) and on *N* participants/*p* items (10 is “widely accepted”; our lowest is 20.81; Kyriazos, 2018). Our sample size is also sufficient for stable EFA results even if the analyses revealed only a few items per factor and low loadings (*N* participants > 300; Kyriazos, 2018).

Internal Consistency

For each assumed and optimized CFA model, we computed model-based estimates of the *reliability* (or *internal consistency*) of unit-weighted composite (total or mean) scores, which is the proportion of variability in scores attributable to true variability (vs. measurement error). For correlated factor models, we computed categorical omega (ω_{cat}) for each factor’s total score with the `compRelSEM` function of the `semTools` package (ver. 0.5-6; Jorgensen et al., 2022). Following Flora (2020), we estimated the observed total score variance (denominator) from the model-implied variance (vs. observed sample variance) to account for the model’s correlated errors.

For the assumed bifactor model, we used `compRelSEM` to compute categorical omega-hierarchical ($\omega_{h\text{-}cat}$) for each general factor’s total score and categorical omega-hierarchical-subscale ($\omega_{h\text{-}ss\text{-}cat}$; i.e., the proportion of variance of the total score for a specific factor explained by that factor, independent from the general factors). We did not compute internal consistency for the assumed higher-order model given that it yielded an improper solution (see below). We did not compute CIs because our planned method (`ci.reliability` function of `MBESS` package; ver. 4.9.3; Kelley, 2023) is currently unable to account for correlated errors.

Omega coefficients are superior to Cronbach’s alpha because alpha makes restrictive assumptions about a scale’s properties (constant item variances; the *essentially tau-equivalent* *model*) that are often unrealistic in psychology, yielding biased reliability estimates, whereas omega (which assumes the *congeneric model*) does not; omega values greater than .70 are recommended (see Dunn et al., 2014).

Factor Determinacy

Given that multiple sets of factor scores can be computed based on a given set of factor loadings, making the factor scores indeterminate (Grice, 2001), we evaluated the *determinacy* of factor scores. Factor score determinacy indicates how well individual differences on the factor scores represent actual individual differences on the latent variable (Rodriguez et al., 2016). For each assumed and optimized CFA model, we computed (a) the correlation of factor scores with their respective factors (ρ) and (b) the minimum possible correlation between two sets of factor scores ($2\rho^2 - 1$) per Rodriguez et al. Determinacy values greater than .90 and minimum possible correlation values greater than .70 are recommended (see Rodriguez et al.). These results will inform the potential use of factor scores as manifest interpretation bias nodes in the network models for Study 2 and the potential use of factor scores in other studies.

Construct Reliability

To evaluate how well a factor's items represent the factor in a latent measurement model (which affects the factor's replicability across studies), we computed the *construct reliability* (or *construct replicability* or *maximal reliability*) for each factor in the assumed and optimized CFA models per Rodriguez et al. (2016). Construct reliability represents the proportion of variability in the latent variable explained by its indicators. Whereas internal consistency is the correlation between values of the latent variable and unit-weighted composite scores, construct reliability (*H*) is the correlation between values of the latent variable and optimally weighted composite scores (based on factor loadings; Rodriguez et al.). Values greater than .70 are recommended

(Rodriguez et al.). These results will inform the potential use of latent interpretation bias nodes in the network models of Study 2 and the use of latent measurement models in other studies.

Missing Data Handling

After excluding ITT participants with no item-level RR data at baseline (e.g., due in part to a server error), as stated above we used pairwise deletion to handle missing item-level data in ordinal CFA and EFA models. Given that internal consistency, factor score determinacy, and reliability of optimally weighted composites are computed from estimated models, this missing data handling method also applies to these metrics. Given the low rate of item-level missingness (0.3% for baseline RR), we assumed data were missing completely at random (MCAR) and did not pursue multiple imputation or search for auxiliary variables that correlate with item-level missingness and item values.

Results for Study 1

Initial Confirmatory Factor Analyses

All 36 Threat and Nonthreat Items

The initial model with four correlated factors (positive threat, positive nonthreat, negative threat, negative nonthreat) based on all 36 items (Model 1) had unacceptable fit (SRMR = .091, RMSEA = 0.098, CFI = .726, TLI = 0.707) per traditional guidelines (SRMR near or < .08; RMSEA near or < 0.06; CFI and TLI near or > .95). The initial bifactor model with two correlated general factors (positive, negative), each with two specific factors (threat, nonthreat), also had unacceptable fit (Model 2; SRMR = .090, RMSEA = 0.098, CFI = .743, TLI = 0.709; Table 2). The initial higher-order model yielded an improper solution (Model 3; standard errors and fit could not be computed, likely due to specifying a model unsupported by the data).

As shown in Figure SA1, although the four correlated factors were well defined by their factor loadings, the positive nonthreat and negative nonthreat factors were highly correlated, $r = .87$, $p < .001$, suggesting that they lack discriminant validity and may be collapsed (when $r \geq .85$ per Brown, 2015, p. 146). Given that these two nonthreat factors were more highly correlated than the two positive factors ($r = .78$, $p < .001$) and the two negative factors ($r = .33$, $p < .001$), the higher-order model was likely misspecified, as this pattern of correlations conflicts with the posited positive and negative second-order factors (see Brown, 2015, p. 290). Moreover, in the bifactor model, although the positive general factor was fairly well defined by its loadings, the negative general factor was not (Figure SA2; see Morin et al., 2020, p. 1060). This suggests that combining the threat and nonthreat items for each valence when scoring is not advisable.

Initial Exploratory Factor Analyses

The oblimin, geomin, and promax rotations yielded similar results (especially for EFA solutions with two or three factors, which were our focus given that they were more interpretable than solutions with four or more factors); thus, we focus on the oblimin results. Fit statistics and factor loadings for all EFA models are in Tables 3 and SA2, respectively.

All 36 Threat and Nonthreat Items

For all 36 items, inspection of the scree plot revealed a breakpoint between the cliff and the scree at 4 potential factors; parallel analysis revealed that 5 eigenvalues exceeded the mean, suggesting up to 5 factors (see Montoya & Edwards, 2021). Given that the upper limit can be ±1 that suggested by parallel analysis (Lim & Jahng, 2019) and our expectation of more than one RR dimension, we specified two-to-six factors in initial EFAs.

Across all solutions, the nonthreat items (positive and negative) loaded onto one factor, and the negative threat items loaded onto other factor(s). In the two-factor solution, the positive threat items tended to load with the nonthreat items onto one factor, and the negative threat items tended to load onto another factor. In the three-factor solution, the positive threat items, negative threat items, and nonthreat items tended to load onto three different factors. In the four-factor solution, the negative threat items and nonthreat items tended to load onto two different factors, and the positive threat items began to separate in loading onto multiple factors. In the five- and six-factor solutions, the negative threat items also began to separate. Given that the solutions with four-to-six factors were less interpretable and that we sought to define both positive threat and negative threat factors, we focused on improving the three-factor solution (positive threat, negative threat, nonthreat) by excluding items one at a time (Rosellini & Brown, 2021).

28-35 Threat and Nonthreat Items

We ultimately excluded a total of 8 items. Parallel analysis after each of the first 6 item exclusions suggested up to four factors (leading us to consider two-to-five factors for EFAs of 30-35 items), and parallel analysis after each of the next two exclusions suggested up to three factors (leading us to consider two-to-four factors for EFAs of 28-29 items). First, from 36 items we excluded a positive threat item (Item 2a) with consistently small loadings across the two-to-four factor solutions. Second, from 35 items we excluded a negative threat item (Item 7a) that consistently cross-loaded onto nonthreat and positive threat factors. Third, from 34 items we excluded three positive threat items in turn that consistently had small loadings on the positive threat factor and loaded onto either the negative threat factor (Item 1d) or the nonthreat factor (Items 3a and 5c). Finally, from 31 items we excluded three positive nonthreat items (Items 7d, 6d, and 9a) that consistently cross-loaded onto the positive threat factor.

The resulting three-factor solution based on 28 items had simple structure, and the positive threat (5 items), negative threat (8 items), and nonthreat (15 items) factors explained

7.8%, 12.3%, and 27.4% of the variance, respectively (47.5% in total). The positive threat factor significantly correlated with the negative threat (r = -.19, p < .001) and nonthreat (r = .18, p < .001) factors. The negative threat and nonthreat factors did not significantly correlate (r = .05, p = .106).[5] Model fit generally improved after each item exclusion (Table SA3); however, although the 28-item three-factor solution (RMSEA = 0.078, CFI = .873) fit better than the 36-item solution (RMSEA = 0.089, CFI = .801), the 28-item fit was still unacceptable per traditional guidelines (RMSEA near or < 0.06; CFI near or > .95).

Revised Confirmatory and Exploratory Factor Analyses

28 Threat and Nonthreat Items

Confirmatory Factor Analyses. Based on the three-factor EFA solution with 28 items, a revised CFA model with three correlated factors (positive threat, negative threat, nonthreat; Model 5; Figure SA6) had mixed evidence of fit per traditional guidelines (SRMR near or < .08; RMSEA near or < 0.06; CFI and TLI near or > .95): SRMR (.085) met these guidelines, but other indices did not (RMSEA = 0.092, CFI = .794, TLI = 0.776; Table 3). However, a CFA model with two correlated factors (negative threat, no negative threat [positive threat and nonthreat items]; Model 6; Figure SA7) fit worse (SRMR = .099, RMSEA = 0.103, CFI = .743, TLI = 0.722); this model was fit based on the two-factor EFA solution with 28 items, which had nearly simple structure, as an incorrect number of factors can yield misfit (Brown, 2015, p. 140).

Inspection of significant (p < .05) standardized residuals for an EFA with the three correlated factors in the CFA framework (Model 4; Figure SA4) revealed that the most extreme

[5] For the geomin rotation, the positive threat factor significantly correlated with the negative threat factor (r = -.36, p < .001) but did not significantly correlate with the nonthreat factor (r = .11, p = .069). The correlation between the negative threat and nonthreat factors remained nonsignificant (r = -.03, p = .491). For the promax rotation, for which significance tests of factor correlations are unavailable, the positive threat-negative threat, positive threat-nonthreat, and negative threat-nonthreat correlations were -.30, .34, and .04, respectively.

residuals (-5.71 to -9.69) were for the relations between the four positive threat and negative threat items from the same scenario (Figure SA5). Modification indices for this kind of model are limited to error correlations (another source of misfit, Brown, 2015, p. 157). Inspection of significant modification indices revealed that correlated errors for these four item pairs had the three highest (and sixth highest) modification indices (13.95-60.01), with medium-to-large completely standardized expected parameter change (SEPC) values (*r*s = .28-.51; Table SA4). These correlated errors suggest item context effects due to scenarios (see Podsakoff et al., 2024). The model revealed no salient cross-loadings.

A CFA model with the three correlated factors (positive threat, negative threat, nonthreat) and correlated errors between the four positive threat and negative threat items from the same scenario (Model 7; Figure SA8) fit better than the three correlated-factor CFA model without the correlated errors, but fit remained mixed (SRMR = .082, RMSEA = 0.083, CFI = .836, TLI = 0.819; Table 3). A CFA model with these correlated factors and correlated errors among any items (threat or nonthreat) from the same scenario (Model 8; Figure 1) also had better, but still mixed, fit (SRMR = .075, RMSEA = 0.082, CFI = .850, TLI = 0.821). In each CFA model with three correlated factors—without correlated errors per scenario or with such correlated errors among threat items or any items—the factors were well defined, the positive threat and negative threat factors significantly correlated ($r = -.23$, $p < .001$; $r = -.13$, $p = .009$; $r = -.13$, $p = .009$; respectively), and the nonthreat factor significantly correlated with the positive threat factor (*r*s = .40, *p*s < .001) and negative threat factor (*r*s = .11-.12, highest $p = .010$).

Given that positive versus negative item wording often yields item characteristic effects (Podsakoff et al., 2024), we also attempted to fit various models with two correlated substantive factors (threat, nonthreat) and with method factors (or correlated errors when method factors yielded improper solutions) among items with positive versus negative wording (Models 18-22), but most of these models yielded improper solutions (for details, see Section SA1.1), prompting us to simplify the models by focusing on all 18 threat items.

All 18 Threat Items

Exploratory Factor Analyses. The oblimin, geomin, and promax rotations based on all 18 threat items yielded similar results for solutions with one or two factors, which were our focus given that they were more interpretable than solutions with additional factors. We again focus on the oblimin results. The scree test suggested up to five factors, and parallel analysis suggested up to four; given the possibility of one substantive threat factor with method effects, we considered one-to-five factors. In the one-factor solution, all positive threat and negative threat items loaded onto the factor, with positive and negative loadings, respectively. In the two-factor solution, the positive and negative items loaded onto separate factors, with simple structure. In solutions with three-to-five factors, the positive and negative items tended to load onto separate factors (one or more), with some positive and negative items from the same scenario loading or cross-loading onto the same factor, but with opposite signs. Given that the solutions with one or two factors were well defined by all items, we did not exclude any items.

The one-factor solution explained 23.2% of the variance. In the two-factor solution, the positive threat and negative threat factors, which significantly correlated (r = -.22, p < .001),[6] explained 14.4% and 20.4% of the variance, respectively (34.8% in total). Both models had poor fit (RMSEA = 0.165, CFI = .425; RMSEA = 0.138, CFI = .651; respectively; Table SA3) per traditional guidelines (RMSEA near or < 0.06; CFI near or > .95).

[6] The factors were also significantly correlated for the geomin rotation (r = -.26, p < .001), with a similar magnitude and direction for the promax rotation (r = -.29).

Confirmatory Factor Analyses. Based on the two-factor EFA solution with 18 threat items, a CFA model with two correlated factors (positive threat, negative threat; Model 10; Figure SA11) had unacceptable fit (SRMR = .091, RMSEA = 0.132, CFI = .637, TLI = 0.586; Table 4) per traditional guidelines (SRMR near or < .08; RMSEA near or < 0.06; CFI and TLI near or > .95). The significant standardized residuals and modification indices for an EFA with the two correlated factors in the CFA framework (Model 9; Figure SA9) again revealed that the most extreme residuals (or nearly so; -6.65 to -13.07) were for relations between the nine positive and negative items from the same scenario (Figure SA10) and that error correlations for these pairs had the seven highest (and 9th and 12th highest) modification indices (19.98-114.19) with medium-to-large SEPC values (*r*s = .33-.62; Table SA5). The model revealed no salient cross-loadings.

A CFA model with the two correlated factors (positive threat, negative threat) and correlated errors between positive and negative items from the same scenario (Model 11; Figure 2) fit better than the two-factor CFA model without the correlated errors, but fit was mixed per traditional guidelines (SRMR near or < .08; RMSEA near or < 0.06; CFI and TLI near or > .95), with only SRMR (.067) meeting these guidelines (RMSEA = 0.081, CFI = .873, TLI = 0.844; Table 5). However, a CFA model with one factor (based on the one-factor EFA solution) and the correlated errors (Model 12; Figure SA12) fit worse (SRMR = .135, RMSEA = 0.159, CFI = .500, TLI = 0.393). In both CFA models with two correlated factors—without correlated errors per scenario or with them—the positive threat and negative threat factors were well defined and significantly correlated ($r = -.29$, $p < .001$; $r = -.17$, $p < .001$; respectively).

Although a CFA model with one substantive threat factor, correlated errors per scenario, and two correlated method factors to control for positive versus negative wording (Model 13;

Figure SA13) had better, but still mixed, fit than the two correlated-factor model with correlated errors per scenario (SRMR = .046, RMSEA = 0.066, CFI = .928, TLI = 0.896), the substantive threat factor was not well defined by its loadings. The substantive threat factor was also not well defined when retaining the method factor for only the positive items (Model 14; Figure SA14) or for only the negative items (Model 15; Figure SA15).[7] This suggests that the positive threat and negative threat items are not positively versus negatively worded indicators of one substantive threat factor, but indicators of two substantive factors (positive threat, negative threat).

Finally, a bifactor model with one substantive general threat factor, two substantive specific factors (positive threat, negative threat), and correlated errors per scenario (Model 16; Figure SA16) also had better, but still mixed, fit than the two correlated-factor model with correlated errors per scenario (SRMR = .049, RMSEA = 0.066, CFI = .926, TLI = 0.894), but the general factor was not well defined by its loadings. Moreover, a higher-order model with one substantive second-order threat factor, two substantive first-order factors (positive threat, negative threat), and correlated errors per scenario (Model 17, Figure SA17) yielded an improper solution. These results further suggest that combining the positive threat and negative threat items, each set of which defines a substantive factor, into a single total score is not advisable.

Internal Consistency, Factor Determinacy, and Construct Reliability

Initial Confirmatory Factor Models

All 36 Threat and Nonthreat Items. For the initial model with four correlated factors (positive threat, positive nonthreat, negative threat, negative nonthreat) based on all 36 threat and nonthreat items (Model 1, Figure SA1), the internal consistency of each factor's unit-weighted composite score (ω_{cat}s = .74, .85, .81, and .87, respectively), the correlation of each factor with

[7] Models with correlated errors for positive vs. negative wording had similar results (Models 23-27; Section SA1.2).

its factor scores (ρs = .92, .97, .93, and .97), the minimum possible correlation between two sets of factor scores for each factor ($2\rho^2 - 1$ values = .70, .89, .74, and .89), and the construct reliability for each factor (*H*s = .81, .91, .86, and .92) were acceptable.

For the initial bifactor model with two correlated general factors (positive, negative), each with two specific factors (threat, nonthreat; Model 2, Figure SA2), the internal consistency of the positive general factor's unit-weighted composite score was acceptable ($\omega_{h\text{-}cat}$ = .74), but the internal consistency of the negative general factor's unit-weighted composite score was not ($\omega_{h\text{-}cat}$ = .44), nor were the internal consistencies of most of the specific factors ($\omega_{h\text{-}ss\text{-}cat}$s = .43, .01, .82, and .01, respectively). The correlation of each general factor with its factor scores (ρs = .97), the minimum possible correlation between two sets of factor scores for each general factor ($2\rho^2 - 1$ values = .87 and .86), and the construct reliability for each general factor (*H*s = .92 and .93) were acceptable. However, for most of the specific factors, the correlation of the specific factor with its factor scores (ρs = .84, .71, .93, and .66), the minimum possible correlation between two sets of factor scores for the specific factor ($2\rho^2 - 1$ values = .41, .01, .73, and -.12), and the construct reliability for the specific factor (*H*s = .69, .30, .86, and .29) were unacceptable.

Given that the initial higher-order model (Model 3, Figure SA3) yielded an improper solution, we did not compute its internal consistency, factor determinacy, or construct reliability.

Revised Confirmatory Factor Models

28 Threat and Nonthreat Items. For the CFA model with three correlated factors (positive threat, negative threat, nonthreat) and correlated errors among any items from the same scenario (Model 8; Figure 1), the internal consistencies of the unit-weighted composite scores were acceptable for the negative threat and nonthreat factors (ω_{cat}s = .80 and .90, respectively) and nearly acceptable for the positive threat factor (ω_{cat} = .69). For the negative threat and nonthreat factors, the correlation of the factor with its factor scores (ρs = .93 and .97) and the minimum possible correlation between two sets of factor scores for the factor ($2\rho^2 - 1$ values = .75 and .89) were acceptable. Although the correlation of the positive threat factor with its factor scores was nearly acceptable (ρ = .88), the minimum possible correlation between two sets of factor scores for this factor was unacceptable ($2\rho^2 - 1 = .56$). The construct reliabilities of the positive threat, negative threat, and nonthreat factors were acceptable (*H*s = .75, .84, and .94).[8]

All 18 Threat Items. For the CFA model with two correlated factors (positive threat, negative threat) and correlated errors per scenario based on all 18 threat items (Model 11, Figure 2), the internal consistency of each factor's unit-weighted composite score was acceptable (ω_{cat}s = .74 and .82, respectively). The correlation of the negative threat factor with its factor scores (ρ = .93) and the minimum possible correlation between two sets of factor scores for this factor ($2\rho^2 - 1 = .72$) were acceptable. Although the correlation of the positive threat factor with its factor scores was nearly acceptable (ρ = .89), the minimum possible correlation between two sets of factor scores for this factor was unacceptable ($2\rho^2 - 1 = .60$). The construct reliabilities for the positive threat and negative threat factors were acceptable (*H*s = .79 and .86, respectively).

Discussion for Study 1

The present study tested and optimized the structural validity of our team's RR measure of interpretation biases using baseline data from a web-based CBM-I trial run in anxious adults. After initial CFA models inferred from the literature (e.g., four correlated factors: positive threat, negative threat, positive nonthreat, negative nonthreat) were unsupported, EFA suggested three factors (positive threat, negative threat, nonthreat) and potential items to exclude. Including

[8] For the CFA model with three correlated factors (positive threat, negative threat, nonthreat) and correlated errors among threat items from the same scenario (Model 7; Figure SA8), results were nearly identical (ω_{cat}s = .69, .80, and .90, respectively; ρs = .88, .92, and .97; $2\rho^2 - 1$ values = .55, .70, and .89; and *H*s = .74, .84, and .94).

correlated errors among items from the same scenario, exploratory CFA models with (a) these three factors based on a subset of 28 threat and nonthreat items or (b) two factors (positive threat, negative threat) based on all 18 threat items had well-defined factors with generally acceptable internal consistency and construct reliability. Although factor determinacy was acceptable for the negative threat and nonthreat factors, it was mixed for positive threat. Overall CFA model fit was mixed. Exploratory CFA models of the threat items supported the positive threat and negative threat factors as two distinct constructs (i.e., not merely method artifacts due to item wording).

Distinct Threat Versus Nonthreat Factors

Although the four correlated factors implied by the RR development paper (Mathews & Mackintosh, 2000) showed acceptable internal consistency (and factor determinacy and construct reliability) in our data, the high correlation between the positive nonthreat and negative nonthreat factors in the initial CFA suggested that they lack discriminant validity and could be collapsed, which highlights the importance of evaluating factor structure. By contrast, the negative threat and negative nonthreat factors were only moderately correlated in this initial correlated factor model (likely leading to the inestimable initial higher-order model), and the negative general factor was not well defined (and had poor internal consistency) in the initial bifactor model. These results suggest that combining all negative items when scoring (as in Baee et al., 2024) is not advisable.

After items with low loadings and salient cross-loadings were removed during EFA, CFA models based on 28 threat and nonthreat items found that three factors (positive threat, negative threat, nonthreat) were well defined and correlated in accordance with theory (e.g., significant negative correlation between positive threat and negative threat; stronger positive correlation of nonthreat with positive threat than with negative threat). These CFA models also revealed the need to control for item context effects. Including correlated errors among items from the same scenario improved model fit, although fit remained mixed per traditional guidelines. Overall, the finding of distinct (yet correlated) threat versus nonthreat factors supports researchers' use of the threat items as measures of threat-relevant (vs. threat-irrelevant) interpretation biases.

Distinct Positive Threat Versus Negative Threat Factors

In addition to item context effects (e.g., tied to scenarios), it is important to consider item characteristic effects (Podsakoff et al., 2024), especially for positively versus negatively worded factors to rule out the possibility that the two factors are merely method artifacts (i.e., variations of a single substantive threat factor due simply to positive vs. negative wording, rather than two substantive positive threat and negative threat factors). After EFA of all 18 threat items revealed no items to exclude, CFA models controlling for scenario found that two factors (positive threat, negative threat) were well defined and negatively correlated (though with mixed fit) and that one threat factor with positive and negative method factors was not well defined. A general threat factor with positive and negative specific factors was also not well defined (and a higher-order model was inestimable). These results suggest that positive threat and negative threat are distinct constructs, consistent with findings that positive and negative threat factors (indicated by multiple measures) have different relations with criterion variables (e.g., anxiety; Steinman et al., 2020), although such external validity relations have not yet been assessed for RR specifically.

Modeling Recognition Ratings

The positive threat and negative threat factors in the two-factor model based on all 18 threat items (with correlated errors per scenario) had acceptable internal consistencies for their unit-weighted composite scores (ω_{cat}s = .74 and .82, respectively), consistent with those in a prior analysis (without correlated errors) of this sample (αs = .73 and .85; Ji et al., 2021). The factors' internal consistencies were slightly lower in the three-factor model based on 28 threat and nonthreat items (controlling for scenario, ω_{cat}s = .69 and .80, respectively, and .90 for the nonthreat factor). The construct reliabilities of all factors in both models were also acceptable, supporting use of these factors in latent measurement models. Although factor determinacy for the negative threat factor (and the nonthreat factor) was acceptable, in both models it was mixed for the positive threat factor, making the use of factor scores for positive threat questionable.

The question of whether to use the two-factor model based on all 18 threat items or the three-factor model based on 28 threat and nonthreat items depends on the application. In the three-factor model, threat items that cross-loaded onto the nonthreat factor had been removed, helping ensure that the 5-item positive threat and 8-item negative threat factors are distinct from the nonthreat factor. However, the imbalanced number of items between these two threat factors yields imbalances (a) in positively versus negatively worded items, complicating analysis of item wording effects, and (b) in threat domains (3 physical, 1 social, and 1 other for positive threat vs. 2 physical, 3 social, and 3 other for negative threat), complicating factor comparisons. The two-factor model based on all 18 threat items may be preferable in most situations given its balance in threat domains, and it is especially preferable when both factors (or scores derived from them) are used in the same statistical model (e.g., multiple regression, network analysis). Most prior research has used all 18 threat items, analyzing separate unit-weighted composite scores for the 9 positive threat and 9 negative threat items (e.g., Ji et al., 2021; Eberle et al., in press; Hohensee et al., 2020; Larrazabal et al., 2024; Schmitt et al., 2023; Steinman et al., 2020; Silverman et al., 2024; Vela de la Garza Evia et al., 2024); the present results support this approach. However, it may be fruitful to compare results based on both the 28-item three-factor model and the 18-item two-factor model. We encourage researchers to clearly report which items they analyze.

Improving Recognition Ratings Modeling and Measurement

Even when controlling for scenario, model fit for both the two-factor model based on all 18 threat items and the three-factor model based on 28 items was mixed. Although none of the scenarios with outlying mean uniquenesses had multiple items with small loadings, leading us to exclude single (vs. all) items from a given scenario at each step of the EFAs leading to the 28-item model, future factor analyses could consider excluding all items from a scenario if any of its items has relatively low loadings or salient cross-loadings. Excluding entire scenarios would also avoid the imbalances described above and, in models including both threat and nonthreat items, provide enough items per scenario to consider using method factors (vs. correlated errors) to control for scenario. The method factors could also be correlated per threat domain, although the present models' standardized residuals and modification indices did not suggest a clear need to control for this. It might also be fruitful to run factor analyses focused on the nonthreat items (e.g., to assess item wording effects), as we did for all 18 threat items, before revisiting an overall model. Note that when we refer to excluding items or scenarios, we are referring to content that we propose to exclude from analyses to improve the measurement model for the present RR measure, not to content that we propose excluding from the RR measure itself.

However, in addition to further optimizing the measurement model of the present RR measure, it might be possible to improve the measure itself. For example, three kinds of items could be revised. The first is items with relatively low loadings (e.g., positive threat Item 2a, for which it might be difficult to imagine an audience "laugh[ing] in appreciation" during one's speech at a wedding reception). The second is items with unexpected primary loadings (e.g., positive threat Item 5c, for which getting a call from a friend running late might not be positive enough with respect to the threat that she does not want to see you, leading to a primary loading on nonthreat in the three-factor EFA). The third is items with salient cross-loadings (e.g., positive nonthreat Item 7d, for which “but” in “but you are happy you are getting exercise” might partially negate the threat of a painful scrape, leading to a cross-loading onto positive threat in the three-factor EFA). Moreover, additional scenarios and items for each threat domain could be given and factor analyzed (Rosellini & Brown, 2021). The present RR measure consists of the 9 scenarios (of 18 piloted) whose negative threat items most highly correlated with an anxiety criterion, which might help explain why the negative threat factor had higher internal consistency and factor determinacy than the positive threat factor. Optimizing a measure per multiple kinds of validity is recommended (see Hussey & Hughes, 2020). Finally, increasing the number of response options may reduce measurement error (see Rosellini & Brown, 2021); for 26 of the 36 items, the most extreme response option was the most frequent option endorsed.[9]

Future Directions

Given that the three-factor CFA model based on 28 threat and nonthreat items and the two-factor CFA model based on all 18 threat items (each with correlated errors per scenario) were fit after the initial CFAs and EFAs, confirmatory tests of these models in an independent sample are an important next step (which we have already preregistered; https://doi.org/mt9r). In addition to further assessing RR’s external validity, future research should evaluate other aspects of structural validity, specifically test-retest reliability (dependability and stability; see Revelle & Condon, 2019; Watson, 2004) and measurement invariance (also already in our preregistration based on age, gender, condition, time point, and the Condition × Time interaction). It would also be interesting to analyze how responses to items might change over the course of completing the RR items, as imagining potential threats when reading the scenarios and then recalling the

[9] For item distributions, see the “results/efa/hists/” folder at https://osf.io/ebn25/.

scenarios' meanings when rating the items may influence participants' ratings.

Finally, given different versions of RR stimuli used across studies (Duken et al., 2024), with psychometric properties (e.g., internal consistency) ranging from poor (e.g., Duken et al., 2024; Reuland & Teachman, 2014) to questionable (e.g., Edwards et al., 2018) to acceptable (for RR in this study), and properties unreported in most studies, we encourage researchers to report the RR stimuli they use and to evaluate their psychometric properties. Notably, we emphasize the need to evaluate more than internal consistency, which assumes unidimensionality and does not guarantee other aspects of structural validity (i.e., factor structure, measurement invariance, test-retest reliability; Hussey & Hughes, 2020), some of which will require large samples to evaluate.

Limitations and Conclusion

The present study has a few limitations. First, although we believe our assumption that such data MCAR is reasonable given our low missing data rate, it is generally preferable to use PML estimation or multiple imputation with ordinal data, which assume the data are missing at random (MAR). Second, our sample was mostly female, White, and not Hispanic or Latino, limiting our results' generalizability. Third, no model fully met traditional fit guidelines, leading us to compare models descriptively rather than using statistical tests to compare nested models.

Claims about the causal effects of interpretation biases on anxiety and about CBM-I's effects on such biases depend on the reliability and validity of interpretation bias measures. The present study provides initial evidence of the factor structure of the RR measure developed by our team and supports its use in assessing positive and negative threat-relevant interpretations as two substantive constructs distinct from one another and from threat-irrelevant interpretations. We noted two approaches that align with this structure. First, the 5-item positive threat and 8-item negative threat factors from our three-factor model can be used when the priority is to ensure that the threat factors are distinct from the nonthreat factor. Second, given complications that arise from the imbalanced number of positive threat versus negative threat items in the three-factor model, the 9-item positive threat and 9-item negative threat factors from our two-factor model may be preferable in many situations. The present study also highlights the importance of assessing structural validity from multiple angles and of modeling method effects. However, ongoing construct validation to further evaluate structural and external validity, to improve CFA model fit and internal consistency (especially for positive threat), and to improve the RR measure itself will maximize its use as a primary measure of target engagement.

Implications of Study 1 Results for Study 2

As stated above, Study 1 aimed to test and, if needed, improve the measurement models for RR for use in Study 2 and in other anxious samples. We planned to use the model fit and factor loading pattern of the best CFA model found in Study 1 to inform whether this model can reasonably be used to model interpretation bias in Study 2. Specifically, we planned to require that model fit meet traditional guidelines for two indices (i.e., SRMR and any one of RMSEA, CFI, or TLI). Second, we required that the factor loading pattern be interpretable and reasonably accord with theory (at minimum, presence of a factor for positive threat or negative threat, with at least three indicators per factor). If these criteria were met, then we planned to use internal consistency, factor determinacy, and construct reliability for the best model in Study 1 to inform how bias would be modeled in Study 2 (i.e., unit-weighted composite scores, factor scores, or latent measurement model).

The fit of the CFA model with two correlated factors (positive threat, negative threat) and correlated errors per scenario based on all 18 threat items in Study 1 met traditional guidelines for only SRMR. However, we decided to model positive threat and negative threat per this model in Study 2 for two reasons. First, the traditional guidelines likely lack generalizability; even after correction for DWLS estimation with categorical data, the traditional guidelines are based on one specific model (see McNeish & Wolf, 2024). Second, our model had interpretable factor loadings and factor correlations, and the internal consistencies and construct reliabilities were acceptable for both factors. We will model these two constructs in Study 2 using unit-weighted composite scores as manifest nodes. We will use the 9 positive threat and 9 negative threat items from the two-factor model in Study 1 given their balance in threat domains (rather than the unbalanced subset of threat items from the three-factor model) and our plan to include both positive and negative biases in the same networks.

Study 2: Network Analysis of Change Mechanisms

In addition to investigating and improving the psychometric properties of interpretation bias measures to facilitate research on cognitive mechanisms of change in anxiety (Study 1), this dissertation uses network modeling to investigate the effects of CBM-I on interpretation biases, individual anxiety symptoms, and anxiety-related impairment, and on the bias-symptom-impairment relations (Study 2). Network theory holds promise for studying change mechanisms by viewing psychopathology as the causal relations among symptoms and studying the effects of treatments on symptoms and their relations (Borsboom & Cramer, 2013; Hofmann et al., 2020). This contrasts with the dominant disease model, which views psychopathology as an underlying latent disease that causes the symptoms (explaining their covariation) and which assumes that the symptoms (e.g., anxiety intensity, avoidance) do not causally affect one another (Borsboom & Cramer). By estimating *partial correlations*, or unique pairwise relations among variables that control for all the other variables in the model (vs. *marginal* or *zero-order* correlations that do not control for other variables), network models reveal conditional (in)dependencies that can suggest potential causal relations among variables (Haslbeck et al., 2022a; van Bork et al., 2024; albeit with caution: Ryan et al., 2022).

Interpretation Bias and Extended Psychopathology Networks

Network models have been used to study not only the relations among symptoms (self-reported experiences and behaviors that are part of diagnostic criteria) of specific syndromes, such as *DSM-5* anxiety disorders (Wilshire et al., 2021), but also the relations among symptoms and components of functional impairment (self-reported interference with activities) tied to broad classes of psychopathology, such as transdiagnostic anxiety (e.g., Curtiss et al., 2018). *Extended networks* have also included nodes for theory-derived mechanistic variables, such as attentional threat bias and attentional control (Heeren & McNally, 2016) and coping strategies (Papini et al., 2020), alongside symptom nodes (Hoffart & Johnson, 2020a; Jones et al., 2017). These variables (e.g., components of functional impairment, mechanistic variables) are not necessarily part of diagnostic criteria (van den Hout, 2014) but are likely key to hypotheses about the causal system (Fried & Cramer, 2017) and hence important elements[10] to include in the network (Bringmann et al., 2022). We also recognize that some diagnostic criteria do include mechanistic variables, making a distinction between symptoms and mechanisms conceptually and practically fraught, so we use these terms as heuristics and acknowledge the imprecision. We model nodes for interpretation biases, anxiety symptoms, and related impairment to study the effects of CBM-I on these nodes and their relations in a transdiagnostic sample of anxious adults.

In these models, we adopt the perspective that more densely connected psychopathology networks are more pathological than less densely connected ones, such that activation of a given

[10] *Elements* may be a more appropriate term for symptom (and other) nodes in a network model given that the term *symptoms* implies that the symptom nodes reflect an underlying disease; although hybrid network models allow for the possibility of common latent causes of symptoms in addition to causal relations among symptoms in a network, pure network models (e.g., those used in Study 2) do not include common causes (Fried & Cramer, 2017).

node (e.g., anxiety intensity) by an external stressor causes other nodes (e.g., avoidance) to activate, given their strong relations (Blanchard & Heeren, 2022). Cross-sectional networks have found mixed results for this *connectivity hypothesis*, but within-person time series networks have found more support (Robinaugh et al., 2020). For example, a cross-sectional network of fear and avoidance ratings of social situations had significantly greater global strength for adults with social anxiety disorder (SAD) than that for adults without SAD (Heeren & McNally, 2018). However, a cross-sectional network of anxiety and depression symptoms for adults at discharge from a partial hospital program had significantly *greater* global strength than that at admission, even though symptom severity had significantly *decreased* (Beard et al., 2016). Within-person networks may be more suitable for testing the connectivity hypothesis in that they do not assume ergodicity (that average effects between people are the same as the average within-person effects for a given person; Robinaugh et al., 2020). For example, the temporal within-person networks of negative emotion ratings for adults with major depressive disorder (MDD) estimated from ecological momentary assessment data have shown significantly greater global strength than the networks for adults without MDD (Pe et al., 2015). When mechanistic nodes (e.g., interpretation biases) are included in a network, they are expected to explain some of a given symptom node's variance that otherwise would be explained only by other symptom nodes (Hoffart & Johnson, 2020a). As a result, an extended network that includes mechanistic nodes in addition to symptom nodes may show weaker relations among the symptom nodes than a pure symptom network that does not include mechanistic nodes. Nevertheless, if higher scores on each node indicate a negative affective state or dysfunction for that node (see Borsboom et al., 2021; Bringmann et al., 2023), then stronger relations among such *negative nodes* for mechanistic variables, symptoms, and impairment should constitute greater pathology.

Effects of Interpretation Bias Training on Psychopathology Networks

Based on theorized intervention effects on networks (see Borsboom, 2017; Borsboom et al., 2021; Hayes & Andrews, 2020; van Bork et al., 2024), we expect CBM-I to have two kinds of effects on a network containing nodes for interpretation biases, anxiety symptoms, and impairment. First, CBM-I, a mechanistic treatment, targets interpretation biases, aiming to decrease negative bias and increase positive bias (or decrease a *lack* of positive bias, which we refer to hereafter to model all nodes as negative nodes) by providing practice assigning a benign interpretation and training a positive contingency between ambiguity and nonthreatening outcomes. As a result, CBM-I may deactivate (reduce the mean levels of) a negative bias node and a lack of positive bias node, which may in turn deactivate symptom and impairment nodes given their expected relations with these bias nodes. Thus, CBM-I may shift mean levels of nodes without necessarily changing their relations; such *node intervention* may reduce the activation of bias, symptom, and impairment nodes but leave the system at risk for relapse if a stressor reactivates the network, whose structure is unchanged (Borsboom et al., 2021). Second, CBM-I may change the relations among nodes, thereby changing the network's structure. That is, CBM-I may *destabilize* relations among biases, symptoms, and impairment, reducing network connectivity (Borsboom, 2017) and weakening a pathological attractor (Hayes & Andrews, 2020; Olthof et al., 2023) that keeps the system in a stable anxious state of high activation (which may characterize anxious people seeking CBM-I) and resists its transition into a stable healthy state of lower activation. Such intervention on the network structure may yield a more resilient dynamical landscape (van Bork et al., 2024).

Network Intervention Analysis

A treatment's effects on mean levels of nodes in a network can be tested with *network* *intervention analyses*. Such analyses include treatment condition as a node in a cross-sectional network at each time point to test condition differences on mean levels of each node at the time point while controlling for relations among all other nodes (Blanken et al., 2019; Blanken et al., 2021). Edges between condition and certain nodes (e.g., bias) are the direct effects of condition on mean levels of those nodes, and edges between these nodes and other nodes (e.g., anxiety symptoms) are potential indirect effects—paths via which condition's activation/deactivation of a given node may spread. (We underscore these as *potential* indirect effects given the cross-sectional nature of the networks, in which case causal relations can be inferred only for the direct effects of condition on given node(s), not for the effects of those nodes on other adjacent nodes.) The effects can differ at each time point, revealing the order in which the effects emerge. Given that cognitive models of anxiety posit that cognitive changes precede and mediate symptom changes (Smits et al., 2012), including in CBM-I (e.g., Steinman & Teachman, 2014), CBM-I can be expected to have direct effects on interpretation biases that emerge at earlier time points and indirect effects on anxiety symptoms and impairment through their relations with biases.

Temporal Network Modeling

An intervention's effects on relations among nodes can be studied with various methods (for systematic review, see Schumacher et al., 2022). Temporal network modeling is an attractive option in that it addresses some limitations of cross-sectional models, which conflate between- and within-person relations and lack temporal precedence needed for Granger-causal inference (where nodes precede and predict others). We will use a lag-1 graphical vector autoregression (GVAR) model to disaggregate temporal, contemporaneous, and between-person networks with nodes for interpretation biases, anxiety symptoms, and related impairment in each condition. The temporal network, which models lagged within-person relations from one time point to the next, has temporal precedence, and the contemporaneous network, which models concurrent within-person relations at the same time point, captures relations on shorter time scales (Epskamp et al., 2018).[11] Thus, these two networks permit more precise tests of the cognitive-symptom relations that cognitive models posit within people over time. Further, a multigroup model can be used to compare the networks between conditions and assess whether the positive CBM-I network has less connectivity than the other conditions' networks.

Temporal networks used to require time series data (≥ 20 time points; Jordan et al., 2020) but can now be fit with panel data (≥ 3 time points; Epskamp, 2020). Notably, the GVAR model assumes that the means and (co)variances of nodes are invariant over time (i.e., *stationary*), an assumption made more plausible by transforming the data prior to analysis (see Method below). Thus, we view network intervention analysis and temporal network modeling as complementary approaches, each with strengths and weaknesses. The former tests condition effects on mean levels of nodes and reveals paths of spreading (de)activation time point by time point, allowing network parameters to change over time but conflating within- and between-person relations; the latter tests condition effects on network structures that disaggregate within- from between-person (and lagged from concurrent) relations but whose parameters must be locally stationary.

Overview and Hypotheses

In the present study, we will apply network intervention analysis and temporal network

[11] Specifically, a lag-1 GVAR model obtains the temporal network by regressing responses for each variable at a given time point on responses for that variable (autoregressive effects) and other variables (cross-lagged effects) at the previous time point and obtains the contemporaneous network by modeling the relations among the residuals for all variables from the temporal network as a Gaussian graphical model (GGM). The GGM represents the variables as nodes connected by edges whose weights reflect partial correlations (for overviews of GGMs and GVAR models, see Burger et al., 2022; Epskamp et al., 2018; Epskamp et al., 2022). The temporal network models how well deviations from person means for each node at a given time point uniquely predict person-mean deviations for all other nodes at the next time point. The contemporaneous network models unique relations among deviations from person means for all nodes at the same time point, after controlling for temporal effects. The between-person network models unique relations among all nodes' person means.

modeling to data from the hybrid efficacy-effectiveness trial (Ji et al., 2021) of CBM-I in Study 1 that our team ran in a large transdiagnostic sample of anxious adults on a public research website called *MindTrails* (https://mindtrails.virginia.edu). For the network intervention analyses, we will model cross-sectional networks with nodes for (CBM-I) condition, interpretation biases, anxiety symptoms, and related impairment at each time point. Per our preregistration (https://doi.org/m292),[12] we expect that the positive CBM-I condition (relative to the no-training condition and to the 50-50 condition) will have direct effects on biases and indirect effects on symptoms and impairment via biases. We also expect that the positive CBM-I condition's direct effects on biases will emerge earlier than any of its direct effects on symptoms.

For the temporal network models, we will fit temporal, contemporaneous, and between-person networks with nodes for biases, symptoms, and impairment in each condition as part of a multigroup model. Given that change mechanisms primarily concern within-person relations (Hoffart & Johnson, 2020a), we will focus on the temporal and contemporaneous networks. For the no-training condition (which involved only imagery primes and did not include any CBM-I training), we expect bias nodes to predict symptom and impairment nodes in both the temporal and contemporaneous networks. We have no hypotheses for condition effects on these specific relations or others but expect lower overall connectivity in the positive CBM-I condition across time than in each of the other (no-training and 50/50) conditions.

To our knowledge, this is the first study to use network models to analyze the effects of cognitive bias modification on cognitive biases or symptoms. Although prior studies have analyzed networks of interpretation bias (along with attentional or memory biases) and total

[12] The network analysis preregistration outlines four sets of analyses, not all of which are for this dissertation. Study 2 of this dissertation includes network intervention analyses and temporal network modeling analyses in Managing Anxiety that are part of the first and third sets of preregistered analyses (Studies 1 and 3 of preregistration).

scores for mental health (Parsons et al., 2021) or anxiety (Claus et al., 2023), this is also the first study to model networks of interpretation biases and individual symptom or impairment nodes.

Method for Study 2

Participants and Procedure

In Managing Anxiety (see Ji et al., 2021, for details), an RCT with two crossed factors, community adults with at least moderate anxiety symptoms were randomly assigned to one of three conditions and completed a baseline assessment, forming the intent-to-treat (ITT) sample. The conditions were positive CBM-I with ambiguous scenarios that ended positively 90% of the time and negatively 10% of the time, a 50-50 CBM-I condition with scenarios that ended positively 50% of the time and negatively 50% of the time, and a no-training control condition. Within each condition, participants were also randomly assigned to complete either an anxious or neutral imagery prime prior to each session. CBM-I groups completed eight training sessions at least 2 days apart, each followed by an assessment, and a 2-month follow-up assessment. Session 1 training and its assessment became available immediately after the baseline assessment. The no-training group completed only imagery primes and assessments.

We focused on ITT participants (807 overall) with available item-level data (N = 729; 230 in positive CBM-I, 243 in 50-50 CBM-I, 256 in no-training). We used these participants' assessment data from baseline (n = 729) through Session 6 (n = 83). Anxiety symptoms and impairment were assessed at every time point, and interpretation bias was assessed at baseline, Session 3, and Session 6. These measures were also included at Session 8 and 2-month follow-up. We exclude the follow-up data to retain equidistant measures in training dose and more equidistant measures in elapsed time. We exclude the Session 8 data given few observations (n = 26 out of the 807 ITT participants), per the main outcomes paper (Ji et al., 2021).

Main outcome analyses showed superior improvement in interpretation bias in the positive CBM-I condition versus the 50/50 and no-training conditions and superior reductions in total anxiety scores in positive CBM-I versus no-training (Ji et al., 2021). Initial temporal network models of the relations among interpretation bias, anxiety symptoms, and impairment were attempted for a course project (Eberle, 2021) but have not been published. Also, the course project models used data that had not been transformed to meet the stationarity assumption and did not converge (connectivity was not computed and no models were compared). We will transform the data in the present study.

Measures

Interpretation Biases

Interpretation biases were assessed with the RR measure based on the 18-item two-factor model from Study 1. We computed positive and negative bias scores as mean ratings of the 9 positive and 9 negative threat interpretations, respectively. We reverse scored the positive bias items to indicate a lack of positive bias. The distinct (yet correlated) negative threat and positive threat factors found in Study 1 support the inclusion of both nodes as sufficiently distinct and separately identifiable (see Bringmann et al., 2022), and different effect sizes of CBM-I (vs. comparators) on these biases (Eberle et al., 2023; Larrazabal et al., 2023; but see also Ji et al., 2021) support them as independently manipulable.

Anxiety Symptoms and Impairment

Anxiety symptoms and components of functional impairment were assessed with the Overall Anxiety Severity and Impairment Scale (OASIS; Norman et al., 2006), a five-item self-report of (a) frequency of anxiety, (b) intensity/severity of anxiety, (c) frequency of anxious avoidance, (d) degree of social impairment, and (e) degree of work impairment in the past week.

Items are rated on a 5-point scale from 0 (*lowest frequency/intensity/severity/degree*) to 4 (*highest frequency/intensity/severity/degree*), with response option wording varying by item (Appendix B). CBM-I trials have used the total (or average item) score as the main outcome for anxiety. We included each of the five items as manifest nodes (following Curtiss et al., 2018) given our interest in isolating each of the five variables and modeling their unique relations. For zero-order Pearson correlations among all seven RR and OASIS nodes for ITT participants at baseline (*r*s = .10-.59), see Table SB1.

Statistical Analysis

For the Study 2 data and analysis code, see https://osf.io/2ytr6/. For our preregistration deviations, see Section SB1 in Supplement B (also at https://osf.io/2ytr6/).

Network Intervention Analysis

To test condition's direct and indirect effects on mean levels of nodes, we used the `mgm` package (ver. 1.2-14; Haslbeck & Waldorp, 2020) in R (ver. 4.4.0; R Core Team, 2024) to fit a cross-sectional mixed graphical network model with manifest nodes for condition, interpretation biases, anxiety symptoms, and impairment at each time point that included both bias and anxiety measures (baseline, Session 3, Session 6).

After estimating the networks using the regularized methods we had preregistered (for details, see Section SB1.1) and realizing issues of interpreting estimates of parameter uncertainty for regularized models, which do not yield CIs (Williams, 2021), and given our interest in specific edges, we estimated saturated networks without regularization (Williams et al., 2019; Williams & Rast, 2020). To do so, we set `mgm`'s least absolute shrinkage and selection operator (LASSO) regularization tuning parameter λ to zero (per Fried et al., 2020), which removes the LASSO penalty function from the maximum likelihood estimate (Blanken et al., 2022, p. 121), and removed `mgm`'s default beta-min threshold `LW` (i.e., Loh & Wainwright, 2013). Because this approach does not test the significance of edges automatically, we then used `mgm`'s `resample` function to fit each model in 500 bootstrap samples and computed 95% and 99% CIs as the quantiles (2.5th and 97.5th; 0.5th and 99.5th; respectively) of the bootstrap distribution. We interpreted significant edges as those whose CIs excluded zero and used thresholding to plot significant edges. We focused on interpreting edges significant at the alpha level of .01, which has been suggested for reducing false positives (Williams & Rast, 2020, p. 200) and is a common default in network psychometrics (Isvoranu & Epskamp, 2021). However, when effects for ITT participants and completers differ, we also report edges significant at the alpha level of .05 (and plot all edges significant at .05 in the supplement), which may yield a better balance between inference and predictive accuracy (Williams & Rodriguez, 2022), though we prioritize inference. We plot and report all edges for nonthresholded networks in the supplement. We plotted networks using the `qgraph` package (ver. 1.9.8; Epskamp et al., 2012).

To maximize the sample size in each group, we compared the positive CBM-I, 50/50 CBM-I, and no-training groups (collapsing across imagery prime conditions in each group). Condition (binary) was coded 1 for positive CBM-I and 0 for the reference group in two models (vs. no-training and vs. 50/50). All other nodes were treated as continuous and standardized using `mgm`'s default method (`scale = TRUE`). (Although bias measured by RR [four response options] may generally need to be treated as ordinal, `mgm` does not have a dedicated method for modeling ordinal data. Hence, prior studies using `mgm` have modeled ordinal data as continuous [e.g., see suppl. of Fried et al., 2020]. Also, `mgm` performs poorly when ordinal data are treated as categorical [Isvoranu & Epskamp, 2021].)

Missing Data Handling. Little work has addressed missing data handling for network modeling (Blanken et al., 2022). Following prior network intervention analyses (e.g., Blanken et al., 2019; Blanken et al., 2021), we used listwise deletion, assuming the data are MCAR and analyzing networks with different sample sizes at each time point. As a completer analysis, we also fit the models in participants with complete data for all nodes across all time points (n = 83). To compare plots across time points, condition comparisons, and samples, we set the maximum edge width to the maximum edge magnitude across all network intervention analyses.

Temporal Network Modeling

To estimate relations among interpretation biases, anxiety symptoms, and impairment in each condition, we used the `panelgvar` function of the `psychonetrics` R package (ver. 0.12; Epskamp, 2024b) to specify a multilevel lag-1 GVAR model for panel data (Epskamp, 2020) based on time points that included both RR and OASIS measures (baseline, Session 3, Session 6; see Section SB1.2). Including only random intercepts, the model assumes that each person has the same network structure but different means. We sought to jointly estimate a directed fixed effect temporal within-person network (with effects then standardized to partial directed correlations), undirected fixed effect contemporaneous within-person network, and between-person network. However, due to estimation issues, we interpreted only the within-person networks.[13] A *directed* edge from one node to another in the temporal network reflects the unique effect of the former node at a given time point on the latter node at the next time point, whereas an *undirected* edge between two nodes in the contemporaneous network reflects the unique relation between the two nodes at a given time point. Each within-person relation is set to

[13] When first trying to model between-person and contemporaneous within-person effects as GGMs, the between-person standard errors were implausibly large and the between-person scaling parameters were nonsignificant (in addition, in the multigroup model we could only approximate standard errors). Therefore, per Epskamp (2023), we modeled between-person effects using a Cholesky decomposition and did not interpret the between-person effects. This approach still yields correct estimates for temporal and contemporaneous within-person effects (Freichel & Epskamp, 2024). The within-person effects and model fit were nearly identical regardless of the between-person model, and using the decomposition allowed us to avoid approximating standard errors in the multigroup model.

equality across all time points.

After estimating the networks using the pruning and model search procedure we had preregistered (for details, see Section SB1.3), we realized issues with estimates of parameter uncertainty with such model selection methods (Williams, 2021). Considering this and our interests in specific edges and in comparing conditions, to avoid model selection bias we estimated saturated networks (using full information maximum likelihood estimation; FIML) and did not apply any data-driven model selection methods. This approach tests the significance of edges automatically (bootstrapping is unneeded) and provides indices of model fit, which we assessed based on the chi-square, RMSEA, CFI, TLI, AIC, and BIC. We again used thresholding to plot significant edges using alpha levels of .01 and .05 and focused on interpreting edges significant at .01. We plot nonthresholded networks in the supplement. After first fitting a single-group model across all conditions to resolve any initial estimation issues, we fit a three-group model, which can handle unequal group sizes, to compare the networks by condition (positive CBM-I vs. 50-50 CBM-I and vs. no-training) on overall connectivity (although the present dissertation includes analysis of only descriptive differences). We plotted the multigroup networks with `qgraph`. To compare plots across conditions, we set the maximum edge width to the maximum edge magnitude across the within-person networks across all three conditions.

We modeled all nodes as manifest[14] and treated them as continuous. (Again, although interpretation biases measured with RR may generally need to be treated as ordinal, `psychonetrics` does not permit modeling of mixed networks. The package does offer a three-stage weighted least squares estimator to model polychoric correlations [treating all nodes as

[14] Although *panel-lvgvar* models can include nodes for not only manifest variables, but also latent variables (e.g., which might eventually be used for interpretation bias nodes; Epskamp, 2020), the sample size required per node is not yet known. Thus, we followed advice to start with a small number of manifest nodes (S. Epskamp, personal communication, July 26, 2020).

ordinal], but this performs similarly to the FIML estimator that treats all nodes as continuous [Isvoranu & Epskamp, 2021].)

Connectivity. We computed connectivity in two ways: (a) *global strength* (mean of absolute values of edge weights), which reflects how much nodes covary, regardless of whether node relations are positive or negative; and (b) *global expected influence* (mean of signed edge weights), which accounts for negative edges and reflects how much nodes positively covary (i.e., activate one another; see Robinaugh et al., 2016). For the temporal network, we computed inter-node connectivity (only cross-lagged effects), intra-node connectivity (only autoregressive effects), and overall connectivity (all effects; see Groen et al., 2019). For the contemporaneous network, which has no autoregressive effects, we computed only inter-node connectivity. We computed multiple measures given lack of consensus on how connectivity should be defined with respect to the connectivity hypothesis (Blanchard & Heeren, 2022). We used all edges to compute connectivity (including nonsignificant edges; following Groen et al., 2019), again to avoid model selection bias (Williams, 2021) when testing differences between conditions in future analyses that are not part of the present dissertation. See Section SB1.4 for details.

Detrending. Given that the model assumes stationarity, we removed linear trends prior to analysis (Burger et al., 2022) by using the `lm` function to regress each variable on a fixed linear effect of time (i.e., assessment point; coded 0 [baseline] to 6 [Session 6]) in ordinary least squares (OLS) regressions and replacing the nonmissing scores with the residuals for each significant effect ($\alpha = .05$ per Epskamp, 2020, and Epskamp et al., 2018).[15] Given condition

[15] Time-varying network models that do not assume stationarity and explicitly model changes in means, variances, and network parameters over time require time-series data (Haslbeck et al., 2022b). In time-series data, detrending has been done using OLS regression for each participant (Epskamp, 2020; Epskamp et al., 2018; Fisher et al., 2017) or across participants (Fried et al., 2021). In panel data, it has been done by using OLS regression across participants (Freichel, 2023; Freichel et al., 2023) or standardizing each variable at each time point (O’Driscoll et al., 2022). Modeling continuous time may be a fruitful future direction (see Ryan & Hamaker, 2022); however, for simplicity, we model discrete assessment points, which are equidistant in training dose but not in elapsed time.

differences in trends (Ji et al., 2021), we tested and removed the trends separately by condition. Although we considered then standardizing each variable across time points and participants (per Epskamp, 2024a; Freichel, 2023, Freichel et al., 2023), because this approach yielded nearly identical panel GVAR results and the model already yields standardized estimates as partial correlations, we removed this nonpreregistered step.

Missing Data Handling. In the panel GVAR models, we assumed the data are MAR and used `psychonetric`'s FIML estimation and default `nlminb` optimizer. The OLS models used for detrending used pairwise deletion, which assumes the data are MCAR.

Results for Study 2

Network Intervention Analysis

Networks showing edges significant at the alpha level of .01 for ITT participants are in Figures 3-4 (for completers, see Figures SB3 and SB8). Networks showing edges significant at the alpha level of .05 (ITT: Figures SB1 and SB6; completers: Figures SB4 and SB9) and all edges (ITT: Figures SB2 and SB7; completers: Figures SB5 and SB10) are in Supplement B. For numeric results of all edges, see Tables SB2-SB3. The significant effects we refer to hereafter are significant at the alpha level of .01 (to reduce false positives and prioritize inference over prediction), unless noted otherwise (i.e., when the direct or indirect effects significant at $\alpha = .01$ differ for ITT participants vs. completers, we also note whether effects are significant at $\alpha = .05$).

Baseline Relations

Although we had no baseline hypotheses, consistent with random condition assignment no significant edges emerged between condition and any nodes at baseline for ITT participants (Figures 3-4; or completers; Figures SB3 and SB8). Both baseline networks (positive CBM-I vs. 50-50 CBM-I; positive CBM-I vs. no-training) had significant edges representing positive unique relations among the anxiety items and a positive unique relation between negative interpretation bias and situational avoidance. In addition, the positive CBM-I versus no-training network had a significant positive unique relation between negative bias and a lack of positive bias (Figure 4).

Condition Differences in Means

As hypothesized, positive CBM-I had significant direct effects on interpretation biases. Relative to 50-50 CBM-I, positive CBM-I significantly reduced negative bias by Session 3 and Session 6, and at Session 6 negative bias had a significant positive unique relation with anxiety frequency (suggesting a possible indirect effect of positive CBM-I on anxiety frequency via bias, although this possible mediator-outcome relation is not temporal, limiting any causal inference; Figure 3). Relative to no-training, positive CBM-I significantly reduced lack of positive bias by Session 3 (when positive CBM-I was higher on positive bias), and lack of positive bias had a significant positive relation with situational avoidance at that time point (tentatively suggesting another possible indirect effect; Figure 4). Unexpectedly, positive CBM-I's direct effect on lack of positive bias at this time point was accompanied by a direct effect on anxiety frequency (i.e., a reduction), contrary to our hypothesis that positive CBM-I's direct effects on bias would emerge earlier than any of its direct effects on symptoms or impairment.

When analyzing participants with complete data across all three time points, the direct effect of positive CBM-I on interpretation bias was only observed relative to 50-50 CBM-I at Session 6, when the positive CBM-I group had significantly lower negative bias (Figure SB3). However, at this time point, negative bias had a significant negative unique relation with anxiety severity (in the corresponding network above, which included one additional participant who had complete data at Session 6 but not at other time points, this edge was significant only at $\alpha = .05$; Figure SB1). Moreover, negative bias's positive unique relation with anxiety frequency was no longer significant at the alpha level of .01 (although it was at $\alpha = .05$; Figure SB4). Additionally, among completers, positive CBM-I's direct effect on lack of positive bias relative to no-training at Session 3 (when positive CBM-I was higher on positive bias) was nonsignificant (Figure SB8; and positive CBM-I's direct effect on anxiety frequency was significant only at $\alpha = .05$; Figure SB9). Instead, at this time point, positive CBM-I had significantly higher social impairment than no-training; however, because the completer sample does not retain random assignment to condition, this difference may be due in part to baseline differences. Although completers in positive CBM-I did not have significantly greater social impairment at baseline than those in no-training (Figure SB8), the full network shows moderately greater baseline social impairment in positive CBM-I ($r_{ij.\mathbf{z}} = .41$; Figure SB10).

Temporal Network Modeling

Networks showing edges significant at the alpha level of .01 are in Figure 5; for networks showing edges significant at the level of .05 and for networks showing all edges, see Figures SB11-SB12, respectively. For numeric results of all edges, see Table SB4. Unless noted otherwise, the significant effects we refer to hereafter are significant at the alpha level of .01.

Within-Person Relations in No-Training

As hypothesized, in the no-training condition, the temporal network revealed a significant positive unique directed edge such that, on average across time points, a greater lack of positive bias at a given time point (relative to one's general, mean level of positive bias) predicted greater work impairment (than one's mean level of work impairment) at the next time point (Figure 5). Unexpectedly, no other significant relations of bias nodes with symptom or impairment nodes were found in the temporal or contemporaneous networks in the no-training condition.

Condition Differences in Overall Connectivity

With respect to hypothesized condition differences in connectivity, positive CBM-I had descriptively lower connectivity in its saturated temporal and contemporaneous networks than 50-50 CBM-I on every metric (range of differences = -.02 to -.09; Table 5). Positive CBM-I's networks also had descriptively lower connectivity than no-training on nearly every metric (range of expected differences = -.02 to -.07), except for intra-node connectivity in the temporal network based on mean strength (unexpected difference = .05). Although these differences may tentatively suggest lower connectivity in positive CBM-I, they need to be statistically tested (van Borkulo et al., 2022). The potentially lower connectivity in positive CBM-I also appears evident when visually comparing the significant edges in its networks with those of 50-50 CBM-I and no-training (Figure 5). However, it is important to note that connectivity was estimated using saturated networks (including significant and nonsignificant edges) and that differences in edge weights also need to be statistically tested (Gelman & Stern, 2006; Nieuwenhuis et al., 2011).

Linear Trends and Model Fit

OLS regressions run prior to the panel GVAR analyses revealed significant linear trends (at $\alpha = .05$) for every variable in every condition except for negative interpretation bias in the 50-50 CBM-I and no-training conditions. The saturated multigroup GVAR model fit to detrended data showed mixed evidence of fit, $\chi^2(420) = 747.58$, $p < .001$; RMSEA = 0.057; CFI = .86, TLI = 0.77, in reference to traditional guidelines (i.e., nonsignificant chi-square, RMSEA near or < 0.06; CFI and TLI near or > .95; Hu & Bentler, 1999).

Discussion for Study 2

The present study used network modeling to test the effects of web-based CBM-I on interpretation biases, anxiety symptoms, and related impairment viewed as a complex system using data from an RCT run in a transdiagnostic sample of anxious adults. Consistent with our hypothesized effects on mean levels of nodes in cross-sectional networks at each time point, in our ITT sample, positive (vs. 50-50) CBM-I had a significant direct effect on negative bias at Sessions 3 and 6 and an indirect effect on anxiety frequency at Session 6. Moreover, at Session 3 positive CBM-I (vs. no-training) had a significant direct effect on lack of positive bias (when positive CBM-I was higher on positive bias) and an indirect effect on situational avoidance, although positive CBM-I also had a significant direct effect on anxiety frequency at the same (rather than a later) time point. Only the direct effect of positive (vs. 50-50) CBM-I on negative bias at Session 6 was significant in completer analyses. Temporal network analyses found that, consistent with our hypothesized within-person relations in no-training, lack of positive bias predicted work impairment. Additionally, consistent with our hypothesized condition differences in connectivity, overall positive CBM-I had descriptively lower network connectivity than 50-50 CBM-I and no-training, although these differences in connectivity need to be statistically tested.

Network Intervention Analysis

Positive CBM-I's direct effects on interpretation biases and indirect effects on anxiety symptoms via biases are consistent with cognitive models of anxiety, which posit that cognitive changes precede and mediate symptom improvements (Smits et al., 2012; Steinman & Teachman, 2014). The direct negative effect of positive CBM-I (vs. 50-50 CBM-I) on negative bias at Session 6 and the positive unique relation between negative bias and anxiety frequency suggest an indirect effect of positive CBM-I on anxiety frequency via negative bias. Although causality cannot be inferred from this cross-sectional relation between negative bias and anxiety frequency, it seems intuitive that a lower tendency to perceive threat would prompt anxiety less frequently. Given that anxiety frequency had overall positive unique relations with other symptom and impairment nodes, this suggests a potential pathway through which positive CBM-

I may improve certain symptom and impairment nodes relative to 50-50 CBM-I. However, the main outcomes paper found that positive CBM-I did not improve significantly more than 50-50 CBM-I on the OASIS total score over time in the full ITT sample (Ji et al., 2021). Although the differences in sample size and missing data handling between the present analyses and the main paper may explain this discrepancy, another possible explanation is that the significant unique relation between negative bias and anxiety frequency did not emerge until Session 6. This raises the possibility that it might take time for users to apply their unique treatment gains in reducing negative threat interpretations to actual anxiety-provoking situations in their daily lives.

Additionally, the direct negative effect of positive CBM-I (vs. no-training) on lack of positive bias at Session 3 (when positive CBM-I was higher on positive bias) and the positive unique relation between lack of positive bias and situational avoidance suggest an indirect effect of positive CBM-I on situational avoidance via lack of positive bias. Again, although the relation between lack of positive bias and situational avoidance is cross-sectional, this may highlight a unique role for envisioning positive outcomes (vs. only not assuming negative outcomes) in not avoiding situations that provoke anxiety. Given situational avoidance’s positive unique relation with social impairment, this suggests a potential pathway through which positive CBM-I may improve symptom and impairment nodes relative to no-training. Such a pathway may help explain the main outcomes paper’s finding that positive CBM-I improved significantly more than no-training on the OASIS total score over time (Ji et al., 2021). Positive CBM-I’s direct effect on anxiety frequency may represent another path to improved symptom and impairment nodes. Although we expected direct effects on bias nodes to emerge earlier than any direct effects on symptom and impairment nodes, not assessing biases until Session 3 (after assessing them at baseline) may have prevented detection of earlier direct effects on biases. The direct effect on anxiety frequency also raises questions about what variables not included in the network could explain this direct effect; comparing 50-50 CBM-I with no-training may provide clues (e.g., common factors that are present in both positive and 50-50 CBM-I vs. factors unique to positive CBM-I). Although the main paper found no significant effects of 50-50 CBM-I (vs. no-training) on biases or the OASIS total score over time (analyzed separately), it may be fruitful to test for any unique effects of 50-50 CBM-I (vs. no-training) on biases and symptom/impairment nodes in a network.

It is important to note that, except for the direct effect of positive CBM-I (vs. 50-50 CBM-I) on negative bias at Session 6, the direct and indirect effects of positive CBM-I (vs. 50-50 CBM-I or no-training) at Sessions 3 and 6 were not significant among participants with complete data at all three time points (*n* in each model for participants with complete data = 55-57). Given that fewer participants had data at Session 6 (*n* in each model = 58) than at Session 3 (*n* in each model = 102-113), we are more confident in the results at Session 3. Specifically, these results are (a) the direct effect of positive CBM-I (vs. 50-50 CBM-I) on negative bias and (b) the direct effects of positive CBM-I (vs. no-training) on lack of positive bias and on anxiety frequency (i.e., participants in positive CBM-I were higher on positive bias and lower on anxiety frequency). Session 3 results also indicated the indirect effect of positive CBM-I on situational avoidance via lack of positive bias (i.e., participants in positive CBM-I were higher on positive bias, and participants higher on positive bias were lower on situational avoidance). It is also important to note that fewer participants had data at Session 3 than at baseline (*n* in each model = 439-469) and that Session 3 participants (and Session 6 participants) may no longer reflect the ITT sample. To further improve confidence in Session 3 results, future research could analyze Session 3 participants' baseline networks to assess whether the condition differences at Session 3 were already present at baseline in this subsample. Future research should also consider potential ways to handle missing data that would retain the ITT sample over time (e.g., multiply impute data with a joint multivariate linear mixed model including condition, time, and their interaction as predictors; then fit the network models in each imputed dataset and pool the results).

Comparing the present direct effects of positive CBM-I on biases to the condition effects in the main outcomes paper (Ji et al., 2021) shows that the direct effects both mostly align with the main outcomes paper's condition effects and extend them by suggesting that positive CBM-I may have unique effects on biases, even after controlling for other nodes in the network. The direct effects of positive CBM-I (vs. 50-50 CBM-I) on negative bias at Sessions 3 and 6 accord with the main outcomes paper's finding that positive CBM-I improved significantly more than 50-50 CBM-I on negative bias over time (Ji et al., 2021) and suggest that completing mostly positive (vs. 50% positive) training scenarios may yield *unique* improvements in negative bias. Positive CBM-I (vs. 50-50 CBM-I) had no significant direct effects on lack of positive bias, consistent with the main paper's finding of no significant differences between these conditions on positive bias over time. The direct effect of positive CBM-I (vs. no-training) on lack of positive bias at Session 3 (when positive CBM-I was higher on positive bias) accords with the main paper's finding that positive CBM-I improved significantly more than no-training on positive bias and suggests that completing positive (vs. no) scenarios yields unique improvements in positive bias. Although the main paper found that positive CBM-I improved significantly more than no-training on negative bias, positive CBM-I (vs. no-training) had no significant direct effects on negative bias in this study. Comparing the main paper's results with the present results are complicated by differences in sample size and missing data handling (e.g., the main paper used FIML to retain the full ITT sample, whereas the present analyses used listwise deletion at each time point). However, the present direct effects suggest that some of positive CBM-I’s effects on individual biases in the main paper may hold after controlling for all other nodes, and suggest that the direct effects emerge by Session 3.

Temporal Network Modeling

In the temporal network for the no-training condition, the positive unique cross-lagged relation from lack of positive bias to anxiety-related work impairment accords with predictions for within-person relations stemming from cognitive models of anxiety (Smits et al., 2012), such as that when threat-related cognitions are higher than one’s average level, anxiety will also be higher than average (e.g., Hoffart, 2016; Hoffart & Johnson, 2020b). However, it is notable that lack of positive bias did not have significant temporal (or contemporaneous) effects on anxiety symptom nodes. If such an effect had been significant, it might help explain how lack of positive bias at a given time point might contribute to later work impairment. However, our assessment schedule may explain this nonsignificant finding. Namely, the temporal effects of interpretation biases on anxiety symptoms may occur over shorter time intervals than those spanning the three time points included in the present analysis (i.e., ≥ 4 days between baseline and Session 3; ≥ 6 days between Session 3 and Session 6). Given that a sampling rate that is too infrequent to assess theoretical changes in variables can obscure estimates of their within-person relations (Slipetz et al., 2023; Haslbeck & Ryan, 2022), future studies should consider assessing nodes in the network on shorter time scales (e.g., using ecological momentary assessment).

Although the condition differences in network connectivity need to be statistically tested, the descriptively lower connectivity for positive CBM-I relative to 50-50 CBM-I and no-training tentatively suggests that, in addition to reducing the mean levels of nodes, positive CBM-I may weaken the relations among nodes. Positive CBM-I's direct and indirect effects on mean levels of certain nodes suggest potential paths via which positive CBM-I reduces the activation of bias, anxiety symptom, and impairment nodes. If the permutation tests we have already preregistered (https://doi.org/m292) show that positive CBM-I has significantly lower connectivity, this would suggest that positive CBM-I may also help destabilize relations that may otherwise resist the network's transition out of a stable state of higher activation. Given that the networks of anxious people seeking CBM-I are likely already at a state of higher activation at baseline, if positive CBM-I only deactivates nodes without changing the network's structure, the relations among the nodes may themselves return the network to the higher state of activation. That is, not changing the nodes' relations may maintain a dynamical landscape in which the attraction of the stable state of higher activation is stronger than that of a stable state of lower activation. A significant reduction of connectivity in positive CBM-I, if found, may therefore decrease the vulnerability of the network and improve the network's resilience, making it easier to remain in a lower state of activation if external stressors reactivate certain nodes (see van Bork et al., 2024, for discussion of features of a network's dynamical landscape; e.g., *vulnerability* and *stable state*).

Because significant differences in overall network connectivity do not entail significant differences in overall network structure (and vice versa; van Borkulo et al., 2022), future work should also test for condition differences in overall structure and in the weights of specific edges (also already preregistered; https://doi.org/m292). Such tests may yield insights about which relations especially contribute to potentially lower connectivity in positive CBM-I. Further, knowing which specific edges significantly differ between conditions may provide clues as to moderating variables that might explain how positive CBM-I potentially changes node relations. The present study's pairwise networks assume that no nodes in the network moderate relations among other nodes—that the unique relation between any two nodes is independent of the values of other nodes. Moderated networks that allow the values of some or all nodes to moderate other nodes' relations (Haslbeck et al., 2021) may be useful for testing whether the network's structure depends on the values of certain nodes (e.g., O'Driscoll et al., 2022). Such networks can be fit with cross-sectional (Haslbeck et al., 2021; Haslbeck, 2022) or time series (Swanson, 2020) data.

Future Directions

Given that the mixed graphical models used for the network intervention analyses and the panel GVAR models used for the temporal network analyses are both fixed effect networks that estimate average, nomothetic effects across participants, future studies should consider collecting time series data that would enable estimation of personalized, idiographic effects per participant. For example, even though the panel GVAR models disaggregate between- from within-person relations, they assume that the within-person relations are the same for all participants (Epskamp, 2020). In addition, the panel GVAR models assume that the means, variances/covariances, and all network parameters are locally stationary across included time points (baseline to Session 6; Epskamp, 2020). Although we removed linear trends before fitting these models to make their assumption of stationary means more plausible, such detrending can also artificially inflate the models' fit because the models themselves constrain parameters to be the same across all time points (Epskamp, 2020; Epskamp, 2024a), and even after detrending, model fit was still mixed. Further, when trends result from an intended manipulation (e.g., an intervention), detrending the variables can artificially reduce the within-person effects among variables (Wang & Maxwell, 2015). Time series models would have the added benefit of enabling estimation of time-varying network models that explicitly model changes in parameters over time (Haslbeck et al., 2022b).

In the present study, we focused on testing conditional dependencies (partial correlations whose magnitudes significantly differ from zero). However, future research should also consider testing conditional independencies (i.e., partial correlations whose magnitudes are significantly smaller than a given effect size of interest) using equivalence tests, given that a nonsignificant edge in a conditional dependence network could indicate an edge that is either truly zero or just undetected (Williams et al., 2021). Identifying conditional independencies may inform causal inference by suggesting variables that may be causally independent (Ryan et al., 2022). Bayesian estimation may be especially useful to quantify the relative evidence (using Bayes factors) that each edge indicates conditional dependence, conditional independence, or ambiguity (i.e., the evidence for conditional dependence vs. conditional independence is inconclusive), although larger samples may be needed to evaluate the edges deemed ambiguous (Williams et al.). Conditionally independent and ambiguous edges are important to consider given the possibility that edges that significantly differ from zero (traditionally visualized in conditional dependence networks) do not significantly differ from edges that do not significantly differ from zero (which are traditionally not visualized in conditional dependence networks; Williams et al.).

Although the present study focused on positive CBM-I's effects on networks over the course of treatment, a network perspective may also be useful for studying maintenance of gains after the end of treatment. Some studies of positive CBM-I have found significant losses of some treatment gains with respect to mean levels of interpretation biases and anxiety symptoms (e.g., OASIS total score; Eberle et al., in press, in which CBM-I participants significantly worsened on most outcomes during 2-month follow-up, although overall improvements from baseline to follow-up still favored CBM-I over a comparison condition). However, other studies of positive CBM-I have shown better maintenance of gains (e.g., see Larrazabal et al., 2023, for follow-up over 6 months). Perhaps evaluating network connectivity after the end of treatment could help explain whether treatment gains from positive CBM-I are maintained (e.g., if node relations that are potentially destabilized during treatment regain their strength after treatment, the network may become more vulnerable to reactivation by external stressors, leading to loss of treatment gains in terms of worsening mean levels of nodes). This perspective highlights the potential utility of considering node relations in addition to mean levels of nodes in developing treatment targets and in testing treatment effects both over the course of treatment and after it ends.

Another future direction is to optimize the selection and measurement of nodes. Although Study 1 informed the structural validity of the interpretation bias nodes, which were modeled as average item scores, the anxiety symptom and impairment nodes were modeled as single items (as is common in network models to date), which assume no measurement error (van Bork et al., 2024). Even though the network itself may require a measurement theory different from existing theories (e.g., classical test theory, latent variable or modern test theory) given that its targets of measurement are different from those of existing theories, existing theories can still be used for the measurement of nodes themselves, depending on one's conceptualization of the nodes (van Bork et al.). (See van Bork et al. for proposed measurement targets for networks.) Thus, just as multiple items were used to measure each interpretation bias node, the measurement of certain nodes (e.g., avoidance) may be improved by using multiple items (e.g., including an item on experiential avoidance in addition to the item on situational avoidance), whereas single items may be sufficient for more straightforward nodes (e.g., anxiety frequency). Future work could also consider refining the nodes included in the network, which should include the minimally complete set of nodes needed to model a given system (Bringmann et al., 2022). For example, certain beliefs (e.g., coping self-efficacy) or other biases (e.g., threat-related attention or memory biases) posited in models of anxiety (Abramowitz et al., 2019, pp. 35-49) could be considered. However, some nodes (e.g., beliefs) may operate on longer time scales than others, a challenge that continuous time models of time series data may help address (Bringmann et al., 2022).

Limitations

In addition to the considerations discussed in prior sections, the present study has a few limitations. First, the mixed graphical models for network intervention analyses and the panel GVAR models for temporal network analyses assumed data were MCAR or MAR, respectively. However, a prior analysis of Managing Anxiety data found an association between confidence that an online training program designed to change one's thinking about situations will reduce one's anxiety (rated at baseline) and whether or not participants dropped out of the study before Session 6 (Hohensee et al., 2020). Predictors of missingness can bias parameter estimates if they also correlate (e.g., $\geq |.40|$) with values of incomplete analysis variables (Enders, 2010, p. 133). Thus, future work should test whether this variable or other potential predictors of missingness (e.g., demographic variables, device used to complete the study, and self-reported importance of reducing anxiety, which were predictors of missingness in other *MindTrails* studies; Eberle et al., 2023; Eberle et al., in press) correlate with the analysis variables. If so, then the feasibility of including the predictors as auxiliary variables in missing data handling (i.e., multiple imputation, maximum likelihood) should be considered.

Second, the CIs in the present study were fairly wide (especially for the mixed graphical models at Session 6), which can limit detection of significant edges (Williams et al., 2021). This underscores the need to replicate the present findings in a larger sample, which is especially key for partial correlation networks given that partial correlations have more sampling variability than zero-order correlations (Williams, 2022). Third, as noted in Study 1, our sample was mostly female, White, and not Hispanic or Latino, limiting generalizability. Fourth, the panel GVAR model assumes multivariate normality (Epskamp, 2020), and this model and the mixed graphical models assume only linear relations. If the true relation among two nodes is nonlinear, this may yield a partial linear correlation of zero (Ryan et al., 2022), and these linear models are unable to represent dynamical landscapes with the different stable states that we have referred to (van Bork et al., 2024). The Ising model, which has binary nodes (active vs. inactive), permits nonlinear dynamics, but it can currently be estimated only from cross-sectional data (van Bork et al.). Finally, the present descriptive differences in connectivity need to be statistically tested.

Conclusion

Network theory holds promise for studying mechanisms of change by conceptualizing psychopathology as a complex system of causally interacting symptoms and related elements. Analyzing cross-sectional networks of interpretation biases, individual anxiety symptoms, and related impairment, the present study tested the effects of positive web-based CBM-I (vs. 50-50 CBM-I and no-training comparators) on the mean levels of nodes at each time point, revealing (causal) direct effects and (possibly causal) indirect effects that suggest potential pathways of network deactivation. Analyzing temporal networks, the study also estimated the within-person relations among nodes in each condition over time, revealing descriptively lower connectivity for positive CBM-I (vs. 50-50 CBM-I and no-training). Although these condition differences need to be tested statistically, they tentatively suggest that positive CBM-I may also destabilize the network's structure, which may reduce the network's vulnerability to a higher state of activation.

General Conclusion

Using cognitive science principles to directly target threat interpretations of ambiguous situations, scalable CBM-I procedures hold promise for reducing the treatment gap for anxious people and for testing basic cognitive theories of anxiety. Inferences about CBM-I's effects on interpretation biases depend on valid measurement of these biases, and conceptualizing anxiety as a complex system of interacting elements holds promise for clarifying the pathways through which CBM-I changes interpretation biases and anxiety. Evaluating the structural validity of a key measure of interpretation biases and testing CBM-I's effects on networks of these biases, anxiety symptoms, and related impairment, the present dissertation strengthens inferences about interpretation biases and advances understanding of cognitive mechanisms of change in anxiety.

References

Abramowitz, J. S., Deacon, B. J., & Whiteside, S. P. H. (2019). *Exposure therapy for anxiety: Principles and practice* (2nd ed.). Guilford. https://psycnet.apa.org/record/2019-19532-000

Baee, S., Eberle, J. W., Baglione, A. N., Spears, T., Lewis, E., Behan, H. C., Wang, H., Funk, D. H., Barnes, L. E., & Teachman, B. A. (2024). *Early attrition prediction for digital mental health studies.* JMIR Preprints. https://doi.org/kqdk

Baxter, A. J., Scott, K. M., Vos, T., & Whiteford, H. A. (2013). Global prevalence of anxiety disorders: A systematic review and meta-regression. *Psychological Medicine, 43*(5), 897-910. http://doi.org/f4s9bw

Beard, C., Millner, A. J., Forgeard, M. J., Fried, E. I., Hsu, K. J., Treadway, M. T., Leonard, C. V., Kertz, S. J., & Björgvinsson, T. (2016). Network analysis of depression and anxiety symptom relationships in a psychiatric sample. *Psychological Medicine, 46*(16), 3359-3369. https://doi.org/10.1017/s0033291716002300

Beard, C., & Peckham, A. D. (2020). Interpretation bias modification. In J. S. Abramowitz & S. M. Blakey (Eds.), *Clinical handbook of fear and anxiety: Maintenance processes and treatment mechanisms* (pp. 359–377). American Psychological Association. https://doi.org/10.1037/0000150-020

Blanchard, A., & Heeren, A. (2022). Ongoing and future challenges of the network approach to psychopathology: From theoretical conjectures to clinical translations. In Asmundson, G. J. G. (Ed.), *Comprehensive clinical psychology* (2nd ed., pp. 32-46). https://doi.org/gjrnph

Blanken, T. F., Isvoranu, A-M., & Epskamp, S. (2022). Estimating network structures using model selection. In A. Isvoranu, S. Epskamp, L. Waldorp, & D. Borsboom (Eds.), *Network psychometrics with R* (pp. 110-131). Routledge.

Blanken, T. F., Jansson-Fröjmark, M., Sunnhed, R., Lancee, J. (2021). Symptom-specific effects of cognitive therapy and behavior therapy for insomnia: A network intervention analysis. *Journal of Consulting and Clinical Psychology*, *89*(4), 364-370. https://doi.org/gjf787

Blanken, T. F., Van Der Zweerde, T., Van Straten, A., Van Someren, E. J. W., Borsboom, D., & Lancee, J. (2019). Introducing network intervention analysis to investigate sequential, symptom-specific treatment effects: A demonstration in co-occurring insomnia and depression. *Psychotherapy and Psychosomatics*, *88*, 52-54. https://doi.org/gghm9d

Bonifay, W., & Cai, L. (2017). On the complexity of item response theory models. *Multivariate Behavioral Research*, *52*(4), 465-484. https://doi.org/10.1080/00273171.2017.1309262

Bornovalova, M. A., Choate, A. M., Fatimah, H., Petersen, K. J., & Wiernik, B. M. (2020). Appropriate use of bifactor analysis in psychopathology research: Appreciating benefits and limitations. *Biological Psychiatry*, *88*, 18-27. https://doi.org/ggjrbf

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, *16*(1), 5-13. https://doi.org/10.1002/wps.20375

Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, *9*, 91-121. http://doi.org/f4vznx

Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., Robinaugh, D. J., Perugini, M., Dalege, J., Constantini, G., Isvoranu, A., Wysocki, A. C., van Borkulo, C. D., van Bork, R., & Waldorp, L. J. (2021). Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers*, *1*(1), 1-18. https://doi.org/10.1038/s43586-021-00055-w

Bringmann, L. F., Albers, C., Bockting, C., Borsboom, D., Ceulemans, E., Cramer, A., Epskamp, S., Eronen, M. I., Hamaker, E., Kuppens, P., Lutz, W., McNally, R. J., Molenaar, P., Tio, P., Voelkle, M. C., & Wichers, M. (2022). Psychopathological networks: Theory, methods and practice. *Behaviour Research and Therapy, 149*. https://doi.org/gnpnpg

Bringmann, L., Helmich, M., Eronen, M., & Voelke, M. (2023). Complex systems approaches to psychopathology. In R. F. Krueger & P. H. Blaney (Eds.), *Oxford textbook of psychopathology* (4th ed., pp. 103-122). Oxford University Press. https://doi.org/10.1093/med-psych/9780197542521.003.0005

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford. https://psycnet.apa.org/record/2015-10560-000

Burger, J., Hoekstra, R. H. A., Mansueto, A. C., & Epskamp, S. (2022). Network estimation from time series and panel data. In A. Isvoranu, S. Epskamp, L. Waldorp, & D. Borsboom (Eds.), *Network psychometrics with R* (pp. 168-192). Routledge.

Cai, L. (2010). A two-tier full-information item factor analysis model with applications. *Psychometrika, 75*(4), 581-612. https://doi.org/10.1007/S11336-010-9178-0

Chisholm, D., Sweeny, K., Sheehan, P., Rasmussen, B., Smit, F., Cuijpers, P., & Saxena, S. (2016). Scaling-up treatment of depression and anxiety: A global return on investment analysis. *Lancet Psychiatry, 3*, 415-424. https://doi.org/f8mghm

Claus, N., Takano, K., & Wittekind, C. E. (2023). The interplay between cognitive biases, attention control, and social anxiety symptoms: A network and cluster approach. *PLOS One, 18*(4). https://doi.org/10.1371/journal.pone.0282259

Curtiss, J., Ito, M., Takebayashi, Y., & Hofmann, S. G. (2018). Longitudinal network stability of the functional impairment of anxiety and depression. *Clinical Psychological Science,* *6*(3), 325-334. https://doi.org/10.1177/2F2167702617745640

Duken, S. B., Moriya, J., Hirsch, C., Woud, M. L., van Bockstaele, B., & Salemink, E. (2024). *Reliability and validity of four cognitive interpretation bias measures in the context of social anxiety*. PsyArXiv. https://doi.org/10.31234/osf.io/t9jbr

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, *105*, 399-412. https://doi.org/10.1111/bjop.12046

Eberle, J. W. (2021). *Network analysis of interpretation bias and anxiety symptoms in online interpretation bias modification* [PSYC 5500 course paper]. University of Virginia.

Eberle, J. W., Boukhechba, M. O., Sun, J., Zhang, D., Funk, D. H., Barnes, L. E., & Teachman, B. A. (2023). Shifting episodic prediction with online cognitive bias modification: A randomized controlled trial. *Clinical Psychological Science, 11*(5), 819-840. https://doi.org/jw42

Eberle, J. W., Daniel, K. E., Baee, S., Silverman, A. L., Lewis, E., Baglione, A. N., Werntz, A., French, N. J., Ji, J. L., Hohensee, N., Tong, X., Huband, J. M., Boukhechba, M., Funk, D. H., Barnes, L. E., & Teachman, B. A. (in press). Web-based interpretation training to reduce anxiety: A sequential multiple-assignment randomized trial. *Journal of Consulting and Clinical Psychology*. https://doi.org/mmrc

Edwards, C. B., Portnow, S., Namaky, N., & Teachman, B. A. (2018). Training less threatening interpretations over the Internet: Impact of priming anxious imagery and using a neutral control condition. *Cognitive Therapy and Research, 42*, 832-843. https://doi.org/gfm9fr

Enders, C. K. (2010). *Applied missing data analysis.* New York: Guilford.

Epskamp, S. (2020). Psychometric network models from time-series and panel data.

Psychometrika, 85(1), 206-231. https://doi.org/10.1007/s11336-020-09697-3

Epskamp, S. [SachaEpskamp] (2023, October 18). *Comment on "warning: inv_sympd(): given matrix is not symmetric"* [GitHub issue comment]. GitHub. https://github.com/SachaEpskamp/psychonetrics/issues/25#issuecomment-1767748032)

Epskamp, S. (2024a). *Psychometric network modeling with psychonetrics.* Modern Modeling Methods Conference. https://osf.io/dax8g/?view_only=

Epskamp, S. (2024b). psychonetrics: Structural Equation Modeling and Confirmatory Network Analysis [Computer software]. https://cran.r-project.org/package=psychonetrics

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). Qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software, 48*(4), 1–18. https://doi.org/10.18637/jss.v048.i04

Epskamp, S., Haslbeck, J. M. B., Isvoranu, A., & van Borkulo, C. D. (2022). Pairwise Markov random fields. In A. Isvoranu, S. Epskamp, L. Waldorp, & D. Borsboom (Eds.), *Network psychometrics with R* (pp. 91-110). Routledge.

Epskamp, S., van Borkulo, C. D., van der Veen, D. C., Servaas, M. N., Isvoranu, A., Riese, H., & Cramer, A. O. J. (2018). Personalized network modeling in psychopathology: The importance of contemporaneous and temporal connections. *Clinical Psychological Science, 6*(3), 416-427. https://doi.org/10.1177/2F2167702617744325

Fisher, A. J., Reeves, J. W., Lawyer, G., Medaglia, J. D., & Rubel, J. A. (2017). Exploring the idiographic dynamics of mood and anxiety via network analysis. *Journal of Abnormal Psychology, 126*(8), 1044-1056. http://dx.doi.org/10.1037/abn0000311

Flake, J. K., Pek, J., & Hehman, E. (2017). Construct validation in social and personality research: Current practice and recommendations. *Social Psychological and Personality*

Science, 1-9. https://doi.org/10.1177/1948550617693063

Flora, D. B. (2020). Your coefficient alpha is probably wrong, but which coefficient omega is right? A tutorial on using R to obtain better reliability estimates. *Advances in Methods and Practices in Psychological Science, 3*(4), 484-501. https://doi.org/10.1177/2515245920951747

Flora, D. B., & Flake, J. K. (2017). The purpose and practice of exploratory and confirmatory factor analysis in psychological research: Decisions for scale development and validation. *Canadian Journal of Behavioural Science*, *49*(2), 78-88. https://doi.org/gbn562

Fodor, L. A., Georgescu, R., Cuijpers, P. Szamoskozi, S., David, D., Furukawa, T. A., & Cristea, I. A. (2020). Efficacy of cognitive bias modification interventions in anxiety and depressive disorders: A systematic review and network meta-analysis. *Lancet Psychiatry, 7*(6), 506-514. http://doi.org/gg27p8

Freichel, R. (2023). Symptom network analysis tools for applied researchers with cross-sectional and panel data – A brief overview and multiverse analysis. *Psychological Reports*. https://doi.org/10.1177/00332941231213649

Freichel, R., & Epskamp, S. (2024). *Handling problematic between-person estimates in panel network models: A comparative simulation study.* PsyArXiv. https://doi.org/10.31234/osf.io/eqaux

Freichel, R., Veer, I. M., Wiers, R. W., McNally, R. J., & Epskamp, S. (2023). *Cross-lagged panel models for studying psychopathology: A comparative overview of structural equation and panel network approaches.* PsyArXiv. https://doi.org/10.31234/osf.io/b94qt

Fried, E. I., & Cramer, A. O. J. (2017). Moving forward: Challenges and directions for psychopathological network theory and methodology. *Perspectives on Psychological*

Science, 12(6), 999-1020. https://doi.org/10.1177/1745691617705892

Fried, E. I., Papanikolaou, F., & Epskamp, S. (2021). Mental health and social contact during the COVID-19 pandemic: An ecological momentary assessment study. *Clinical Psychological Science, 10*(2), 340-354. https://doi.org/10.1177/2F21677026211017839

Fried, E. I., von Stockert, S., Haslbeck, J. M. B., Lamers, F., Schoevers, R. A., & Penninx, B. W. J. H. (2020). Using network analysis to examine links between individual depressive symptoms, inflammatory markers, and covariates. *Psychological Medicine, 50*, 2682-2690. https://doi.org/gg6378

Gäde, J. C., Schermelleh-Engel, K., & Klein, A. G. (2017). Disentangling the common variance of perfectionistic strivings and perfectionistic concerns: A bifactor model of perfectionism. *Frontiers in Psychology, 8*. https://doi.org/10.3389/fpsyg.2017.00160

Gelman, A., & Stern, H. (2006). The difference between "significant" and "not significant" is not itself statistically significant. *The American Statistician, 60*(4), 328-331. https://doi.org/10.1198/000313006X152649

Green, D. P., Goldman, S. L., & Salovey, P. (1993). Measurement error masks bipolarity in affect ratings. *Journal of Personality and Social Psychology, 64*(6), 1029-1041. https://psycnet.apa.org/doi/10.1037/0022-3514.64.6.1029

Grice, J. W. (2001). Computing and evaluating factor scores. *Psychological Methods, 6*(4), 430-450. https://psycnet.apa.org/doi/10.1037/1082-989X.6.4.430

Groen, R. N., Snippe, E., Bringmann, L. F., Simons, C. J. P., Hartmann, J. A., Bos, E. H., & Wichers, M. (2019). Capturing the risk of persisting depressive symptoms: A dynamic network investigation of patients' daily symptom experiences. *Psychiatry Research, 271*, 640-648. https://doi.org/10.1016/j.psychres.2018.12.054

Haslbeck, J. M. B. (2022). Estimating group differences in network models using moderation analysis. *Behavior Research Methods*, *54*, 522-540. https://doi.org/gmpfh6

Haslbeck, J. M. B., Borsboom, D., & Waldorp, L. J. (2021). Moderated network models. *Multivariate Behavioral Research, 56*(2), 256-287. https://doi.org/ghn6x5

Haslbeck, J. M. B., Ryan, O., Robinaugh, D. J., Waldorp, L. J., & Borsboom, D. (2022a). Modeling psychopathology: From data models to formal theories. *Psychological Methods, 27*(6), 930–957. https://doi.org/10.1037/met0000303

Haslbeck, J. M. B., Ryan, O., van der Maas, H. L. J., & Waldorp, L. J. (2022b). Modeling change in networks. In A. Isvoranu, S. Epskamp, L. Waldorp, & D. Borsboom (Eds.), *Network psychometrics with R* (pp. 192-209). Routledge.

Haslbeck, J. M. B., & Ryan, O. (2022). Recovering within-person dynamics from psychological time series. *Multivariate Behavioral Research*, *57*(5). https://doi.org/10.1080/00273171.2021.1896353

Haslbeck, J. M. B., & Waldorp, L. J. (2020). Estimating time-varying mixed graphical models in high-dimensional data. *Journal of Statistical Software, 93*(8), 1-46. https://doi.org/ggtsd2

Hayes, A. M., & Andrews, L. A. (2020). A complex systems approach to the study of change in psychotherapy. *BMC Medicine, 18*. https://doi.org/10.1186/s12916-020-01662-2

Heeren, A., & McNally, R. J. (2016). An integrative network approach to social anxiety disorder: The complex dynamic interplay among attentional bias for threat, attentional control, and symptoms. *Journal of Anxiety Disorders, 42*, 95-104. http://dx.doi.org/10.1016/j.janxdis.2016.06.009

Heeren, A., & McNally, R. J. (2018). Social anxiety disorder as a densely interconnected network of fear and avoidance for social situations. *Cognitive Therapy and Research, 42*, 103-113. https://doi.org/10.1007/s10608-017-9876-3

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

Hoffart, A., & Johnson, S. U. (2020a). Latent trait, latent-trait state, and a network approach to mental problems and their mechanisms of change. *Clinical Psychological Science*, 1-19. http://doi.org/ggzgf2

Hoffart, A., & Johnson, S. U. (2020b). Within-person networks of clinical features of social anxiety disorder during cognitive and interpersonal therapy. *Journal of Anxiety Disorders, 76*, https://doi.org/10.1016/j.janxdis.2020.102312

Hofmann, S. G., & Hayes, S. C. (2019). The future of intervention science: Process-based therapy. *Clinical Psychological Science, 7*(1), 37-50. http://doi.org/gdj5n5

Hofmann, S. G., Curtiss, J. E., & Hayes, S. C. (2020). Beyond linear mediation: Toward a dynamic network approach to study treatment processes. *Clinical Psychology Review, 76*. http://doi.org/ggh7wz

Hohensee, N., Meyer, M. J., & Teachman, B. A. (2020). The effect of confidence on dropout rate and outcomes in online cognitive bias modification. *Journal of Technology in Behavioral Science, 5*, 226-234. https://doi.org/10.1007/s41347-020-00129-8

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55. https://psycnet.apa.org/doi/10.1080/10705519909540118

Hussey, I., & Hughes, S. (2020). Hidden invalidity among 15 commonly used measures in social and personality research. *Advances in Methods and Practices in Psychological Science*, 1-19. https://doi.org/10.1177/2515245919882903

Isvoranu, A-M., & Espkamp, S. (2021). Which estimation method to choose in network psychometrics? Deriving guidelines for applied researchers. *Psychological Methods.* http://dx.doi.org/10.1037/met0000439

Ji, J. L., Baee, S., Zhang, D., Calicho-Mamani, C. P., Meyer, M. J., Funk, D., Portnow, S., Barnes, L., & Teachman, B. A. (2021). Multi-session online interpretation bias training for anxiety in a community sample. *Behaviour Research and Therapy, 142*. https://doi.org/g62s

Ji, J. L., Salemink, E., & Teachman, B. A. (2024). Association between interpretation flexibility and emotional health in an anxious sample: The challenge of measuring flexible adoption of multiple perspectives. *Journal of Experimental Psychopathology*. https://doi.org/10.1177/20438087241226642

Jones, P. J., Heeren, A., & McNally, R. J. (2017). Commentary: A network theory of mental disorders. *Fronteirs in Psychology, 8*. https://doi.org/10.1002/wps.20375

Jordan, D. G., Winer, E. S., & Salem, T. (2020). The current status of temporal network analysis for clinical science: Considerations as the paradigm shifts? *Journal of Clinical Psychology*, 1-22. https://doi.org/10.1002/jclp.22957

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2022). semTools: Useful tools for structural equation modeling [Computer software]. https://CRAN.R-project.org/package=semTools

Kelley, K. (2023). MBESS: The MBESS R package [Computer software]. https://CRAN.R-project.org/package=MBESS

Kyriazos, T. A. (2018). Applied psychometrics: Sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology*, *9*, 2207-2230. https://doi.org/10.4236/psych.2018.98126

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology, 49*(4), 764-766. https://psycnet.apa.org/doi/10.1016/j.jesp.2013.03.013

Larrazabal, M. A., Eberle, J. W., Vela de la Garza Evia, A., Boukhechba, M., Funk, D. H., Barnes, L. E., Boker, S. M., & Teachman, B. A. (2024). Online cognitive bias modification for interpretation to reduce anxious thinking during the COVID-19 pandemic. *Behaviour Research and Therapy, 173*. https://doi.org/mdg9

Lim, S., & Jahng, S. (2019). Determining the number of factors using parallel analysis and its recent variants. *Psychological Methods, 24*(4), 452. https://doi.org/gf4jsz

Loh, P., & Wainwright, M. J. (2013). Structure estimation for discrete graphical models: Generalized covariance matrices and their inverses. *The Annals of Statistics, 41*(6), 3022-3049. https://doi.org/gjs9pt

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, *33*(3), 335–343. https://doi.org/10.1016/0005-7967(94)00075-U

MacLeod, C., & Mathews, A. (2012). Cognitive bias modification approaches to anxiety. *Annual Review of Clinical Psychology*, *8*, 189-217. https://doi.org/grzp

Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors? *Journal of Personality and Social Psychology, 70*(4), 810.

https://psycnet.apa.org/doi/10.1037/0022-3514.70.4.810

Mathews, A., & Mackintosh, B. (2000). Induced emotional interpretation bias and anxiety. *Journal of Abnormal Psychology, 109*(4), 602-615. https://doi.org/ccgdjs

McNally, R. J. (2019). Attentional bias for threat: Crisis or opportunity? *Clinical Psychology Review, 69*, 4-13. https://doi.org/10.1016/j.cpr.2018.05.005

McNeish, D. (2023). Dynamic fit index cutoffs for categorical factor analysis with Likert-type, ordinal, or binary responses. *American Psychologist, 78*(9), 1061–1075. https://doi.org/10.1037/amp0001213

McNeish, D., & Wolf, M. G. (2024). Direct discrepancy dynamic fit index cutoffs for arbitrary covariance structure models. *Structural Equation Modeling: A Multidisciplinary Journal*, 1-28. https://doi.org/10.1080/10705511.2024.2308005

Morin, A. J., Myers, N. D., & Lee, S. (2020). Modern factor analytic techniques: Bifactor models, exploratory structural equation modeling (ESEM), and bifactor-ESEM. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of sport psychology* (4th ed., pp. 1044-1073). Wiley. https://doi.org/10.1002/9781119568124.ch51

Nieuwenhuis, S., Forstmann, B. U., & Wagenmakers, E. (2011). Erroneous analyses of interactions in neuroscience: A problem of significance. *Nature Neuroscience, 14*(9), 1105-1107. https://doi.org/10.1038/nn.2886

Norman, S. B., Hami Cissell, S., Means-Christensen, A. J., & Stein, M. B. (2006). Development and validation of an Overall Anxiety Severity and Impairment Scale (OASIS). *Depression and Anxiety, 23*(4), 245-249. https://doi.org/10.1002/da.20182

O'Driscoll, C., Epskamp, S., Fried, E. I., Saunders, R., Cardoso, A., Stott, J., Wheatley, J., Cirkovic, M., Naqvi, S. A., Buckman, J. E. J., & Pilling, S. (2022). Transdiagnostic symptom dynamics during psychotherapy. *Scientific Reports, 12*. https://doi.org/h63r

Olthof, M., Hasselman, F., Maatman, F. O., Bosman, A. M. T., & Lichtwarck-Aschoff, A. (2023). Complexity theory of psychopathology. *Journal of Psychopathology and Clinical Science, 132*(3), 314–323. https://psycnet.apa.org/doi/10.1037/abn0000740

Parsons, S., Songco, A., Booth, C., & Fox, E. (2021). Emotional information-processing correlates of positive mental health in adolescence: A network analysis approach. *Cognition and Emotion, 35*(5), 956-969. https://doi.org/10.1080/02699931.2021.1915752

Papini, S., Dainer-Best, J., Rubin, M., Zaizar, E. D., Telch, M. J., & Smits, J. A. J. (2020). Psychological networks can identify potential pathways to specific intervention targets for anxiety in response to Coronavirus Disease 2019 (COVID-19) [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/4zr25

Pe, M. L., Kircanski, K., Thompson, R. J., Bringmann, L. F., Tuerlinckx, F., Mestdagh, M., Mata, J., Jaeggi, S. M., Buschkuehl, M., Jonides, J., Kuppens, P., & Gotlib, I. H. (2015). Emotion-network density in major depressive disorder. *Clinical Psychological Science, 3*(2), 292-300. https://doi.org/10.1177/2167702614540645

Podsakoff, P. M., Podsakoff, N. P., Williams, L. J., Huang, C., & Yang, J. (2024). Common method bias: it's bad, it's complex, it's widespread, and it's not easy to fix. *Annual Review of Organizational Psychology and Organizational Behavior, 11*, 17-61. https://doi.org/10.1146/annurev-orgpsych-110721-040030

R Core Team (2023). R: A language and environment for statistical computing [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

R Core Team (2024). R: A language and environment for statistical computing [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Reuland, M. M., & Teachman, B. A. (2014). Interpretation bias modification for youth and their parents: A novel treatment for early adolescent social anxiety. *Journal of Anxiety Disorders*, *28*(8), https://doi.org/10.1016/j.janxdis.2014.09.011

Revelle, W. (2023). psych: Procedures for Psychological, Psychometric, and Personality Research [Computer software]. https://cran.r-project.org/package=psych

Revelle, W., & Condon, D. M. (2019). Reliability from α to ω: A tutorial. *Psychological Assessment, 31*(12), 1395-1411. https://psycnet.apa.org/doi/10.1037/pas0000754

Robinaugh, D. J., Millner, A. J., & McNally, R. J. (2016). Identifying highly influential nodes in the complicated grief network. *Journal of Abnormal Psychology, 125*(6), 747-757. https://doi.org/10.1037/abn0000181

Robinaugh, D. J., Hoekstra, R. H. A., Toner, E. R., & Borsboom, D. (2020). The network approach to psychopathology: A review of the literature 2008–2018 and an agenda for future research. *Psychological Medicine, 50*, 353-366. https://doi.org/gghndf

Rodebaugh, T. L., Scullin, R. B., Langer, J. K., Dixon, D. J., Huppert, J. D., Bernstein, A., & Zvielli, A. (2016). Unreliability as a threat to understanding psychopathology: The cautionary tale of attentional bias. *Journal of Abnormal Psychology, 125*(6), 850-851. https://doi.org/10.1037/abn0000184

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods, 21*(2), 137-150. https://psycnet.apa.org/doi/10.1037/met0000045

Rosellini, A. J., & Brown, T. A. (2021). Developing and validating clinical questionnaires. *Annual Review of Clinical Psychology, 17*, 55-81. https://doi.org/gh4pw3

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of*

Statistical Software, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Ryan, O., Bringmann, L. F., & Schuurman, N. K. (2022). The challenge of generating causal hypotheses using network models. *Structural Equation Modeling: A Multidisciplinary Journal*, *29*(6), 953-970. https://doi.org/10.1080/10705511.2022.2056039

Ryan, O., & Hamaker, E. L. (2022). Time to intervene: A continuous-time approach to network analysis and centrality. *Psychometrika*, *87*(1), 214-252. https://doi.org/gkx59v

Salemink, E., & van den Hout, M. (2010). Validation of the ‘‘recognition task’’ used in the training of interpretation biases. *Journal of Behavior Therapy and Experimental Psychiatry*, *41*, 140-144. https://doi.org/10.1016/j.jbtep.2009.11.006

Savalei, V. (2021). Improving fit indices in structural equation modeling with categorical data. *Multivariate Behavioral Research*, *56*(3), 390-407. https://doi.org/10.1080/00273171.2020.1717922

Schumacher, L., Burger, J., Echterhoff, J., & Kriston, L. (2022). *Methodological and statistical practices of using symptom networks to evaluate mental health interventions: A systematic review*. PsyArXiv. https://doi.org/10.31234/osf.io/ea7fu

Schmitt, T. N., Larrazabal, M. A., Eberle, J. W., Boukhechba, M., Barnes, L. E., & Teachman, B. A. (2023). *Relationship between interpretation bias, racial or ethnic identity, and COVID-19 anxiety*. PsyArXiv. https://doi.org/10.31234/osf.io/hxc9j

Silverman, A. L., Vela de la Garza Evia, A. F., Larrazabal, M. A., Khutsisvili, L., Boukhechba, M., Barnes, L. E., & Teachman, B. A. (2024). *Pilot feasibility and acceptability of a Spanish translation of web-based cognitive bias modification for interpretation among Spanish-speaking Latinx individuals with anxiety*. PsyArXiv. https://doi.org/10.31234/osf.io/3ck26

Slipetz, L. R., Eberle, J. W., Levinson, C. A., Falk, A., Cusack, C. E., & Henry, T. R. (2023). *Analyzing ecological momentary assessment data with state-space models: Considerations and recommendations.* ArXiv. https://doi.org/10.48550/arXiv.2305.03214

Smits, J. A. J., Julian, K., Rosenfield, D., & Powers, M. B. (2012). Threat reappraisal as a mediator of symptom change in cognitive-behavioral treatment of anxiety disorders: A systematic review. *Journal of Consulting and Clinical Psychology, 80*(4), 624-635. https://doi.org/10.1037/a0028957

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

Steinman, S. A., & Teachman, B. A. (2014). Reaching new heights: Comparing interpretation bias modification to exposure therapy for extreme height fear. *Journal of Consulting and Clinical Psychology, 82*(3), 404-417. https://doi.org/10.1037/a0036023

Swanson, T. J. (2020). *Modeling moderators in psychological networks* (Publication No. 28000912) [Doctoral dissertation, University of Kansas]. ProQuest Dissertations.

Teachman, B. A. (2014). No appointment necessary: Treating mental illness outside the therapist's office. *Perspectives on Psychological Science, 9*(1), 85-87. http://doi.org/f5sbkz

van Bork, R., Lunansky, G., & Borsboom. D. (2024). Measurement targets for network constructs in psychopathology. *Measurement, 232*. https://doi.org/10.1016/j.measurement.2024.114643

van Borkulo, R., Boschloo, L., Kossakowski, J. J., Tio, P., Schoevers, R. A., Borsboom, D., & Waldorp, L. J. (2022). Comparing network structures on three aspects: A permutation test. *Psychological Methods.* http://dx.doi.org/10.1037/met0000476

van den Hout, M. (2014). Psychiatric symptoms as pathogens. *Clinical Neuropsychiatry, 11*(6), 153-159. https://dspace.library.uu.nl/handle/1874/307019

Vela de la Garza Evia, A., Eberle, J. W., Baee, S., Wolfe, E. C., Boukhechba, M., Funk, D. H., Teachman, B. A., & Barnes, L. E. (2024). *Behavioral engagement and psychosocial outcomes in web-based interpretation bias training for anxiety.* PsyArXiv. https://doi.org/10.31234/osf.io/yucqf

Wang, L., & Maxwell, S. E. (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. *Psychological Methods, 20*(1), 63-83. https://psycnet.apa.org/doi/10.1037/met0000030

Watson, D. (2004). Stability versus change, dependability versus error: Issues in the assessment of personality over time. *Journal of Research in Personality, 38*, 319-350. https://doi.org/10.1016/j.jrp.2004.03.001

Wilshire, C. E., Ward, T., & Clack, S. (2021). Symptom descriptions in psychopathology: How well are they working for us? *Clinical Psychological Science, 9*(3), 323-339. https://doi.org/10.1177/2F2167702620969215

Williams, D. R. (2021). *The confidence interval that wasn't: Bootstrapped "confidence intervals" in l1-regularized partial correlation networks.* PsyArXiv. https://doi.org/jdcx

Williams, D. R. (2022). Learning to live with sampling variability: Expected replicability in partial correlation networks. *Psychological Methods, 27*(4). https://psycnet.apa.org/doi/10.1037/met0000417

Williams, D. R., Briganti, G., Linkowski, P., & Mulder, J. (2021). *On accepting the null hypothesis of conditional independence in partial correlation networks: A Bayesian* *analysis*. PsyArXiv. https://doi.org/10.31234/osf.io/7uhx8

Williams, D. R., & Rast, P. (2020). Back to the basics: Rethinking partial correlation network methodology. *British Journal of Mathematical and Statistical Psychology, 73*, 187-212. https://doi.org/10.1111/bmsp.12173

Williams, D. R., Rhemtulla, M., Wysocki, A. C., & Rast, P. (2019). On nonregularized estimation of psychological networks. *Multivariate Behavioral Research, 54*(5), 719-750. https://doi.org/10.1080/00273171.2019.1575716

Williams, D. R., & Rodriguez, J. E. (2022). Why overfitting is not (usually) a problem in partial correlation networks. *Psychological Methods, 27*(5), 822-840. https://doi.org/10.1037/met0000437

Table 1

Demographic Characteristics

Characteristic	Value
n	749
Age	
Years: *M* (*SD*)	33.82 (13.56)
Prefer not to answer or missing: *n* (%)	2 (0.3)
Gender: *n* (%)	
Female	542 (72.4)
Male	187 (25.0)
Transgender	6 (0.8)
Other	8 (1.1)
Prefer not to answer or missing	6 (0.8)
Race: *n* (%)	
American Indian/Alaska Native	5 (0.7)
Black/African origin	20 (2.7)
East Asian	30 (4.0)
Native Hawaiian/Pacific Islander	1 (0.1)
South Asian	32 (4.3)
White/European origin	582 (77.7)
Other or Unknown	47 (6.3)
Prefer not to answer or missing	32 (4.3)
Ethnicity: *n* (%)	
Hispanic or Latino	56 (7.5)
Not Hispanic or Latino	622 (83.0)
Unknown	28 (3.7)
Prefer not to answer or missing	43 (5.7)
Country: *n* (%)	
United States	575 (76.8)
Canada	34 (4.5)
United Kingdom	27 (3.6)
Other [a]	110 (14.7)
Prefer not to answer or missing	3 (0.4)
Education: *n* (%)	
Elementary School	0 (0.0)
Junior High	0 (0.0)
Some High School	13 (1.7)
High School Graduate	48 (6.4)
Some College	173 (23.1)
Associate’s Degree	45 (6.0)
Bachelor’s Degree	184 (24.6)
Some Graduate School	57 (7.6)
Master’s Degree	142 (19.0)
M.B.A.	13 (1.7)
J.D.	9 (1.2)

Characteristic	Value
M.D.	6 (0.8)
Ph.D.	34 (4.5)
Other Advanced Degree	14 (1.9)
Prefer not to answer or missing	11 (1.5)
Employment Status: *n* (%)	
Student	218 (29.1)
Homemaker	22 (2.9)
Unemployed or laid off	17 (2.3)
Looking for work	34 (4.5)
Working part-time	85 (11.3)
Working full-time	315 (42.1)
Retired	22 (2.9)
Other	25 (3.3)
Prefer not to answer or missing	11 (1.5)
Annual Income: *n* (%)	
Less than $5,000	45 (6.0)
$5,000 through $11,999	45 (6.0)
$12,000 through $15,999	25 (3.3)
$16,000 through $24,999	41 (5.5)
$25,000 through $34,999	49 (6.5)
$35,000 through $49,999	72 (9.6)
$50,000 through $74,999	88 (11.7)
$75,000 through $99,999	82 (10.9)
$100,000 through $149,999	72 (9.6)
$150,000 through $199,999	28 (3.7)
$200,000 through $249,999	16 (2.1)
$250,000 or greater	17 (2.3)
Unknown	76 (10.1)
Prefer not to answer or missing	93 (12.4)
Marital Status: *n* (%)	
Single	244 (32.6)
Dating	113 (15.1)
Engaged	29 (3.9)
In marriage-like relationship	64 (8.5)
Married	227 (30.3)
In domestic or civil union	9 (1.2)
Separated	5 (0.7)
Divorced	31 (4.1)
Widow/widower	7 (0.9)
Other	4 (0.5)
Prefer not to answer or missing	16 (2.1)

Note. Each characteristic has missing data for 1-3 participants due to a server error.

[a] Countries with fewer than 10 participants were collapsed into Other due to space constraints.

Table 2

Fit of Initial Confirmatory Factor Analysis Models Based on All 36 Threat and Nonthreat Items

Model	χ^2	*df*	*P*	SRMR	RMSEA	CFI	TLI
1. 4 correlated factors (pos. threat, pos. nonthreat, neg. threat, neg. nonthreat)	2,674.10	588	< .001	.091	0.098	.726	0.707
2. Bifactor with 2 general factors (pos., neg.), each with 2 specific factors (threat, nonthreat)	3,085.98	557	< .001	.090	0.098	.743	0.709
3. Higher-order with 2 second-order factors (pos., neg.), each with 2 first-order factors (threat, nonthreat)	Improper solution						

Note. Models were fit to polychoric correlations using diagonally weighted least squares (DWLS) estimation with robust standard errors and a mean- and variance-adjusted χ^2 (`WLSMV` estimator with pairwise deletion). Robust RMSEA, CFI, and TLI (Savalei, 2021) are shown. Traditional guidelines for "relatively good" fit are nonsignificant χ^2, SRMR near or < .08, RMSEA near or < 0.06, and CFI and TLI near or > .95 (Hu & Bentler, 1999). No index met the guidelines. SRMR = standardized root mean square residual; RMSEA = root mean square error of approximation; CFI = comparative fit index; TLI = Tucker-Lewis index.

Table 3

Fit of Revised Confirmatory Factor Analysis (CFA) Models Based on 28 Threat and Nonthreat Items

Model	χ^2	*df*	*P*	SRMR	RMSEA	CFI	TLI
EFA in CFA Model							
4. 3 correlated factors (pos. threat, neg. threat, nonthreat)	817.49	297	< .001	**.043**	0.078	.873	0.838
CFA Models							
5. 3 correlated factors (pos. threat, neg. threat, nonthreat)	1,449.00	347	< .001	**.085**	0.092	.794	0.776
6. 2 correlated factors (neg. threat, not neg. threat)	2,126.68	349	< .001	.099	0.103	.743	0.722
7. 3 correlated factors (pos. threat, neg. threat, nonthreat) and correlated errors per scenario for threat items	1,293.94	343	< .001	**.082**	0.083	.836	0.819
8. 3 correlated factors (pos. threat, neg. threat, nonthreat) and correlated errors per scenario for all items	1,139.67	316	< .001	**.075**	0.082	.850	0.821

Note. Models were fit to polychoric correlations using diagonally weighted least squares (DWLS) estimation with robust standard errors and a mean- and variance-adjusted χ^2 (`WLSMV` estimator with pairwise deletion). Robust RMSEA, CFI, and TLI (Savalei, 2021) are shown. Traditional guidelines for "relatively good" fit are nonsignificant χ^2, SRMR near or < .08, RMSEA near or < 0.06, and CFI and TLI near or > .95 (Hu & Bentler, 1999). Indices meeting the guidelines are in boldface. EFA = exploratory factor analysis; SRMR = standardized root mean square residual; RMSEA = root mean square error of approximation; CFI = comparative fit index; TLI = Tucker-Lewis index.

Table 4

Fit of Revised Confirmatory Factor Analysis (CFA) Models Based on All 18 Threat Items

Model	χ^2	*df*	*P*	SRMR	RMSEA	CFI	TLI
EFA in CFA Model							
9. 2 correlated factors (pos. threat, neg. threat)	1,179.46	118	< .001	**.083**	0.138	.651	0.548
CFA Models							
10. 2 correlated factors (pos. threat, neg. threat)	1,068.64	134	< .001	.091	0.132	.637	0.586
11. 2 correlated factors (pos. threat, neg. threat) and correlated errors per scenario	547.39	125	< .001	**.067**	0.081	.873	0.844
12. 1 factor (threat) and correlated errors per scenario	2,148.21	126	< .001	.135	0.159	.500	0.393
13. 1 factor (threat), 2 correlated method factors (pos., neg.), and correlated errors per scenario	323.23	107	< .001	**.046**	**0.066**	**.928**	0.896
14. 1 factor (threat), 1 method factor (pos.), and correlated errors per scenario	512.49	117	< .001	**.059**	0.076	.893	0.861
15. 1 factor (threat), 1 method factor (neg.), and correlated errors per scenario	560.38	117	< .001	**.065**	0.082	.879	0.841
16. Bifactor with 1 general factor (threat), 2 specific factors (pos., neg.), and correlated errors per scenario	366.15	108	< .001	**.049**	**0.066**	**.926**	0.894
17. Higher-order with 1 second-order factor (threat), 2 first-order factors (pos., neg.), and correlated errors per scenario	Improper solution						

Note. Models were fit to polychoric correlations using diagonally weighted least squares (DWLS) estimation with robust standard errors and a mean- and variance-adjusted χ^2 (`WLSMV` estimator with pairwise deletion). Robust RMSEA, CFI, and TLI (Savalei, 2021) are shown. Traditional guidelines for "relatively good" fit are nonsignificant χ^2, SRMR near or < .08, RMSEA near or < 0.06, and CFI and TLI near or > .95 (Hu & Bentler, 1999). Indices meeting the guidelines are in boldface. EFA = exploratory factor analysis; SRMR = standardized root mean square residual; RMSEA = root mean square error of approximation; CFI = comparative fit index; TLI = Tucker-Lewis index.

Table 5

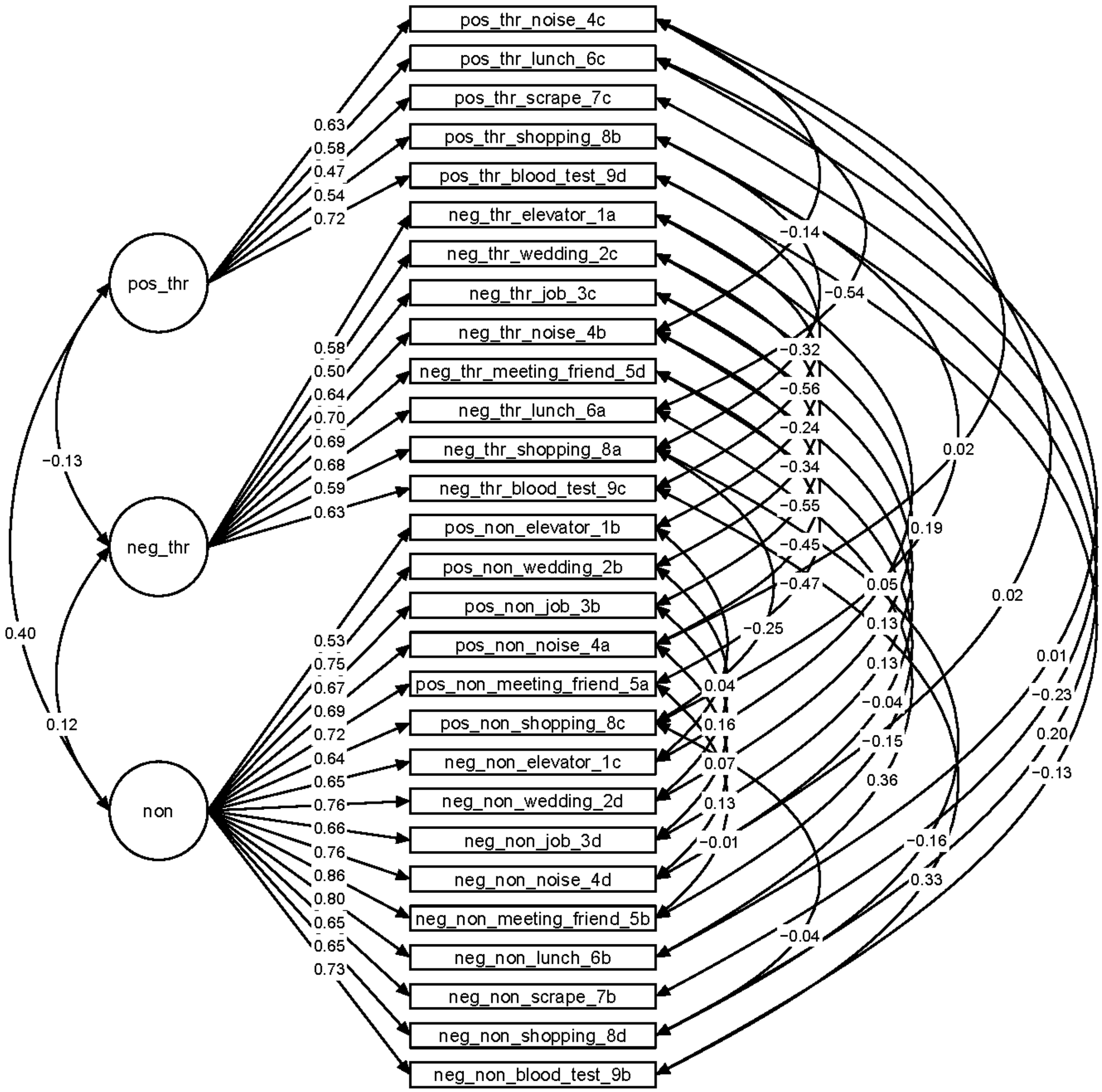
Connectivity Metrics for Saturated Temporal and Contemporaneous Networks from Multigroup GVAR Model

	Condition			Difference	
Metric	Pos. CBM-I	50-50 CBM-I	No-Training	Pos. CBM-I – 50-50 CBM-I	Pos. CBM-I – No-Training
Mean Strength					
Internode	.11	.16	.14	-.05	-.03
Intranode	.15	.19	.10	-.04	.05
Overall	.12	.17	.13	-.05	-.02
Mean Expected Influence					
Internode	-.02	.07	.05	-.09	-.07
Intranode	.03	.11	.05	-.08	-.02
Overall	-.01	.07	.05	-.09	-.07
Mean Strength					
Internode	.17	.24	.24	-.06	-.07
Mean Expected Influence					
Internode	.09	.11	.13	-.02	-.04

Note. Connectivity metrics were computed using both significant and nonsignificant edges.

Figure 1

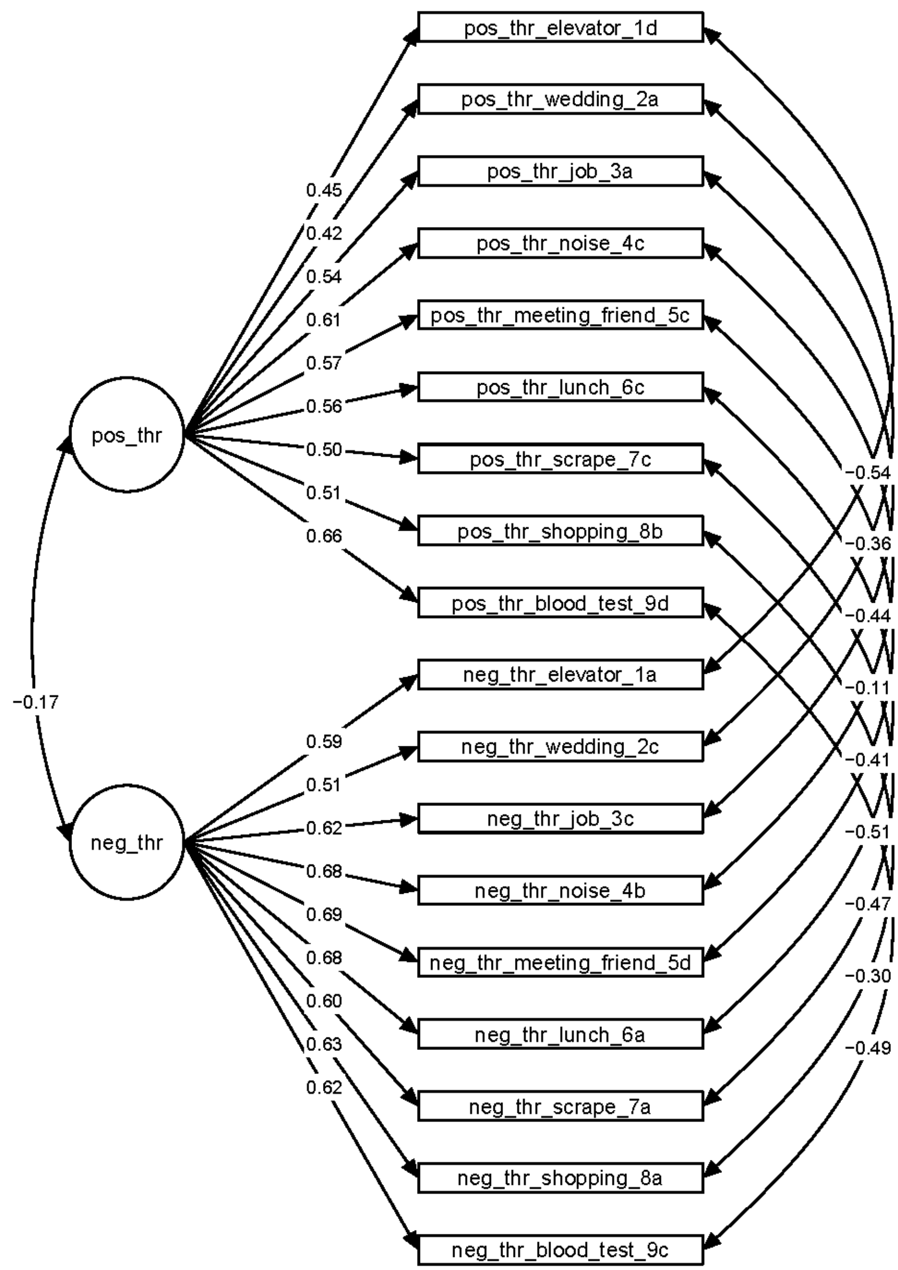
Model 8: CFA Model With 3 Correlated Factors (Positive Threat, Negative Threat, Nonthreat) and Correlated Errors per Scenario for All Items, Based on 28 Threat and Nonthreat Items



Note. Completely standardized estimates are shown. Model was fit to polychoric correlations using diagonally weighted least squares estimation (`WLSMV` estimator with pairwise deletion).

Figure 2

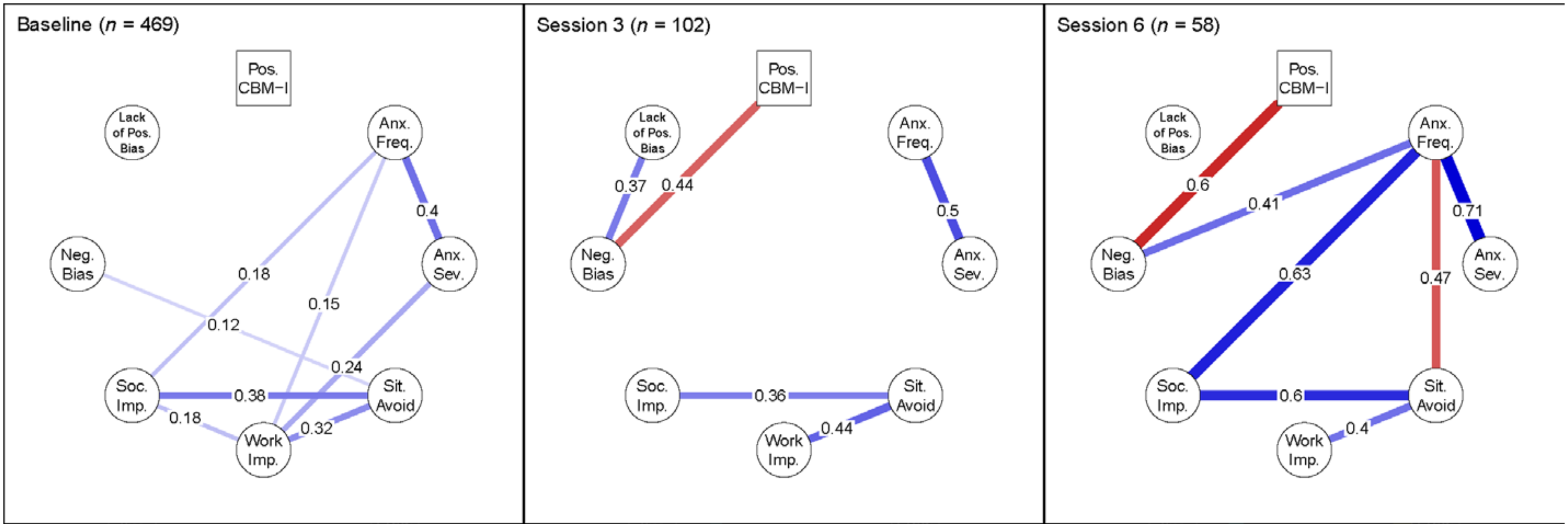
Model 11: CFA Model With 2 Correlated Factors (Positive Threat, Negative Threat) and Correlated Errors per Scenario, Based on All 18 Threat Items



Note. Completely standardized estimates are shown. Model was fit to polychoric correlations using diagonally weighted least squares estimation (`WLSMV` estimator with pairwise deletion).

Figure 3

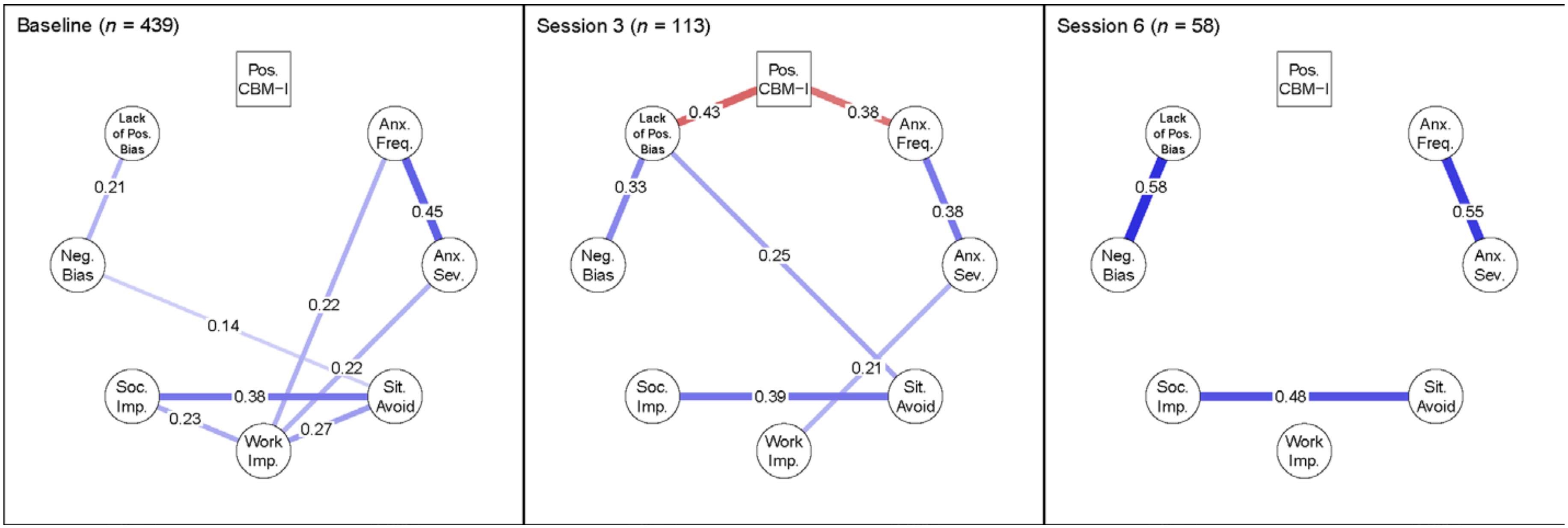
*Cross-Sectional Mixed Graphical Models at Each Time Point With Node Contrasting Positive CBM-I and 50-50 CBM-I for Intent-To-Treat Sample, Showing Thresholded Edges (*p *< .01)*



Note. Only edges significant at *p* < .01 threshold are shown. Sign of edge weight is mapped to edge color, with positive weights in blue and negative weights in red. Magnitude of edge weight is mapped to edge width and color saturation, with maximum width and saturation mapped to maximum edge magnitude (.71) across all mixed graphical models. All nodes were modeled as manifest. Plots were generated in `qgraph` using a circle layout. Pos. CBM-I = positive CBM-I (vs. 50-50 CBM-I; coded 1 and 0, respectively). Anx. Freq. = anxiety frequency; Anx. Sev. = anxiety severity; Sit. Avoid = situational avoidance; Work Imp. = work impairment; Soc. Imp. = social impairment; Neg. Bias = negative bias; Lack of Pos. Bias = lack of positive bias (i.e., reverse-scored positive bias).

Figure 4

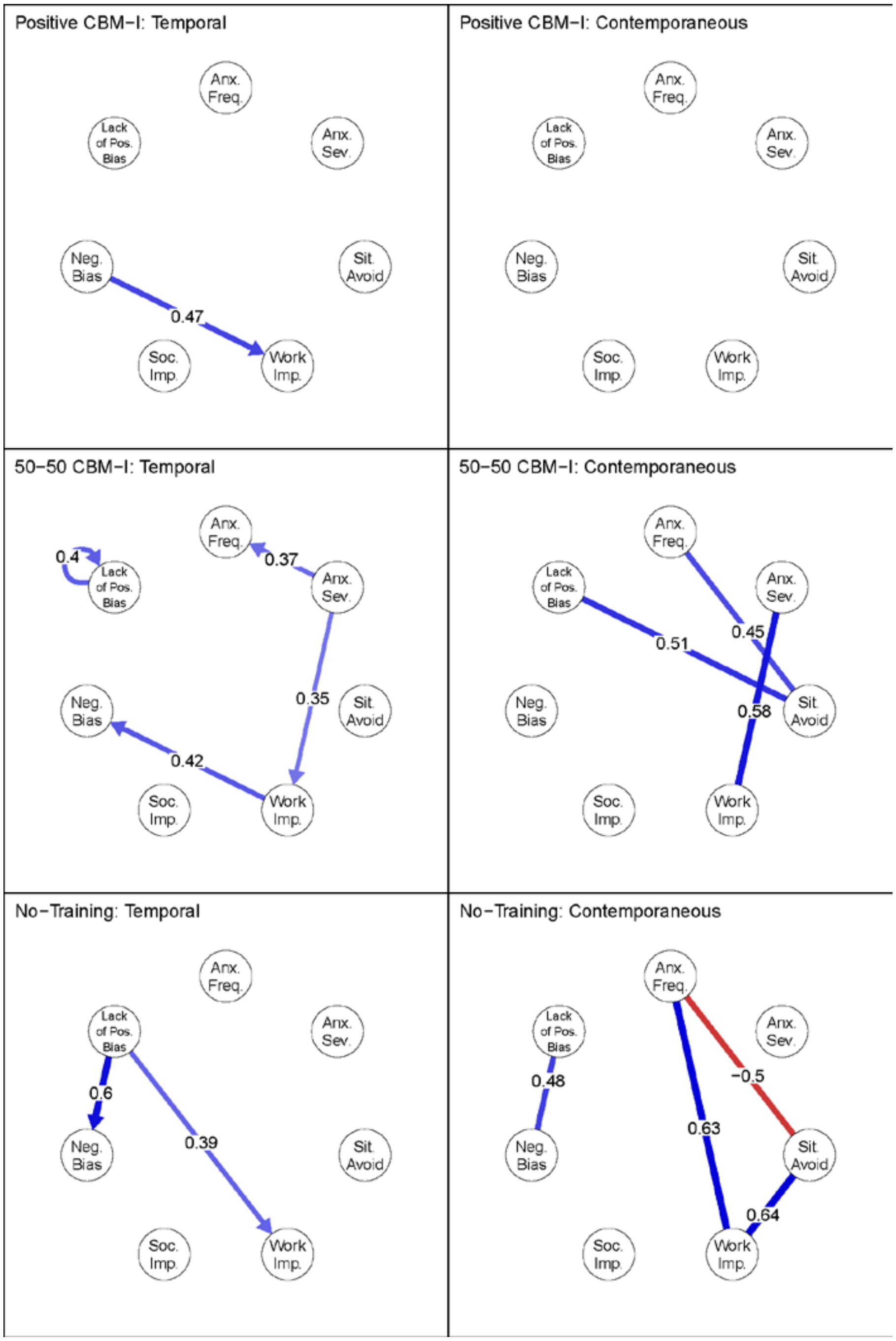
*Cross-Sectional Mixed Graphical Models at Each Time Point With Node Contrasting Positive CBM-I and No-Training for Intent-To-Treat Sample, Showing Thresholded Edges (*p *< .01)*



Note. Only edges significant at $p < .01$ threshold are shown. Sign of edge weight is mapped to edge color, with positive weights in blue and negative weights in red. Magnitude of edge weight is mapped to edge width and color saturation, with maximum width and saturation mapped to maximum edge magnitude (.71) across all mixed graphical models. All nodes were modeled as manifest. Plots were generated in `qgraph` using a circle layout. Pos. CBM-I = positive CBM-I (vs. no-training; coded 1 and 0, respectively). Anx. Freq. = anxiety frequency; Anx. Sev. = anxiety severity; Sit. Avoid = situational avoidance; Work Imp. = work impairment; Soc. Imp. = social impairment; Neg. Bias = negative bias; Lack of Pos. Bias = lack of positive bias (i.e., reverse-scored positive bias).

Figure 5

*Multigroup GVAR Model for Intent-To-Treat Sample, Showing Thresholded Edges (*p *< .01)*



Note. Only edges significant at $p < .01$ are shown. Edge weight (positive in blue, negative in red) is mapped to edge width and color saturation, with maximum width and saturation mapped to maximum edge magnitude (.64) across conditions. All nodes were modeled as manifest. Plots use `qgraph`'s circle layout. Anx. Freq. = anxiety frequency; Anx. Sev. = anxiety severity; Sit. Avoid = situational avoidance; Work Imp. = work impairment; Soc. Imp. = social impairment; Neg. Bias = negative bias; Lack of Pos. Bias = lack of positive bias (i.e., reverse-scored positive bias).

Appendix A

Modified Recognition Ratings

(The measure used in the Managing Anxiety trial is shown below. Modified wording for two scenarios used in Mathews & Mackintosh, 2000 [in boldface] are shown in footnotes.)

Scenarios and Comprehension Questions

In this part of the assessment, you will read a series of very short stories. Pay attention to the title of each story because after you have read all the stories, you will be asked more questions about them.

For each story:

- Read each sentence carefully and really imagine yourself in the situation described.
- Even if the story describes you reacting in a way that you would not usually react, please try to picture yourself responding in the way the story describes.
- There will be an incomplete word at the end of each paragraph.
- Press the key on the keyboard that complete the word.
- When you correctly complete the word you will move on to the next screen and be asked a question about the story.

THE ELEVATOR
You are in the lobby of your friend's new apartment building. You press the button to the elevator to go up. The building looks old. As you get on the elevator you think about its…s_fety.

Did you think about the elevator's safety? Yes/No

THE WEDDING RECEPTION
Your friend asks you to give a speech at her wedding reception. You prepare some remarks and when the time comes, get to your feet. As you speak, you notice some people in the audience start to…l_ugh.

Does the audience laugh when you speak? Yes/No[16]

THE JOB
You are currently working as a contractor for a company. Once this job is finished, you will be without employment until you can find your next job. You think about not having an income for a few weeks and about your future…fin_nces.

Will you be without an income soon? Yes/No

[16] Mathews and Mackintosh (2000) use a different comprehension question: "Did you stand up to speak?"

THE LOUD NOISE
You are woken up in the middle of the night by a loud noise. You are not sure what caused the noise and leave your bedroom to see what happened. You walk…downs_airs.

Have you been woken up in the middle of the night? Yes/No

MEETING A FRIEND
In the street you bump into an old friend you haven't seen for a long time. She is too busy to stop, so you arrange to meet later in a bar. You arrive a little late but the bar is empty and a few minutes later she is still not…th_re.

Did your friend stop? Yes/No[17]

THE LUNCH
You are eating lunch with a friend. As you start eating your salad, you describe your plans for the weekend. You accidentally drop a piece of lettuce, and your friend looks at…y_u.

Do you have salad for lunch? Yes/No

THE SCRAPE
You are playing basketball with some friends. While running toward the ball, you trip and scrape your knee. The scrape hurts a bit, but does not…bl_ed.

Are you playing soccer with your friend? Yes/No

THE SHOPPING TRIP
You are at the mall with your friend. While you shop, she tells you how several of her friends have recently come down with a strange illness. You think about your recent health, and wonder if you will get…si_k.

Are you shopping alone? Yes/No

THE BLOOD TEST
You are at a routine doctor's appointment. At the appointment, the doctor decides to run a few blood tests to check your health. The doctor says he will call you in a few weeks, and you will find out your test results at that…t_me.

Did you take blood tests? Yes/No

[17] Mathews and Mackintosh (2000) use a different comprehension question: “Was anyone else in the bar?”

Recognition Items

Instructions:

Below you will see the titles of the nine short stories you just read and four sentences about each story. None of these sentences are worded identically to the sentences in the stories. Please rate each sentence for how similar it is to the way you understood the original story.

THE ELEVATOR: The building looks old, and as you get on the elevator…

- You think that the elevator will probably break down while you are on it.
- You think that you are going to like your friend's new apartment.
- You think about how smelly the lobby is.
- You think that riding the elevator will be safe.

THE WEDDING RECEPTION: As you enter the room…[18]

- People in the audience laugh appreciatively.
- You notice a friend you were hoping to see walk into the reception.[19]
- People in the audience find your efforts laughable.
- You notice someone you do not like just walked into the reception.[20]

THE JOB: You think about not having an income for a few weeks...

- And know that you can rely on your savings.
- And are excited about not having to set an alarm.
- And worry about becoming broke.
- And are sad about leaving your current coworkers.

THE LOUD NOISE: As you walk downstairs…

- You feel happy, and think about how lovely your house is.
- You feel afraid, and worry that you cannot handle the fear.
- You feel afraid, but you know that you can tolerate the feeling.
- You feel cold, and think about how the house needs better heating.

MEETING A FRIEND: You arrive a little late, and…[21]

- Order your favorite snack.[22]

---

[18] In Mathews and Mackintosh (2000), each item is instead preceded by "As you speak,"
[19] Mathews and Mackintosh (2000) use a different item: "People in the audience applaud your comments."
[20] Mathews and Mackintosh (2000) use a different item: "Some people in the audience start to yawn in boredom."
[21] In Mathews and Mackintosh (2000), no items are preceded by this phrase.
[22] Mathews and Mackintosh (2000) use another item: "You are busy but your friend insists on meeting you in a bar."

- Notice the bar smells gross.[23]
- Get a call from your friend who is on her way, but running late.[24]
- Think your friend decided she did not want to see you.[25]

THE LUNCH: Your friend looks at you…

- Because she thinks you are a slob.
- And you frown because you forgot to bring water to lunch.
- Because she is paying attention as you describe your weekend plans.
- And you smile because your lunch tastes good.

THE SCRAPE: The scrape hurts a bit...

- And you think it will probably get seriously infected.
- And you are frustrated because you tore your shorts.
- But you know you will be okay.
- But you are happy that you are getting exercise.

THE SHOPPING TRIP: You think about your recent health...

- And think you are probably coming down with the strange illness.
- And think you are unlikely to catch the strange illness.
- And smile because you enjoy shopping.
- And feel bored of shopping.

THE BLOOD TEST: The doctor says he will call you in a few weeks...

- And you think about how nice your doctor is.
- And you are annoyed because your doctor is not very friendly.
- And you think that you will not be able to tolerate your anxiety while you wait.
- And you know that you can handle your anxiety while you wait.

[23] Mathews and Mackintosh (2000) use a different item: "Your friend tells you that she does not want to meet you."
[24] In Mathews and Mackintosh (2000), this item is "You arrange to meet a friend in a bar but your friend is late."
[25] In Mathews and Mackintosh (2000), this item is "You arrange to meet in a bar but your friend stands you up."

Appendix B

Overall Anxiety Severity and Impairment Scale

(The measure used in the Managing Anxiety trial, matching that from Norman et al., 2006, is shown below.)

Instructions:

The following items ask about anxiety and fear. For each item, circle the number for the answer that best describes your experience over the past week.

In the past week, how often have you felt anxious?

- No anxiety in the past week.
- Infrequent anxiety. Felt anxious a few times.
- Occasional anxiety. Felt anxious as much of the time as not. It was hard to relax.
- Frequent anxiety. Felt anxious most of the time. It was very difficult to relax.
- Constant anxiety. Felt anxious all of the time and never really relaxed.

In the past week, when you have felt anxious, how intense or severe was your anxiety?

- Little or None: Anxiety was absent or barely noticeable.
- Mild: Anxiety was at a low level. It was possible to relax when I tried. Physical symptoms were only slightly uncomfortable.
- Moderate: Anxiety was distressing at times. It was hard to relax or concentrate, but I could do it if I tried. Physical symptoms were uncomfortable.
- Severe: Anxiety was intense much of the time. It was very difficult to relax or focus on anything else. Physical symptoms were extremely uncomfortable.
- Extreme: Anxiety was overwhelming. It was impossible to relax at all. Physical symptoms were unbearable.

In the past week, how often did you avoid situations, places, objects, or activities because of anxiety or fear?

- None: I do not avoid places, situations, activities, or things because of fear.
- Infrequent: I avoid something once in a while, but will usually face the situation or confront the object. My lifestyle is not affected.
- Occasional: I have some fear of certain situations, places, or objects, but it is still manageable. My lifestyle has only changed in minor ways. I always or almost always avoid the things I fear when I’m alone, but can handle them if someone comes with me.
- Frequent: I have considerable fear and really try to avoid the things that frighten me. I have made significant changes in my life style to avoid the object, situation, activity, or place.
- All the Time: Avoiding objects, situations, activities, or places has taken over my life. My lifestyle has been extensively affected and I no longer do things that I used to enjoy.

In the past week, how much did your anxiety interfere with your ability to do the things you needed to do at work, at school, or at home?

- None: No interference at work/home/school from anxiety
- Mild: My anxiety has caused some interference at work/home/school. Things are more difficult, but everything that needs to be done is still getting done.
- Moderate: My anxiety definitely interferes with tasks. Most things are still getting done, but few things are being done as well as in the past.
- Severe: My anxiety has really changed my ability to get things done. Some tasks are still being done, but many things are not. My performance has definitely suffered.
- Extreme: My anxiety has become incapacitating. I am unable to complete tasks and have had to leave school, have quit or been fired from my job, or have been unable to complete tasks at home and have faced consequences like bill collectors, eviction, etc.

In the past week, how much has anxiety interfered with your social life and relationships?

- None: My anxiety doesn’t affect my relationships.
- Mild: My anxiety slightly interferes with my relationships. Some of my friendships and other relationships have suffered, but, overall, my social life is still fulfilling.
- Moderate: I have experienced some interference with my social life, but I still have a few close relationships. I don’t spend as much time with others as in the past, but I still socialize sometimes.
- Severe: My friendships and other relationships have suffered a lot because of anxiety. I do not enjoy social activities. I socialize very little.
- Extreme: My anxiety has completely disrupted my social activities. All of my relationships have suffered or ended. My family life is extremely strained.