# Title

# Abstract

# Introduction

Cognitive biases are important in the development and maintenance of emotional disorders (for reviews, see (Cisler & Koster, 2010; Gotlib & Joormann, 2010; Mathews & MacLeod, 1994, 2005). Automatic tendencies to selectively process negative, relative to benign or positive material, is associated with anxiety and depression, and these biases have been documented in attention, interpretation of ambiguity, as well as in memory. Studies examining selective processing biases in relation to emotional vulnerability have tended to examine a single process in isolation with few studies examining more than one bias in a single study (Everaert, Koster, & Derakshan, 2012; Hirsch, Clark, & Mathews, 2006). The combined cognitive bias hypothesis (CCBH: Hirsch et al., 2006), however, proposes that cognitive biases do not work in isolation to influence emotional vulnerability, but rather, they influence each other and interact to influence other variables, including emotional vulnerability. For example, increased attentional bias towards threatening stimuli might influence the degree to which less clearly negative information is interpreted more negatively, which in turn could plausibly influence memory for that stimulus. Such a series of causally related negative biases would be expected to propagate emotional vulnerability. The CCBH also suggests that targeting one process in order to produce a change in another is one way in which psychological interventions (e.g., cognitive behaviour therapy, cognitive bias modification) could be utilized to reduce emotional vulnerability. One study conducted in adults, for instance, concluded that memory bias may be more effectively modified by targeting emotional processing in another domain, such as interpretation bias (Vrijsen, van Oostrom, Isaac, Becker, & Speckens, 2014, also see Hertel & Mathews, 2011). A further study in adults investigated the functional relationships among cognitive biases in a subclinical depressed sample and found that while attentional bias was not directly associated with memory bias, there was an indirect association via interpretation bias (Everaert, Tierens, Uzieblo, & Koster, 2013). These results provide some indication that biased cognition in one domain may not act in isolation, but rather may influence other processing biases to ultimately influence emotional vulnerability.

In one study (Everaert, Duyck, & Koster, 2014) a novel computerised version of the scrambled sentences task (Wenzlaff & Bates, 1998) was used to investigate the direct impact of attention bias on interpretation bias to memory bias. Participants are required to unscramble a scrambled sentence to form a grammatically correct sentence of 5 or 6 words. Eye tracking is used to measure attentional bias to negative words based on fixation time on those words, while the ratio of negatively valenced to total unscrambled sentences is used to index interpretation bias. Finally, participants complete an incidental free recall test in which they recall as many of the sentences they constructed as possible, which provides an index of memory bias. From the theoretical framework of the CCBH, incorporating these biases into the same task allows more precise inferences to be made as to how attention bias directly influences interpretation bias, which directly influences memory bias. Two path models were tested, the first omitting relationships between each of the biases, and the second including paths between the biases. The results indicate that the models including functional relationships amongst cognitive biases provided a superior fit of the data. This supports the central notion of the CCBH, namely, cognitive biases in different domains do not act in isolation, but influence each other in sub-clinical depression. The authors conclude that future research should further investigate and take into account the interrelations among cognitive biases.

The research described thus far has addressed the CCBH in adults by examining modelled relationships between biases and symptoms, using tasks that specifically aim to capture the direct effects of one bias on another. Another approach to examine the CCBH is to explore differences in the relationships between biases, in clinical and non-clinical groups. If cognitive biases are found to hold different relationships to one and another in a clinical sample, but not in a normative sample, this would suggest that the interrelationships and interactions among biases might contribute to the pathology itself. In one study, the factor structure of interrelations among cognitive bias measures in attention, inhibition, imagery, and memory were found to differ between a formerly depressed, and non-clinical sample (Vrijsen et al., 2014). Coherence between attention and memory bias was found in the non-clinical sample, but not in the formerly depressed sample. This suggests that functional relationships among cognitive biases may differ between vulnerable and less-vulnerable groups.

With substantive neural, cognitive and social changes occurring during adolescence, this period of development is a key stage for the first-onset and development of emotion disorders. Research to date has been performed in almost exclusively adult samples. To our knowledge only a single study to date has examined the combined cognitive bias hypothesis in an adolescent sample (Klein, de Voogd, Wiers, & Salemink, 2017). Three cognitive bias measures were used; an emotional visual search task, a dot-probe task, and the interpretation recognition task for children. The results indicate that each cognitive bias predicted unique variance in anxiety and depression, separately, supporting the CCBH proposition that cognitive biases in different domains contribute separately to emotional vulnerability. Further exploration of adolescent mental health from a CCBH perspective is likely to be highly informative and this is the aim of the current study.

### Cognitive bias approaches to positive mental health and resilience

While information-processing approaches have been widely used to investigate the cognitive mechanisms of emotion dysfunction (for reviews, see Gotlib & Joormann, 2010; Lau & Waters, 2017; Mathews & MacLeod, 2005; Yiend, 2010) relatively little research has examined the role of selective information processing in positive mental health in adults (Carl, Soskin, Kerns, & Barlow, 2013; Parsons, Kruijt, & Fox, 2016) and even less in adolescents. Positive mental health and mental illness are considered to represent two distinct, albeit inversely correlated, continua (Keyes, 2002, 2005). Mental illness and low mental health have been found to have additive adverse effects on an individual’s functioning in life, including academic impairment and suicidal ideation (Keyes et al., 2012), as well as all-cause mortality (Keyes & Simoes, 2012). While some research has examined factors related to positive mental health within a cognitive-experimental framework, as with the CCBH, the majority of this research has been conducted in adults. Trait happiness and satisfaction with life, for example, which are both core components of positive mental health, have been associated with attentional biases towards positive images especially in the later stages of passive viewing (Raila, Scholl, & Gruber, 2015). An implication of the dual continua model is that they may be characterised by distinct patterns of selective processing styles or biases, just as the ‘symptoms’ of mental health and mental illness differ from each other. A network approach is a useful methodology to explore whether or not distinct networks of cognitive biases distinguish between individuals with high and low levels of self-reported mental health. To our knowledge, this is the first study to use network analyses to examine the role that connections in selective processing of emotional information plays in positive mental health in an adolescent sample.

### Psychological Network approaches

A network perspective on psychopathology views emotional disorders, such as anxiety and depression, as a system of interacting symptoms (Fried, 2017; Fried et al., 2017). As such, rather than individual symptoms acting alone to influence a disorder, the interrelations between them also play a key role. Borsboom and colleagues have been the driving force behind initiating network analyses in clinical psychology (Borsboom, Cramer, Schmittmann, Epskamp, & Waldorp, 2011; Schmittmann et al., 2013), and this has resulted in an increasing application of network analysis approaches to psychopathology (e.g. Bernstein, Heeren, & McNally, 2017; Borsboom & Cramer, 2013; Heeren & McNally, 2016; McNally et al., 2015). A common aim of network approaches is to identify plausible, and potentially causal, connections amongst measured variables. Typically in examining psychopathology the relationships between individual symptoms of a disorder would be investigated (e.g. McNally et al., 2015). Of particular relevance to the current work are two studies that extended the network approach to investigate laboratory measures of cognition and behaviour (Bernstein et al., 2017; Heeren & McNally, 2016). Heeren and McNally investigated the interplay between symptoms, attentional bias, and attentional control in social anxiety disorder. Their analysis indicated that the orienting of attention was strongly linked to self-reported fear of social situations, which in turn was strongly related to avoidance of those situations. The network analysis has a potentially important clinical implication in that it suggests that interventions targeting attention orientation would positively influence other processes and propagate those benefits throughout the psychological network, resulting in therapeutic effects (McNally et al., 2015). Similarly, Bernstein et al. investigated components of executive control and components of rumination. The networks suggested that self-criticism was particularly central to the networks and had strong down-stream effects on negativity and brooding in particular. Thus, reducing self-criticism may have a wide reaching beneficial effect on other components of rumination, and represents a potentially useful therapeutic target. Network analysis approaches therefore provide an informative perspective of the interplay between cognitive processes and components of psychopathology and may help to inform the development of novel clinical interventions. In the current study, we employ network analyses to investigate the CCBH in relation to positive mental health and wellbeing in adolescence.

A network theory of mental disorders has recently been proposed, with several core principles (Borsboom, 2017). The *complexity* principle proposes that mental disorders are characterised in terms of the interactions among different components in a psychopathology network. That is to say that while the individual symptoms influence mental disorders, such symptoms are also related to one another. This relates to the *direct causal connections* principle, which states that direct causal connections among symptoms form the network structure. These principles are comparable to the, albeit less formal, core elements of the CCBH; namely that we would expect different biases to interact with one another to influence emotional vulnerability and that biases may reinforce one another reciprocally to influence emotional vulnerability. The similarities are such that theoretically applying the principles of the network approach may prove fruitful and provide a more formal and systematic approach to investigation of the CCBH. We propose that cognitive biases directly and indirectly influence one another (*direct causal connections*), and that these interactions among different biases are likely to influence emotional vulnerability as well as emotional wellbeing (*complexity*). Borsboom’s fourth principle *mental disorders follow a network structure* can be described as the presence of symptoms that are more closely linked than others or as groups of symptoms that often arise together. Thus, we might expect that cognitive biases for negative information may be more highly clustered and interlinked in more vulnerable populations. Network analysis has been used to examine the structure amongst laboratory measures of attention bias and core symptoms of social anxiety disorder (Heeren & McNally, 2016), as well as functional relationships amongst components of executive control and rumination (Bernstein et al., 2017). In the current study, we sought to expand upon this work by utilising network approaches as a potentially informative theoretical and methodological framework to investigate the CCBH in adolescent mental health.

Visualising the network provides a detailed image of the multivariate dependencies that exist in the dataset. In a psychometric? network, *nodes* that represent observed psychological variables (e.g. psychometric tests or indices of cognitive bias) are connected by *edges,* which represent some statistical relationship between them, such as the correlation. The strength, or *edge weight* of these edges can differ to incorporate information about the direction of the relationship, typically represented with positive associations in green and negative associations in red. Additionally, stronger relationships are represented with thicker edges, whereas weaker relationships are denoted with thinner less saturated edges. In contrast to correlational or latent variable approaches, inference methods from graph theory can then be used to assess the relative importance of each node with several *centrality indices*. This enables inferences to be made as to which symptoms have a greater effect on the network overall, or are potential targets for intervention (Fried et al., 2017; Fried & Cramer, 2017). Network approaches therefore differ from latent variable models that explain co-occurrence amongst symptoms, or biases as in the case of the CCBH, with an underlying unobserved latent variable as a common cause of the symptoms. Under this theoretical position, emotional disorders are denoted by a network in which some symptoms interact to stimulate or inhibit one another (Costantini et al., 2015). If we consider that cognitive biases represent a vulnerability factor for emotional disorders, similar to symptoms of a particular disorder, then this theoretical position is promising in the examination of the CCBH. Under the assumptions here, and the principles of network theory described above, a psychological network analysis approach is a valuable way to analyse cognitive bias data to examine the CCBH. We note that some authors have questioned the replicability of psychopathology networks (Forbes, Wright, Markon, & Krueger, 2017). However, this conclusion has been refuted strongly with evidence demonstrating the high degree of precision with which the networks were replicated (Borsboom et al., 2017).

In the current study, we employed a psychological network analysis approach to investigate the CCBH in positive mental health in adolescents. We were interested in differences between the networks of cognitive biases in adolescents reporting high and low mental health. Therefore, we selected two groups from a large sample of adolescents based on self-reported mental health scores with the aim of examining groups that differ in their self-reported mental health. The study focussed on attention, interpretation, and memory bias, to limit the scope to the processes previously implicated in the CCBH (Hirsch et al., 2006). Our primary aim was to explore differences in cognitive bias networks between adolescents reporting high mental health and those reporting low mental health on a standardized measure. To achieve this we computed weighted and directed networks of cognitive biases in high and low mental health samples to provide an initial investigation of the CCBH in positive mental health in adolescence, and to generate novel hypotheses for future research.

# Methods

The data analysed and presented in this paper are drawn from the CogBIAS longitudinal study (Booth et al., 2017; Booth et al., 2019), which has recruited approximately five hundred 12-14 year old secondary school children in the UK. Each young person completed a series of cognitive bias measures (including attention, interpretation, and memory) and was followed up for two further waves of testing at the age of 14 and again at the age of 16. This project is the only one known to the authors to incorporate a longitudinal design, with a range of cognitive biases measured, and at three time-points in an adolescent sample. The CogBIAS project presents an ideal opportunity to examine the CCBH as it applies to adolescents, specifically with respect to the role these cognitive biases play in positive mental health. In this paper, we use data from wave 1 of the CogBIAS project.

### Participants

We first excluded all participants without complete data in all of the measures described below (and removing DPT accuracy below 70% before this), resulting in a sample of 448 students (from the original sample of 504). For our first analysis we selected two groups of participants for the network analyses based on scores on the Mental Health Continuum. We performed a tertile split to yield low and high MH groups that differ substantially on MH. The low-MH group consisted of 145 participants, scoring under 37 on the MHC and the high-MH group consisted of 150 participants scoring above 47 on the MHC.

Table 1 presents the demographics and summary data from the study measures, for the entire sample. Ethical approval for this study was given by the National Research Ethics Service (NREC; REC reference: 14/SC/0128; IRAS project ID: 141833).

[Table 1. About here]

**Measures**

As we were interested in differences in network structure of cognitive biases in relation to high and low positive mental health we analyse only a subset of the measures included in the CogBIAS project. The combined cognitive bias hypothesis typically describes the relationship between attention, interpretation, and memory biases. We therefore analysed data only from tasks targeting these processes. We present a brief description of these measures below, however, a complete description of the sample, methods, and design used in the project can be in the protocol paper (Booth et al, 2017).

**Mental health.** The MHC-SF contains 14 items that index emotional, psychological, and social wellbeing, in order to create a composite measure of positive mental health. Participants are asked to rate how often they have experienced each of the items in the past month, on a 6-point Likert scale from "never" to "every day". The MHC-SF has shown high internal consistency and discriminant validity (Keyes, 2009; Lamers, Westerhof, Bohlmeijer, ten Klooster, & Keyes, 2011; in the current sample MacDonald’s Omega = .95, and Cronbach’s alpha = .94).

**Attention Bias.** An emotional face (angry, happy, and pain) dot-probe task was used to index attention bias (MacLeod et al., 1986) with stimuli from the STOIC faces database (Roy et al., 2007). As with other papers using data from the CogBIAS project, we have opted to omit the attention bias data from our analyses. The internal consistency of each of the attention bias indices was below any acceptable threshold, in this sample; angry = -.07, 95%CI [-.20, .06]; happy = .12, 95%CI [.00, .24]; pain = -.04, 95%CI [-.17, .09]. These outcome measures are unsuitable for any analyses based on correlational measures and are therefore omitted from any further analyses. For full details about the task see Booth et al. (2017; 2019).

**Interpretation bias.** The Adolescent Interpretation and Belief Questionnaire (AIBQ; Miers, Blöte, Bögels, & Westenberg, 2008)contains ten hypothetical situations, five of which are socially oriented and five are non-socially oriented, that are intended to reflect events that are likely to be experienced by adolescents. Participants read the scenario and are presented with a question that addresses a point of ambiguity in the scenario. A positive, a neutral, and a negative interpretation of the scenario are presented and participants rate how likely that interpretation would pop into their mind on a 5-point Likert scale. Participants then choose which interpretation of the situation they believe to be the most correct. Scenarios are presented in a pseudo-random order. Bias scores were calculated by calculating the mean likelihood ratings for positive and negative interpretations of social and non-social situations separately, resulting in four bias indices – positive social; positive non-social; negative social; negative non-social (table 1 presents means and standard deviations od scores, as we as reliability indices).

**Memory bias.** In the Self-referential encoding task (SRET), participants were presented with 22 positive and 22 negative words in a random order (the word lists were drawn from Hammen & Zupan, 1984). Each word was presented for 200ms before a prompt "Describes me?" was presented on screen, after which participants responded ‘yes’ or ‘no’. After all words had been presented a short distraction task was administered consisting of three simple mathematics questions. Finally, in the incidental recall phase, participants were given three minutes to recall and type in as many words as they could remember. Two bias indices were calculated. Positive and negative memory bias indices were calculated as the number of positive and negative words, respectively, that were endorsed (participants responded that the word described them) and subsequently recalled (Asarnow, Thompson, Joormann, & Gotlib, 2014). We are not aware of a suitable procedure for estimating the reliability of the memory bias indices due to the process of calculating the scores and so we interpret associations with this outcome with some caution.

### Procedure

Participants were tested in groups of between 13 and 50 students in computer labs either in their own school, or at the University of Oxford. Testing consisted of two, one-hour sessions which were either back-to-back or on different days, depending on school availability. In each session, participants completed three tasks, in the same order, followed by a battery of questionnaires (see Booth et al, 2017, for further information on measures not analysed in this paper). Participants were asked to complete both sessions in exam conditions, i.e. not talking or looking at neighbours’ computer screens. At least one researcher was present throughout the testing sessions to answer any questions and ensure adequate testing conditions were maintained.

### Data analysis

We first present a comparison of the network structure of interpretation and memory biases for a high mental health and a low mental health group, following a tertile split of the data by mental health. This analysis formed the basis of the earlier version of this paper (in the next section we explain why a different approach is preferable). We computed a ‘graphical LASSO’ (glasso; Epskamp & Fried, 2016, also Friedman, Hastie, & Tibshirani, 2008) estimation procedure with EBIC model selection (Foygel & Drton, 2010). The glasso algorithm is implemented in the *glasso* package (Friedman, Hastie, & Tibshirani, 2014), and is called by the *bootnet* package (Epskamp et al., 2017), which we used for this paper. The glasso algorithm estimates a partial correlation network by directly penalising elements of the variance-covariance matrix. We set the tuning parameter gamma to 0.5 to generate a sparser network, due to the removal of potentially spurious associations. We then used the NCT() function from the *NetworkComparisonTest* package (van Borkulo, 2016) to compare our high mental health and low mental health netwotks. The function tests for differences in the overall connectivity (as the sum of all edge weights in the network, or global strength) between networks.

The downsides to this approach are; fewer partcipants’ data used, lack of nuance in interpretations.

We used the mgm package () to compute a moderated network. This enabled us to use mental health as a linear variable, thus removing a potentially misleading dichotomy as well as allowing us to use the full sample of participants (that completed all measures).

We also include information on predictability, i.e. the variance explained by all other variables in the network.

We bootstrapped / resampled the network estimation using resample() in mgm using 5000 resamples.

# Results

Table 1 - demographics and reliabilities of measures (possibly correlation matrix?).

**Comparing networks of cognitive biases in a low mental health group to a high mental health group**

First we estimated a graphical LASSO network (tuning parameter gamma was set to .5 to generate a sparser network) for the high and low mental health groups separately. Figure 1 presents a visualisation of both networks. We compared the global strength of each network (the sum of edge strengths in the network) using NCT() from the NetworkComparisonTest package (van Borkulo, 2016). We set NCT to run 1000 iterations resampling

The low-MH network (global strength = 1.70) was more connected than the high-MH network (global strength = .37), and this difference was statistically significant, *p* = .014.

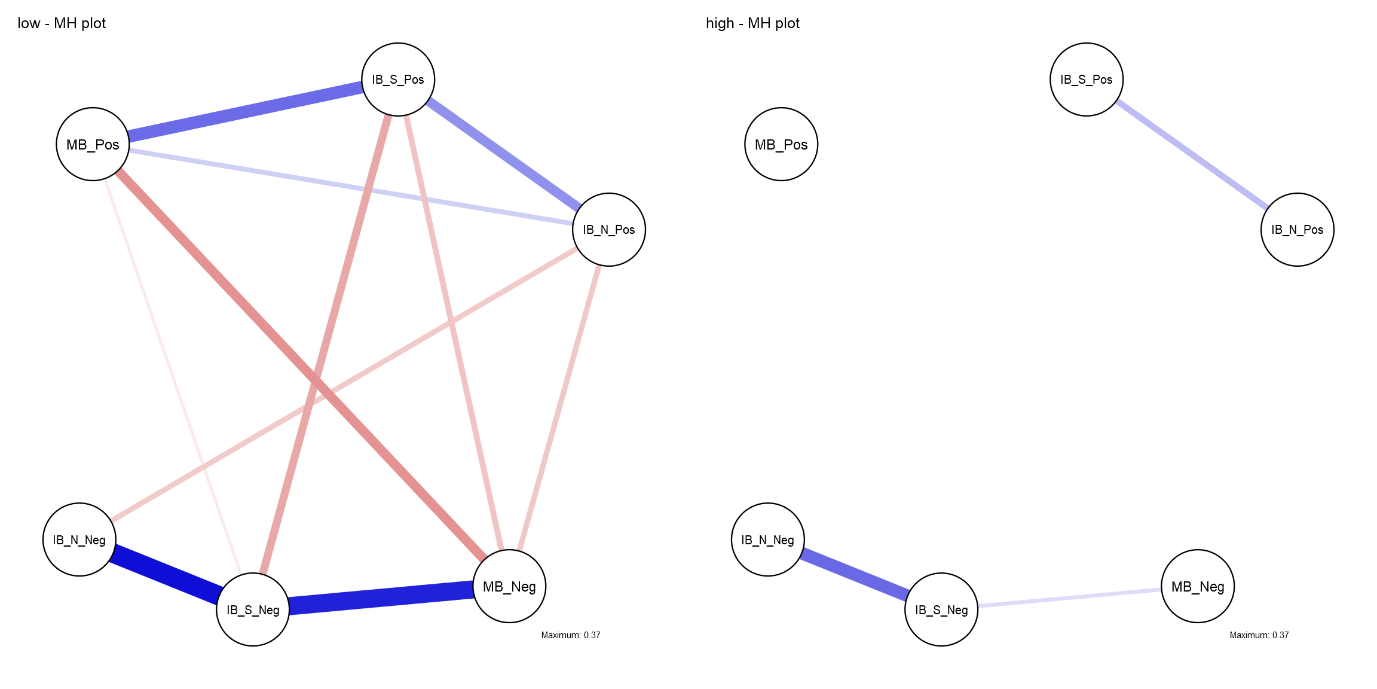
These results give us some impression that high and low mental health groups may differ in the interrelationships amongst cognitive biases in interpretation and memory.

There are limitations to this approach, however. First, we only gain an impression of the overall difference in connectivity between networks. We could investigate differences in individual edges, except we quickly run into a large multiple comparisons issue with NCT. Thus, this analysis lacks specificity. Second, we lose one third of the full sample with this split and treat a continuous variable as categorical – both presenting their own issues.

An improvement would be to use a moderated network approach (). The results thusfar support the idea that mental health may have some moderating effect on the network structure.

This gives several advantages; first, we can use our full sample; second, we can treat mental health as a continuous variable; third, we can estimate the moderating effect of mental health on each edge in the network.

Glasso with high and low MH group



*Figure 1.* Graphical LASSO Networks. The left and right panels present the graphical LASSO network from the low mental health group and the high mental health group, respectively. Each node represents a cognitive bias measure and each edge represents the (partial) correlation between the nodes it connects, after controlling for all other variables in the network. Larger correlations are represented by thicker edges. Blue edges indicate positive relationships, whereas red edges indicate negative relationships.

Note: IB\_S\_Pos = Positive interpretation bias in social scenarios in the Adolescent Interpretation and Belief Questionnaire (AIBQ); IB\_S\_Neg = Negative interpretation bias in social scenarios in the AIBQ; IB\_N\_Pos = Negative interpretation bias in non-social scenarios in the AIBQ; IB\_N\_Neg = Negative interpretation bias in non-social scenarios in the AIBQ; MB\_Pos = endorsed and recalled positive items from the self-referential encoding task; MB\_Neg = endorsed and recalled negative items from the self-referential encoding task.

We took a moderated network approach

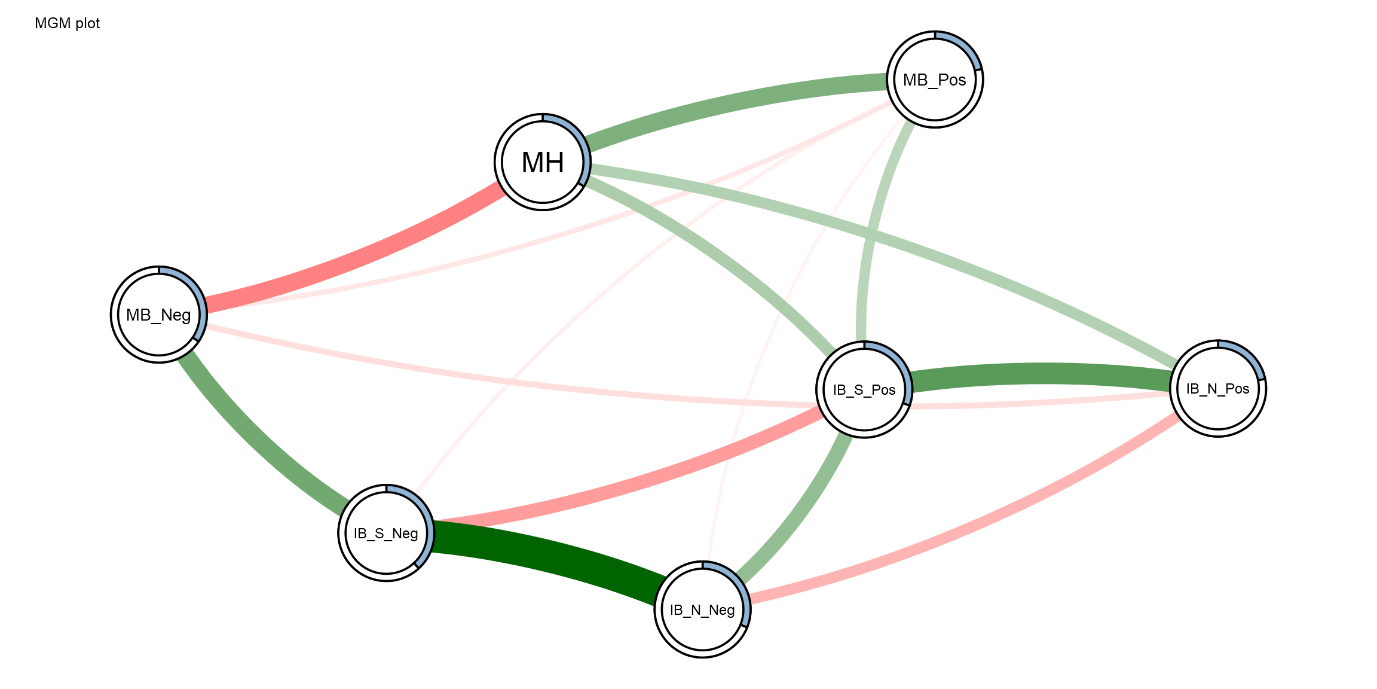


Figure 2. estimated moderated network of positive mental health, and interpretation and memory cognitive biases. Blue edges represent positive associations, red edges represent negative associations; the width of the edge indicates the strength of this relationship. The shaded ares of the pie surrounding each node represents the predictability of that variable, i.e. the variance explained by all other variables in the network. Note that this figure does not visualise the degree of moderation in the networks.

Note: MH = Positive mental health; IB\_S\_Pos = Positive interpretation bias in social scenarios in the Adolescent Interpretation and Belief Questionnaire (AIBQ); IB\_S\_Neg = Negative interpretation bias in social scenarios in the AIBQ; IB\_N\_Pos = Negative interpretation bias in non-social scenarios in the AIBQ; IB\_N\_Neg = Negative interpretation bias in non-social scenarios in the AIBQ; MB\_Pos = endorsed and recalled positive items from the self-referential encoding task; MB\_Neg = endorsed and recalled negative items from the self-referential encoding task.

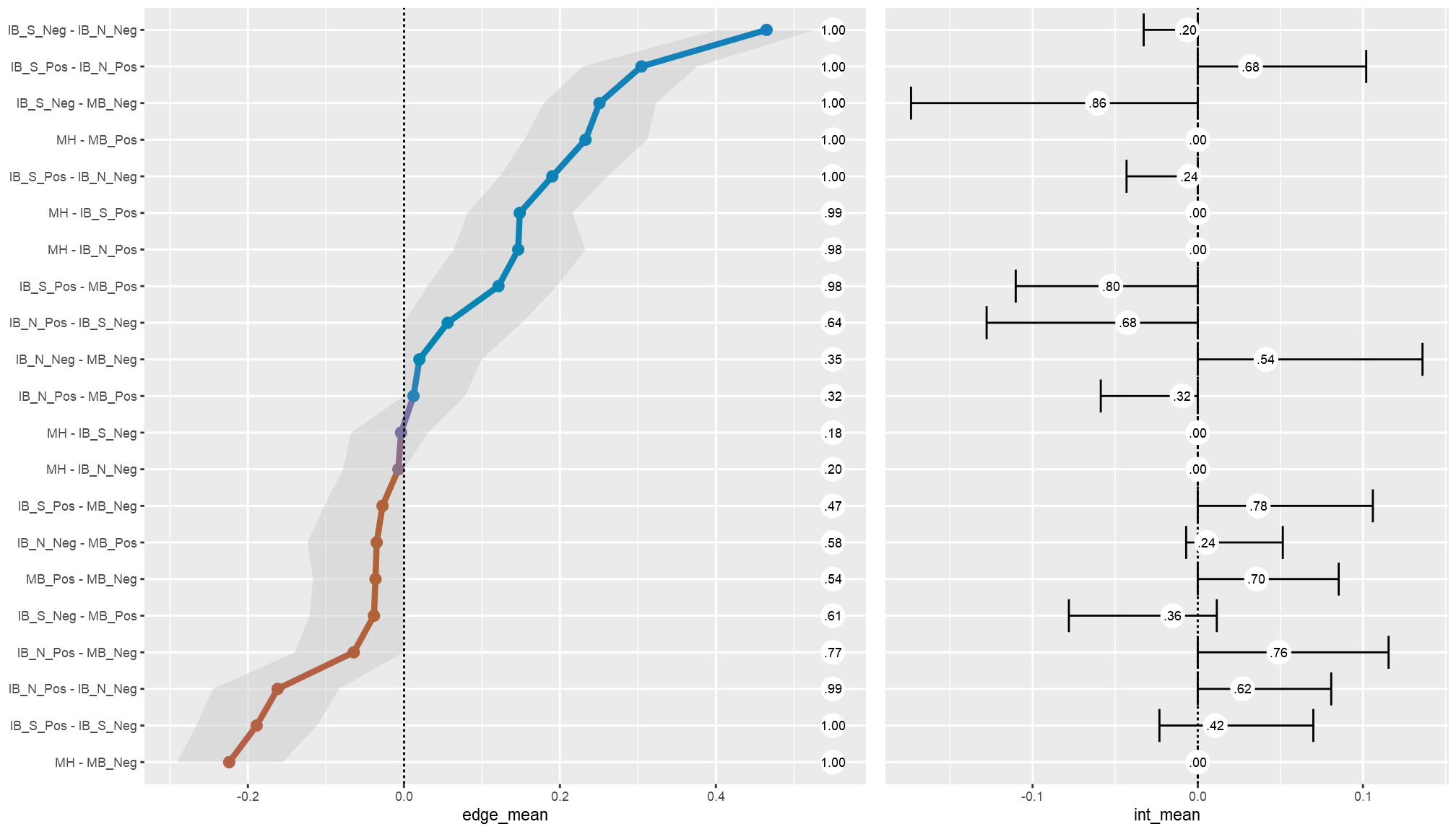


Figure 3. edge strength and degree of moderation by mental health. The left panel presents the estimated edge strengths within the network (these correspond to the network visualization in figure 2) from 5000 resamples of the mgm network estimation procedure. The shaded area represents the 95% CI around the estimate. Numbers running down the centre of the figure represent the proportion of non-zero estimates for each edge. The right panel presents the estimated moderating effect of mental health on each edge; the ticks represent the 95% CI around the estimate. The circled numbers represent the number of non-zero moderation effects arising across the resamples.

Thoughts:

* It does seem that there is some moderation going on. But it is fairly nuanced. In general, higher mental health seems to push the network towards reduced connectivity, but with exceptions. So, overall our

