Original Article

CAMBRIDGE
UNIVERSITY PRESS

Investigating change in network structure of eating disorder symptoms after delivery of a smartphone app-based intervention

Jake Linardon[1,2], Christopher J. Greenwood[1,2,3], Stephanie Aarsman[1,2] and Matthew Fuller-Tyszkiewicz[1,2]

[1]School of Psychology, Deakin University, Geelong, Victoria, Australia; [2]Center for Social and Early Emotional Development, Deakin University, Burwood, Victoria, Australia and [3]Department of Paediatrics, University of Melbourne, Melbourne, Australia

Abstract

Background. Eating disorder (ED) research has embraced a network perspective of psychopathology, which proposes that psychiatric disorders can be conceptualized as a complex system of interacting symptoms. However, existing intervention studies using the network perspective have failed to find that symptom reductions coincide with reductions in strength of associations among these symptoms. We propose that this may reflect failure of alignment between network theory and study design and analysis. We offer hypotheses for specific symptom associations expected to be disrupted by an app-based intervention, and test sensitivity of a range of statistical metrics for identifying this intervention-induced disruption.

Methods. Data were analyzed from individuals with recurrent binge eating who participated in a randomized controlled trial of a cognitive-behavioral smartphone application. Participants were categorized into one of three groups: waitlist ($n = 155$), intervention responder ($n = 49$), and intervention non-responder ($n = 77$). Several statistical tests (bivariate associations, network-derived strength statistics, network invariance tests) were compared in ability to identify change in network structure.

Results. Hypothesized disruption to specific symptom associations was observed through change in bivariate correlations from baseline to post-intervention among the responder group but were not evident from symptom and whole-of-network based network analysis statistics. Effects were masked when the intervention group was assessed together, ignoring heterogeneity in treatment responsiveness.

Conclusion. Findings are consistent with our contention that study design and analytic approach influence the ability to test network theory predictions with fidelity. We conclude by offering key recommendations for future network theory-driven interventional studies.

Introduction

The network perspective of psychopathology proposes that psychiatric disorders can be conceptualized as a complex system of interacting symptoms (McNally, 2016). Functioning of the network may be explored globally, in terms of overall network strength (alternatively referred to as global network connectivity), as well as at a more localized level, exploring the contributions of individual symptoms. At the individual-symptom level, statistical networks enable identification of 'central' symptoms, characterized by strength and/or number of associations with other symptoms within the network (Fried et al., 2017, although see Bringman et al., 2019 for discussion of potential issues with centrality metrics). These central symptoms are thought to be viable targets for intervention through which broader symptom improvement may proliferate. Beyond this symptom-level focus, it is argued that vulnerability to psychopathology, resistance to *v.* success of treatment, and likelihood of relapse may potentially be discerned through the overall strength of the network (Borsboom, 2017).

Within the context of eating disorder (ED) research, network analyses of cross-sectional data offer support for the central role of overvaluation with shape/weight and desire to lose weight across distinct age groups and diagnoses, while also identifying several non-specific ED symptoms that might serve as central symptoms in the network (Monteleone & Cascino, 2021). These findings with respect to overvaluation and desire for weight loss accord with cognitive behavioral therapy (CBT) approaches to eating disorders, where these cognitive factors are a key focus of treatment (Fairburn, Cooper, & Shafran, 2003).

Despite these promising cross-sectional results, efforts to longitudinally explore changes in network structure in response to treatment have yielded mixed and somewhat surprising results. Some changes from baseline to post-intervention were observed at the symptom-level, though the general pattern was for increased node strength (i.e. increased connections to other

symptoms in the network) rather than a reduction in line with diminished impact of the symptoms in response to treatment. These variables with increased strength include dissatisfaction with weight and shape, BMI, shape concerns, eating concerns, binge eating, and depressive symptoms (Forrest & Grilo, 2022; Forrest, Franko, Thompson-Brenner, & Grilo, 2023; Hilbert et al., 2020).

At the global level, network theory-predicted reductions in network connectivity in response to treatment were found for one study (Calugi, Dametti, Chimini, Dalle Grave, & Dalle Grave, 2021), yet non-significant differences (Forrest & Grilo, 2022 baseline *v.* post-intervention comparison; Forrest et al., 2023; Smith et al., 2019) or increases in overall connectivity (Forrest & Grilo, 2022 baseline *v.* follow-up comparison; Hilbert et al., 2020) were found in others. Smith et al. (2019) also found in a transdiagnostic sample that treatment non-responders had greater network connectivity at pre-intervention relative to responders, but this effect was not found by Hilbert et al. (2020) among a binge eating disorder sample. It is worth noting that significant symptom improvements based on traditional tests of efficacy (e.g. ANOVAs) were observed across each of these longitudinal ED network studies, and hence null findings are not attributable to an ineffectual treatment.

Several factors may account for these unexpected findings with respect to global network connectivity in response to treatment. First, network theory predictions about reduced connectivity may be poorly matched to treatment contexts in these existing studies. Forrest and Grilo (2022) hypothesized that treatment would lead to reduced centrality of overvaluation of shape/weight, yet their behavioral weight loss treatments did not directly target this symptom. The CBT-based intervention studies could be reasonably argued to target symptoms included in their network analyses, but these studies did not offer clear hypotheses about which variable inter-relations would be disrupted by specific treatment content. If the treatment is expected to target only a few of the overall number of inter-relations within the network, one might expect small changes to the overall network structure and to symptom (node) strength that necessitate larger samples than these prior studies to detect structural changes. These targeted disruptions to specific symptom relationships may also be better evaluated at the bivariate (or partial) correlation level than through metrics of global network strength or node strength.

A related potential concern is ambiguity in use of the term 'disruption' in the context of these networks. Presence, severity, or frequency of a target symptom may be *disrupted* by reducing the presence, severity, or frequency of a known antecedent symptom. This improvement to the target symptom arises from a recognized causal relationship between the two symptoms. However, the expectation from this interpretation of disruption is that the target symptom reduces in severity because the antecedent symptom's severity has also been reduced. In such a case, we would see that the correlation remains, yet symptom levels reduce. A second interpretation of *disruption* is more in line with the network theory prediction that treatment would reduce the strength of associations among variables. In such a case, the antecedent symptom's level of severity or frequency becomes irrelevant to the likelihood of the target symptom occurring. In this case, successful treatment would entail reduced target symptom severity/frequency (with or without coinciding reduction in antecedent symptom severity/frequency) and reduced strength of the target causal relationship. Thus, in applications of network analysis, we need clearer articulation regarding the type of disruption that is predicted, the specific associations in the model expected to be disrupted, and plausible explanation for why such a disruption is anticipated.

Finally, statistical artifacts may also threaten internal validity of global network strength evaluations in complex, yet potentially predictable ways. Psychological treatment effects are often heterogeneous (Hilbert et al., 2019); analyzing all treatment group participants together may thus obfuscate true impacts on network structure at varying levels of treatment responsiveness. Increases in variability over time – such as may arise from heterogeneous treatment effects – could also lead to increased correlation magnitude (Bland & Altman, 2011). Indeed, Monteleone et al. (2021) found an average correlation of 0.40 between variance predicted in a symptom (node) in a network and the size of its variability (Monteleone and colleagues argue that these associations become non-significant after correction for Type I error inflation. However, we argue that given the small sample size of this study and the choice of a conservative adjustment approach (Bonferroni), this skews results in favor of failing to find an effect). Changes to the distribution of symptom scores may also arise due to treatment effects, in turn resulting in correlation magnitudes that are more or less affected by skew due to attenuation bias or (mis)match of distributions. Consideration of these potential statistical artifacts should be routinely incorporated into network analysis papers to contextualize results of hypothesis testing.

The present study

This study demonstrates our proposal for more appropriate hypothesis formulation and more comprehensive testing to contextualize results of hypothesis tests. Our substantive aim in this study is to evaluate whether app-based intervention responsiveness is associated with changes to network structure for a collection of ED symptoms and broader mental health correlates thought to co-occur with and influence these symptoms.

The app-based intervention evaluated is based on CBT principles that are designed to target binge eating via normalizing eating behaviors, decreasing perceived body image importance, and better tolerating adverse moods. The main intended mechanism for change is reduction in antecedents which, in turn, are expected to reduce key symptoms. There are, however, two specific exceptions, where the intervention targets *disruption* (in the network analysis sense of the word) to the relationship between symptoms. The first of these attempts to educate individuals about the role of dieting in binge eating, and seeks to reduce both dieting attempts and subsequent binge eating after these attempts. The second exception educates participants about the role of negative mood states in precipitating binge eating episodes as a means to self-soothe, provide comfort, or distract. Participants are taught to disrupt the functional relationship between negative thoughts and binge eating as a reaction.

Consistent with traditional formulations of efficacy, it is predicted that individuals assigned to the intervention will exhibit greater reductions in these symptoms post-intervention relative to a wait-list group (Hypothesis 1). Consistent with network theory, it is further hypothesized that specific changes in the network structure of these symptoms are expected for intervention participants for relationships between: (i) dietary restraint and binge eating, (ii) depressive symptoms and binge eating, and (iii) body dissatisfaction and binge eating (Hypothesis 2). As the relationships expected to reduce form a small subset of the overall number of possible relationships (3 of 28), hypothesized reductions

will be tested both in terms of reduction in global strength metrics and through evaluation of change in bivariate correlations, the latter expected to be more sensitive to these proposed changes.

Method

Study design and population

Participants were those with self-reported recurrent binge eating (defined as one episode of objective binge eating per every two weeks, on average, over the past three months) enrolled in a three-arm RCT comparing a waitlist to either a broad (app-based program designed to target multiple mechanisms symptoms) or focused (web-based program designed to isolate one key mechanism) self-guided digital intervention (Linardon, Shatte, McClure, & Fuller-Tyszkiewicz, 2023). Participants were recruited from advertisements distributed throughout the first author's psychoeducational website for eating disorders, which showcases passive information about eating disorders, including their causes, consequences, prevalence, and help options (see Linardon, Rosato, & Messer, 2020a; Linardon, Shatte, Messer, Firth, & Fuller-Tyszkiewicz, 2020b; Linardon, Shatte, Rosato, & Fuller-Tyszkiewicz, 2020c for further detail). Those allocated to the focused intervention were excluded from subsequent analyses due to the program's lack of focus on targeting multiple symptoms and their relationships hypothesized to change in network models (i.e. it did not target body image and mood-related variables like the broad program did). Baseline and four-week follow-up (immediately post-intervention) are analyzed in this study. Ethical approval was obtained through Deakin University.

Intervention

The intervention (*Break Binge Eating*) was delivered through a smartphone application. It was self-guided, based on CBT principles, and contained four sequential modules designed to target three key binge eating maintaining mechanisms: dietary restraint, body image, and negative mood. Modules took between 30 and 90 min to complete and interactive exercises (e.g. quizzes, symptom tracking, and digital diary) were provided to help users practice skills taught throughout the program. The app was designed to be completed in four weeks, with users encouraged to engage in one module per week. However, as the app was purely self-guided, participants were free to go at a self-suited pace. For more information about the intervention, see Linardon et al. (2023)

Measures

Each network included items or subscales from the Eating Disorder Examination Questionnaire (EDE-Q; Fairburn and Beglin, 1994) and Patient Health Questionnaire–4 (PHQ-4; Kroenke, Spitzer, Williams, and Löwe, 2009). The EDE-Q is a 28-item measure that assesses attitudinal and behavioral symptoms of eating disorders. Attitudinal symptoms are rated along a 7-point scale, with higher scores reflecting greater severity. Behavioral symptoms are rated as the number of episodes experienced in the past month. For each network, we used the following EDE-Q symptoms: dietary restraint (five item subscale), eating concerns (five item subscale), overvaluation (five item composite score), dissatisfaction (two item composite score), fear of weight gain (single item number ten), preoccupation with weight and shape (single item number eight), and objective binge eating.

We assessed individual constructs of body image rather than the broad shape and weight concerns subscales in light of prior recommendations and empirical evidence highlighting the distinctiveness of these body image components (Lydecker, White, & Grilo, 2017; Mitchison et al., 2017). We also used the depression subscale from the PHQ-4, which contains two items rated along a 4-point scale.

Analytical approach

Data preparation and preliminary analyses

All analyses were conducted in R version 4.2.2 (R Core Team, 2022). There was less than 5% missing data across all variables; this was handled with a single imputation of modeled variables and demographics using chained equations with the *mice* package (van Buuren & Groothuis-Oudshoorn, 2011). Non-normality was evident for the binge-eating variable. This variable was transformed based on Box–Cox transformation recommendations of log transformation. The goldbricker function (Jones, 2023) was used to assess potential redundancy among symptoms intended for our network analysis. We failed to find any redundancy searching for correlations > 0.5, 0.25 as significant proportion for flagging problematic variables, and p set at 0.01 (Mullen & Jones, 2021).

To address heterogeneity of treatment response, we separate the treatment sample into two groups (treatment responders *v.* non-responders) based on the criterion of at least 50% reduction in objective binge eating episodes (Williams, Watts, & Wade, 2012). It is expected that reductions in magnitude of relationships between binge eating and both dietary restraint and the negative mood variables would be present for the responder group only (not waitlist or non-responder groups), reflective of their overall responsiveness to intervention.

Given the two groups (responders and non-responders) were not randomly assigned, one-way ANOVAs were conducted to evaluate potential differences between the intervention subgroups and wait-list control at baseline for demographics and variables included in our network analyses. Levene's test was used to evaluate potential differences in variance across groups for the same variables due to concern that lower variance estimates for one group may produce reduced correlation magnitude through range restriction (Bland & Altman, 2011). No correction to an alpha level of 0.05 was applied as such corrections would favor a null finding which is advantageous for our conclusions about data quality.

Hypothesis testing

Between-group differences in symptom severity at post-intervention (Hypothesis 1) were tested using one-way ANOVAs. Tukey's Honestly Significant Differences test was used to conduct post-hoc comparisons between the three groups. Bivariate correlations (with Fisher's z test for comparisons) and network analyses (symptom strength centrality and global network strength) were used to evaluate changes in symptom interrelation over time between groups (Hypothesis 2).

Bivariate correlations were run separately by intervention group and timepoint, with Fisher's z test approach used to evaluate whether differences in correlation strength within group but across time (e.g. difference in correlation between body dissatisfaction and binge eating at baseline *v.* post-intervention for the responder group) were significant ($p < 0.05$). In addition to significance testing, we adopted Ferguson's (2009) guidelines for

evaluating practical significance of these changes in correlation magnitude; differences greater than 0.2 were considered the minimum effect size reflective of practical significance.

Network analyses were conducted using the graphical least absolute shrinkage and selection operator (GLASSO) algorithm via *qgraph* for network construction (Epskamp, Cramer, Waldorp, Schmittmann, & Borsboom, 2012), and *NetworkComparisonTest* to evaluate possible differences in global network structure and symptom strength across groups at each timepoint and within group over time (van Borkulo et al., 2022).

Additional analyses

Though not directly informing hypotheses, we visually present networks to enable readers to visually inspect symptom interrelation across groups and time. Stability of the network strength metrics that inform Hypothesis 2 were evaluated using correlation stability coefficients (CS-coefficients) via the bootnet package with 1000 case-dropping bootstraps (Epskamp, Borsboom, & Fried, 2018). CS-coefficient values of > 0.25 indicate minimally acceptable stability of network strength metrics, whereas >0.50 indicates desirable stability levels. Finally, we evaluated whether there was a systematic relationship between a symptom's level of variability around its mean and the variance explained in this symptom by all other symptoms in the network. This was achieved through bivariate correlations, and run separately by group (responders, non-responders, wait-list).

Results

Preliminary analyses

Group differences

At post-test, the overall completion rate was 63% (37% did not complete study assessments) for the group allocated to the app intervention. Treatment responders, non-responders, and wait-list participants did not significantly differ on demographic factors (Table 1), nor for any measure at baseline used for network analyses, except binge eating frequency for which responders reported more frequent episodes and higher variability around their mean score (see online Supplementary Tables S1 and S2). The responders did not significantly differ from the other two groups in correlation magnitude for baseline associations between key variables for Hypothesis 2 (dietary restraint, body dissatisfaction, and depressive symptoms all in relation to binge eating; online Supplementary Tables S5–S8).

More differences in variability were evident at post-intervention, with significant differences for overvaluation, dissatisfaction, fear of weight gain, depressive symptoms, and binge eating. In most instances, the wait-list control group had less variability than the intervention groups (see online Supplementary Table S3). Possible implications of these differences for network comparisons are covered in the Discussion.

Network structure

Online Supplementary Fig. S1 provides the networks at baseline and post-intervention for the eating disorder-relevant variables for responders, non-responders, and wait-list control groups, respectively. Broadly, we see that the number of edges (i.e. associations among two variables, conditioned on other variables in the model) remained relatively stable for all groups, and there was no evidence of change in edges involving binge eating for the responder group. Online Supplementary Fig. S2 shows a slight

increase in number of edges when all intervention participants are modeled together, though none of the changes are consistent with expectation of reduced edges involving binge eating.

Online Supplementary Fig. S3 summarizes the strength of all variables within the networks expressed as z scores. At both time points, many of the strength indices are highly similar for the three groups, with strength values within 0.5 units of each other. Most relevant to hypotheses about treatment efficacy, strength values for binge-eating were at zero for the responder group, but the other groups also had small–albeit non-zero–strength values for binge-eating. This pattern held when comparing full intervention group against the waitlist control group (online Supplementary Fig. S4). Further, analyses revealed at least minimally acceptable levels of stability in strength metrics across groups and time (correlation-stability coefficients values: 0.267 and 0.289, 0.359 and 0.519, and 0.441 and 0.594 at baseline and post-intervention for the responder, non-responder, and wait-list groups, respectively).

The relationship between variance explained in a symptom and the variability of that symptom was non-significant but ranged from moderate to large at post-intervention for all groups ($r_{wait-list} = -0.62$, $p = 0.115$; $r_{responder} = -0.50$, $p = 0.204$; $r_{non-responder} = -0.33$, $p = 0.428$). For the responder group–for whom reductions in correlations were hypothesized–the association between variance explained and variability was negative, and hence unlikely to have created a statistical artifact in favor of our hypothesis.

Hypothesis testing

Group differences post-intervention (Hypothesis 1)

There were significant differences between groups at post-intervention for all variables (Table 2). In all cases, except for dietary restraint, the responder group significantly differed from the wait-list control group on these variables at post-intervention, and also differed from the non-responder group on eating concerns, overvaluation, fear of weight gain, and binge eating. The responder group tended to have the lowest scores on these variables, consistent with expectation of efficacy for this group.

Differences in symptom relationships (Hypothesis 2)

Bivariate correlations. Table 3 provides magnitude of change in correlations across time, within each group. Significant changes over time (with anticipated reduction in initially positive correlations) were found in the responder group for one of the three relationships hypothesized to be disrupted by the intervention (dietary restraint and binge eating), and a second (between binge eating and depressive symptoms) was borderline significant ($p < 0.10$) but above the minimal practice significance threshold of 0.2 (Ferguson, 2009). Magnitude of change in relationship between binge eating and body dissatisfaction was not significant. There was a significant decrease in the magnitude of association between body dissatisfaction and preoccupation, which was not hypothesized.

For non-responsive individuals, five bivariate symptom relationships exhibited a (borderline) significant change in magnitude. For each of these relationships involving binge eating, the observed pattern was strengthening (rather than reduction) of these problematic associations. Finally, in the absence of intervention, there were two significant differences and one borderline significant difference observed over time for the wait-list control

Table 1. Demographic breakdown by group

	Non-responder	Responder	Wait-list	
	(N = 77)	(N = 49)	(N = 155)	p Value
Gender (female)	75 (97.4%)	46 (93.9%)	147 (94.8%)	0.643
Age (mean, s.d.)	35.7 (9.86)	35.3 (8.60)	34.6 (10.40)	0.736
BMI (mean, s.d.)	30.2 (7.92)	30.5 (8.77)	30.5 (8.64)	0.967
Race				
Black	1 (1.3%)	2 (4.1%)	3 (1.9%)	0.148
Asian	4 (5.2%)	4 (8.2%)	5 (3.2%)	
White	68 (88.3%)	37 (75.5%)	141 (91.0%)	
Multiracial	0 (0%)	1 (2.0%)	3 (1.9%)	
Other	4 (5.2%)	5 (10.2%)	3 (1.9%)	
Education				
> Year 12	72 (93.5%)	45 (91.8%)	139 (89.7%)	0.600
Less than Year 12	0 (0%)	1 (2.0%)	1 (0.6%)	
Year 12	5 (6.5%)	3 (6.1%)	15 (9.7%)	
Past AN	3 (3.9%)	4 (8.2%)	17 (11.0%)	0.185
Past BN	11 (14.3%)	9 (18.4%)	18 (11.6%)	0.462
Past BED	20 (26.0%)	14 (28.6%)	47 (30.3%)	0.778
Past OSFED	3 (3.9%)	1 (2.0%)	7 (4.5%)	0.790
Current ED				
AN	0 (0%)	1 (2.0%)	2 (1.3%)	0.947
BED	17 (22.1%)	11 (22.4%)	43 (27.7%)	
BN	5 (6.5%)	4 (8.2%)	8 (5.2%)	
No diagnosis	52 (67.5%)	31 (63.3%)	97 (62.6%)	
OSFED	3 (3.9%)	2 (4.1%)	5 (3.2%)	
Current treatment	17 (22.1%)	12 (24.5%)	35 (22.6%)	0.957

Notes. s.d., standard deviation; BMI, body mass index; AN, anorexia nervosa; BN, bulimia nervosa; BED, binge eating disorder; OSFED, other specified feeding or eating disorder; ED, eating disorder.

group. These differences were not in the direction of reduced symptom association.

Online Supplementary Table S4 shows none of the three target correlations reduce from either a significance or effect-size perspective from baseline to post-intervention when the two intervention subgroups are combined and heterogeneity is ignored. Online Supplementary Tables S5–S8 provide correlations for each group across the two timepoints.

Network analyses. Invariance tests failed to find significant reduction in global network strength for the responder intervention group from baseline to post-intervention (*p*s > 0.623). Nor were there significant differences between the responder group and non-responders (*p*s > 0.053) or wait-list (*p*s > 0.105) at post-intervention. Full results are shown in online Supplementary Table S9.

Strength statistics for individual symptoms in the network (reflecting the aggregation of a symptoms associations with all other symptoms in the network) did not significantly change from baseline to post-intervention for the responder group for the four key variables involved in Hypothesis 2 (binge eating, body dissatisfaction, depressive symptoms, and dietary restraint; all *p*s > 0.204).

Discussion

Despite considerable research interest in network perspectives for psychopathology, most studies have failed to find that symptom reductions post-intervention coincide with reductions in overall strength of associations among these symptoms (e.g. Beard et al., 2016; Smith et al., 2019). We argue that this may, at least in part, reflect failure of alignment between network theory and study design and analysis. Thus, we articulated specific symptom associations expected to be reduced in magnitude by successful intervention, and evaluated the capacity of different statistical tests (bivariate associations, network-derived strength statistics, and network invariance tests) to detect such changes in a group of individuals with ED who were responsive to our CBT-based treatment.

Our findings showed significant improvements for the responder group at the symptom-level for all variables, consistent with traditional conceptualizations of efficacy. Furthermore, inspection of bivariate correlations showed practically significant (two associations: dietary restraint and negative mood with binge eating) and statistically significant (one association; restraint with binge eating) reductions among the three

Table 2. Breakdown by group of post-intervention scores on variables included in network analysis

	Non-responder	Responder	Wait-list	
	(*N* = 77)	(*N* = 59)	(*N* = 155)	*p* Value
Restraint	2.21 (1.43)[WL]	1.89 (1.56)[WL]	3.07 (1.54)[NR,R]	<0.001
Eating concerns	2.93 (1.48)[WL,R]	2.03 (1.22)[WL,NR]	3.43 (1.25)[NR,R]	<0.001
Overvaluation	4.40 (1.63)[R]	3.63 (1.84)[WL,NR]	4.81 (1.38)[R]	<0.001
Dissatisfaction	4.77 (1.42)	4.34 (1.56)[WL]	5.07 (1.23)[R]	0.009
Preoccupation	2.82 (2.02)	2.00 (1.98)[WL]	3.36 (2.05)[R]	<0.001
Fear weight gain	4.42 (1.84)[WL,R]	3.45 (2.21)[WL,NR]	5.18 (1.52)[NR,R]	<0.001
Depression	2.86 (2.06)	2.13 (1.72)[WL]	2.95 (1.80)[R]	0.018
Binge eating	12.90 (6.57)[WL,R]	6.16 (5.35)[WL,NR]	17.80 (16.2)[NR,R]	<0.001

WL, wait-list group; R, responder group; nr, non-responder group.
Notes. Superscript specifies groups that a specific group significantly differs from using Tukey's HSD post-hoc comparison approach (*p* < 0.05).

Table 3. Change in correlations among modeled variables at baseline v. follow-up for the responder group

Responders	1	2	3	4	5	6	7
1. Restraint							
2. Eating concerns	−0.07						
3. Overvaluation	−0.04	−0.10					
4. Dissatisfaction	−0.17	−0.05	−0.12				
5. Preoccupation	0.04	0.01	−0.12	−0.30*			
6. Fear weight gain	−0.02	−0.02	0.05	0.07	−0.16		
7. Depression	0.27†	0.21	0.26†	0.22	−0.04	0.24	
8. Binge eating	−0.36*	0.07	−0.07	0.00	0.00	0.04	−0.28†
Non-responders	1	2	3	4	5	6	7
1. Restraint							
2. Eating concerns	0.04						
3. Overvaluation	−0.01	0.06					
4. Dissatisfaction	0.10	0.12	0.16				
5. Preoccupation	0.16	0.04	0.08	0.11			
6. Fear weight gain	0.10	0.02	−0.27**	−0.03	0.05		
7. Depression	0.04	−0.17	0.16	0.28**	−0.15	−0.06	
8. Binge eating	0.13	0.19†	0.35***	0.09	0.14	0.21†	0.00
Wait-list	1	2	3	4	5	6	7
1. Restraint							
2. Eating concerns	0.04						
3. Overvaluation	0.10	0.07					
4. Dissatisfaction	0.04	0.15†	0.11				
5. Preoccupation	−0.01	0.17**	−0.02	0.08			
6. Fear weight gain	−0.16†	0.00	0.05	0.20*	0.05		
7. Depression	−0.01	0.02	0.08	0.17†	−0.11	0.06	
8. Binge eating	0.03	0.11	0.09	0.18*	0.17†	−0.06	−0.01

Notes. Positive values indicate correlation became increasingly positive over time.
*p < 0.05, ** *p* < 0.01, ***p < 0.001, †*p* < 0.10.
Changes in correlation greater than 0.2 may be deemed practically significant.

associations proposed to be disrupted by the intervention, with a null difference for the body dissatisfaction-binge eating relationship. These disruptions were not evident through standard metrics of network analysis (symptom strength metrics and invariance testing of global strength metrics). Symptom-level and global strength statistics from network analysis may be insensitive to changes in specific associations, especially when these constitute a small fraction of the overall number of connections in the network.

Separating intervention participants into responder and non-responder subgroups highlights that the intervention achieved a desirable outcome for approximately 40% of participants, but that many who received the intervention did not benefit to this level. Moreover, changes in correlation patterns from baseline to post-intervention were quite different for these two subgroups. Modeling the intervention participants as a single group would have missed the hypothesized disruption to key symptom associations that were detected in the present study for the responder subgroup, and may partially account for unexpected results from prior attempts to investigate reductions in networks post-intervention (e.g. Beard et al., 2016; Smith et al., 2019).

More explicit declaration of the pathways expected to be disrupted by intervention enables broader interrogation of whether the intervention worked as intended. Though replication is needed to establish stability of key findings, we make several preliminary points about what the current study's results may suggest for our CBT-based intervention. The intervention seems to produce positive changes in severity levels for a range of ED symptoms and mental health correlates, but has a modest impact on reducing strength of association among these symptoms. As articulated earlier, we hypothesized only three of 28 associations to be disrupted in this way by the intervention, and there were some positive findings supportive of this. Even so, if disruption to strength of symptom associations is important to initial and sustained symptom improvement for participants as proposed by network theory (Hilbert et al., 2020), then efforts may be needed to incorporate more treatment content that may reduce associations among a wide array of symptoms than demonstrated for our CBT-based intervention.

Further, the intervention was ineffective in reducing the association between body dissatisfaction and binge eating. Inspection of this correlation at baseline and post-intervention shows that the relationship was already close to zero, and may thus have not benefitted from interventional messages to reduce the functional relationship between this type of domain-specific negative emotion and binge eating. The increased strength of associations between depressive symptoms and both dietary restraint and overvaluation certainly was not an intended outcome for this intervention, and thus also merits consideration. Given that these symptoms reduced for the responder group, the increased correlations may reflect closer coupling of low levels of depressive symptoms and low levels of dietary restraint at post-intervention. Further, reduced correlation between body dissatisfaction and preoccupation was surprising as the intervention had limited content attempting to reduce preoccupation.

Findings should be interpreted with caveats of several study limitations. First, as intervention groups were defined by change over time, there is potential non-equivalence of groups in this study. While we observed no significant group differences in our key correlations (binge eating with dietary restraint, body dissatisfaction, and negative mood) at baseline, baseline mean and variability of binge eating was higher for the responder group.

Reductions in symptomsdue to the intervention may thus also influence associations with other variables. Even so, negative associations were observed at post-intervention between variability in symptoms and level of variance explained, suggesting that this difference in variability may not have had a pronounced effect on hypothesis testing in the current study. Second, although acceptable levels of stability were observed in our present study for network centrality metrics, sample sizes were small (especially for the intervention arms). Present findings should thus be viewed as preliminary until replication in larger, independent samples.

Third, as is common for intervention studies, efficacy was evaluated at the group level rather than per participant. Ecological momentary assessment (EMA) paradigms in which participant data are collected throughout the intervention period may provide more fine-grained and individual-level (idiographic; e.g. Howe, Bosley, & Fisher, 2020) data to enable evaluation of for whom and how many individuals exhibit the proposed disruptions to intervention-targeted symptom associations. Such an individual-level perspective might enable tailoring of the intervention to specific relationships that may differ across individuals.

Fourth, a limitation is that our analyses are based on cross-sections of data, enabling snapshots in time rather than modeling change as a dynamic process. Analytic approaches (e.g. network slope analysis; Nicholas, Cusack, & Levinson, 2023) and study designs (e.g. intensive longitudinal designs; Howe et al., 2020) may allow researchers to address complementary questions about timing of a change in symptom and symptom relationships consistent with a 'disruption' perspective of symptom networks.

A fifth limitation concerns the limited number of symptoms assessed as part of the network models. For example, we used a brief two-item measure of depression as a proxy for negative mood states, which may be an imperfect measure of this construct and its features that are thought to contribute to binge eating. Future research would benefit from replicating our approach by testing a broader range of symptoms, risk factors, and putative mechanisms.

A sixth limitation concerns the homogenous sample with respect to sociodemographic characteristics. Present findings may not necessarily generalize to persons of different genders, racial backgrounds, or educational history. Replication among broader samples is required.

Despite this, findings are broadly consistent with our contention that study design and analytic approach influence ability to test network theory predictions with fidelity. We thus make several recommendations for future network theory-driven interventional studies:

(i) More detailed articulation of which symptom associations are expected to be disrupted by one's intervention. This should include detailing the treatment content designed to produce this disruption. Some studies may fail to observe network theory-predicted disruptions to the network because the intervention's mechanism of action is antecedent reduction rather than disruption of symptom associations. Others may fail to observe the effect, particularly if using global strength metrics, because the number of associations expected to be disrupted is a small subset of the total number of symptom associations within one's network.

(ii) It is possible that these disruption-focused interventional approaches are uncommon in psychological treatments. Appropriate testing of this component of network theory will require a careful inventory of treatments that purport

to work in ways proposed by network theory. In the absence of plausible candidates, researchers may also need to devise new interventions to enable testing of network theory predictions about the importance of disrupting symptom associations.

(iii) Broadening the range of statistical tests conducted in testing of these interventional effects. Global statistics may be too insensitive to detect meaningful changes in the symptom network post-intervention. Changes in variability and heterogeneity in treatment responsiveness may also create statistical artifacts that mask effects that are present in one's data. In combination, checks for variability and its association with variance explained, bivariate correlations among symptoms within and across time, and more standard network analysis statistics are likely to yield a more complete picture than network analysis statistics alone.

(iv) Other study designs are needed to further interrogate the value of disrupting symptom association in a treatment context. A common network theory-based argument is that the high rate of relapse in psychological conditions may be due to failure to disrupt symptom associations (Lorimer, Delgadillo, Kellett, & Brown, 2020); though symptoms may initially reduce, gradual increases in antecedents could ultimately lead to return of symptom levels. Head-to-head comparisons between interventions designed to cause disruption *v.* those designed to reduce symptoms would help to confirm or disconfirm the hypothesis that disruption to symptom associations is important for preventing symptom (or condition) relapse. Longer-term follow-ups in studies employing these disruptive interventions would also help to determine whether disruptions at post-intervention promote maintained symptom improvements across longer time horizons.

References

Beard, C., Millner, A. J., Forgeard, M. J., Fried, E. I., Hsu, K. J., Treadway, M. T., …Björgvinsson, T. .. (2016). Network analysis of depression and anxiety symptom relationships in a psychiatric sample. *Psychological Medicine*, *46* (16), 3359–3369.

Bland, J. M., & Altman, D. G. (2011). Correlation in restricted ranges of data. *BMJ*, *342*, d556–d556. https://doi.org/10.1136/bmj.d556

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, *16*(1), 5–13.

Bringman, L. F., Elmer, T., Epskamp, S., Krause, R. W., Schoch, D., Wichers, M., … Snippe, E. (2019). What do centrality measures measure in psychological networks? *Journal of Abnormal Psychology*, *128*(8), 892–903.

Calugi, S., Dametti, L., Chimini, M., Dalle Grave, A., & Dalle Grave, R. (2021). Change in eating-disorder psychopathology network structure in patients with anorexia nervosa treated with intensive cognitive behavior therapy. *International Journal of Eating Disorders*, *54*(10), 1800–1809.

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, *50*, 195–212.

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.

Fairburn, C. G., & Beglin, S. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, *16*, 363–370. doi: 10.1002/1098-108X

Fairburn, C. G., Cooper, Z., & Shafran, R. (2003). Cognitive behaviour therapy for eating disorders: A "transdiagnostic" theory and treatment. *Behavior Research and Therapy*, *41*, 509–528. doi: 10.1016/S0005-7967(02)00088-8

Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, *40*(5), 532–538.

Forrest, L. N., Franko, D. L., Thompson-Brenner, H., & Grilo, C. M. (2023). Examining changes in binge-eating disorder network centrality and structure in patients treated with cognitive-behavioral therapy versus interpersonal psychotherapy. *International Journal of Eating Disorders*, *56*(5), 944–955.

Forrest, L. N., & Grilo, C. M. (2022). Change in eating-disorder psychopathology network structure in patients with binge-eating disorder: Findings from treatment trial with 12-month follow-up. *Journal Of Consulting And Clinical Psychology*, *90*(6), 491.

Fried, E. I., van Borkulo, C. D., Cramer, A. O., Boschloo, L., Schoevers, R. A., & Borsboom, D. (2017). Mental disorders as networks of problems: A review of recent insights. *Social Psychiatry and Psychiatric Epidemiology*, *52*(1), 1–10.

Hilbert, A., Herpertz, S., Zipfel, S., Tuschen-Caffier, B., Friederich, H.-C., Mayr, A., & de Zwaan, M. (2020). Psychopathological networks in cognitive-behavioral treatments for binge-eating disorder. *Psychotherapy And Psychosomatics*, *89*(6), 379–385.

Hilbert, A., Petroff, D., Herpertz, S., Pietrowsky, R., Tuschen-Caffier, B., Vocks, S., & Schmidt, R. (2019). Meta-analysis of the efficacy of psychological and medical treatments for binge-eating disorder. *Journal of Consulting and Clinical Psychology*, *87*(1), 91–105.

Howe, E., Bosley, H. G., & Fisher, A. J. (2020). Ideographic network analysis of discrete mood states prior to treatment. *Counselling & Psychotherapy Research*, *20*(3), 470–378. doi: 10.1002/capr.12295

Jones, P. J. (2023). *networktools: Tools for identifying important nodes in networks.* [statistical software].

Kroenke, K., Spitzer, R. L., Williams, J. B., & Löwe, B. (2009). An ultra-brief screening scale for anxiety and depression: The PHQ–4. *Psychosomatics*, *50*, 613–621. doi: 10.1016/S0033-3182(09)70864-3

Linardon, J., Rosato, J., & Messer, M. (2020a). Break binge eating: Reach, engagement, and user profile of an internet-based psychoeducational and self-help platform for eating disorders. *International Journal of Eating Disorders*, *53*, 1719–1728. doi: 10.1002/eat.23356

Linardon, J., Shatte, A., McClure, Z., & Fuller-Tyszkiewicz, M. (2023). A broad v. focused digital intervention for recurrent binge eating: A randomized controlled non-inferiority trial. *Psychological Medicine*, *53*(10), 4580–4591.

Linardon, J., Shatte, A., Messer, M., Firth, J., & Fuller-Tyszkiewicz, M. (2020b). E-mental health interventions for the treatment and prevention of eating disorders: An updated systematic review and meta-analysis. *Journal Of Consulting And Clinical Psychology*, *88*, 994–1007. doi: 10.1037/ccp0000575

Linardon, J., Shatte, A., Rosato, J., & Fuller-Tyszkiewicz, M. (2020c). Efficacy of a transdiagnostic cognitive-behavioural intervention for eating disorder psychopathology delivered through a smartphone app: A randomized controlled trial. *Psychological Medicine*, *52*, 1–12. doi: 10.1017/S00332917 20003426

Lorimer, B., Delgadillo, J., Kellett, S., & Brown, G. (2020). Exploring relapse through a network analysis of residual depression and anxiety symptoms after cognitive behavioural therapy: A proof-of-concept study. *Psychotherapy Research*, *30*(5), 650–661. doi: 10.1080/10503307.2019.1650980

Lydecker, J. A., White, M. A., & Grilo, C. M. (2017). Form and formulation: Examining the distinctiveness of body image constructs in treatment-seeking patients with binge-eating disorder. *Journal Of Consulting And Clinical Psychology*, *85*, 1095–1103. doi: 10.1037/ccp0000258

McNally, R. J. (2016). Can network analysis transform psychopathology? *Behaviour Research And Therapy*, *86*, 95–104. https://doi.org/10.1016/j.brat.2016.06.006

Mitchison, D., Hay, P., Griffiths, S., Murray, S. B., Bentley, C., Gratwick-Sarll, K.,…Mond, J. (2017). Disentangling body image: The relative associations of overvaluation, dissatisfaction, and preoccupation with psychological distress and eating disorder behaviors in male and female adolescents. *International Journal of Eating Disorders*, *50*(2), 118–126.

Monteleone, A. M., Cardi, V., Ambwani, S., Cascino, G., Albano, G., Pellegrino, F., & Treasure, J. (2021). Network intervention analysis to assess the trajectory of change and treatment effects associated with the use of online guided self-help for anorexia nervosa. *Early Intervention in Psychiatry*, *15*(5), 1210–1216.

Monteleone, A. M., & Cascino, G. (2021). A systematic review of network analysis studies in eating disorders: Is time to broaden the core psychopathology to non specific symptoms. *European Eating Disorders Review*, *29*(4), 531–547.

Mullen, R., & Jones, E. S. (2021). Network anlaysis of competitive state anxiety. *Frontiers in Psychology*, *11*, 586976.

Nicholas, J., Cusack, C., & Levinson, C. (2023). Eating disorder symptom and fear change trajectories during imaginal exposure therapy: A slope network analysis. *Behavior Therapy*, *54*(2), 346–360. doi: 10.1016/j.beth.2022.09.006

R Core Team. (2022). *R: a language and environment for statistical computing. R foundation for statistical computing.* Vienna, Austria: R Core Team.

Smith, K. E., Mason, T. B., Crosby, R. D., Cao, L., Leonard, R. C., Wetterneck, C. T., … Wonderlich, S. A. (2019). A comparative network analysis of eating disorder psychopathology and co-occurring depression and anxiety symptoms before and after treatment. *Psychological Medicine*, *49*(2), 314–324.

van Borkulo, C., Boschloo, L., Kossakowski, J., Tio, P., Schoevers, R., Borsboom, D., & Waldorp, L. (2022). Comparing network structures on three aspects: A permutation test. *Psychological Methods*, *28*(6):1273–1285. doi: 10.1037/met0000476.

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*(3), 1–67.

Williams, S. E., Watts, T. K. O., & Wade, T. D. (2012). A review of the definitions of outcome used in the treatment of bulimia nervosa. *Clinical Psychology Review*, *32*, 292–300. doi: 10.1016/j.cpr.2012.01.006