Mapping Post-Traumatic Stress Disorder: A Proof of Concept for Network-Based Case Formulation

Elizabeth Blomfield *

Royal Holloway University of London

Nimrod Hertz *

MRC Cognition and Brain Sciences Unit, Cambridge University

Sharif El Leithy

Traumatic Stress Service, South West London and St George's Mental Health NHS Trust, London, UK

Tim Dalgleish

MRC Cognition and Brain Sciences Unit and Department of Psychology, Cambridge University

Anna Bevan

MRC Cognition and Brain Sciences Unit, Cambridge University

Gary Brown *

Royal Holloway University of London

*Joint lead authors

Corresponding authors

Nimrod Hertz

MRC Cognition and Brain Sciences Unit,

University of Cambridge, UK

15 Chaucer Road, Cambridge CB2 7EF

Nimrod.Hertz@mrc-cbu.cam.ac.uk

Gary Brown

Department of Psychology

Royal Holloway University of London, UK

Wolfson Building, RHUL, Egham, Surrey, TW20 0EX

Gary.Brown@rhul.ac.uk

Abstract

Background and Objective. The network approach to psychopathology views mental disorders as dynamic interactions between symptoms, challenging traditional latent-cause models. We applied this approach to individuals with PTSD, using ecological momentary assessment (EMA) data to build personalized contemporaneous and directed symptom networks. Our primary aim was to assess whether these network-derived formulations aligned with clinician-generated cognitive formulations. We also tested whether cognitive variables predicted affective variables over time, consistent with the Ehlers and Clark (2000) cognitive model of PTSD.

Method. Five participants completed 14-21 days of EMA (response rate=90.2±5.48%), yielding 348 observations total. Items were tailored to each patient's lived experience and mapped to Ehlers and Clark components. Written formulations from patients and clinicians were compared to EMA-based networks generated by three network algorithms.

Results. Network-derived formulations shared some features with patient and clinician-derived formulations but often diverged in complexity and specific predictions. Data-driven networks had good precision (0.64-0.86) but low sensitivity (0.31-0.35), showing better agreement with patient self-formulations than with clinician models. Evidence for lead-lag effects between cognitive and affective variables was limited. Trauma memory emerged as the most central component, with its association to sense of current threat and matching triggers being the most prominent.

Conclusion. Network analysis may refine clinical hypotheses but has limitations in replicating comprehensive formulations. Instead, we recommend its use for clinical hypothesis testing of specific predictions. Further studies, including randomized trials, are needed to test its clinical utility and contribution to treatment outcomes.

Introduction

The widespread use of mobile technology offers new opportunities for continuous in vivo collection of ecological momentary assessment (EMA) data with virtually limitless applications. This has particular promise within healthcare (Balaskas et al., 2021). However, to date, the resulting abundance of data has often been a mixed blessing, as researchers seek the most fitting approach for structuring and making sense of what it provides. Where largely data-driven, machine learning approaches have been applied to these data, there has been success reported in predicting, for example, depression symptoms (e.g., Jacobson & Chung, 2020). However, when *a priori* theoretical predictions have been tested, the outcome has frequently not supported predictions (e.g., Schreuder et al., 2022). The increased take-up of EMA collection methods paralleled development of the network approach to psychological measurement (see Bringmann et al., 2022) in which variables of interest are analyzed as nodes in a network of related variables, and these and related techniques are often drawn upon to analyze EMA data. Early network studies were cross-sectional, that is, based on a single point of data collection, and premised upon causal claims about the centrality parameters that can be calculated from network data, a line of research that was ultimately thrown into question (Bringmann et al., 2019). Longitudinal network studies, many of which have employed EMA data collection, offer the opportunity to study patterns of change over time (Bringmann et al., 2017). Temporal precedence permits the estimation of

directional relationships (Eskamp et al, 2018), comparable to the long established cross-lagged panel design, but are only suggestive of causality as they cannot inherently rule out sources of confounding. Recent efforts that have sought to identify parameters drawn from complex systems approaches within these streams of data, such as hysteresis and critical slowing (Smit et al., 2023), have also been inconclusive.

The network approach has continued to expand and to hold out the prospect of a "new science of mental disorders" (Roefs et al., 2022), frequently with the tacit promise of overcoming barriers that had limited more traditional approaches but perpetually running up against the inherent limits of causal inference based on data alone. In this regard, Pearl and colleagues (Pearl, 2009) and others (Hernan & Robins, 2024) have set out frameworks addressing fundamental issues of causality that clarify when causal inferences can be drawn from observational data as an alternative approach to experimental designs and randomized trials (Vandenbroucke et al., 2016). These approaches take as a starting point a grounding in real world knowledge of the area of interest sufficient for establishing a plausible account of the processes that generate the data in question so that study design, measurement, and analysis can be carried out accordingly. As Rohrer (2018) has stated, "Drawing valid causal inferences on the basis of observational data is not a mechanistic procedure but rather always depends on assumptions that require domain knowledge and that can be more or less plausible."

The relevant domain knowledge in mental health is, in part, embedded in the diagnostic criteria for different mental disorders. The diagnosis serves as a heuristic for the clinician to adapt a treatment model for the diagnosis in question to the circumstances of the specific case through case formulation. There have been

longstanding efforts to better operationalize and test formulation, but these have encountered a set of perennial challenges, namely difficulty establishing inter-judge reliability (Bieling & Kuyken, 2003) and lack of standardized approaches to developing formulations (Mumma et al., 2018); see Brown (2022) for a review. Hallam (2013) recently proposed a comprehensive method for developing clinical formulations grounded in functional analysis, a well-established systematic method that evolved alongside behaviorism and is integral to clinical practice within that orientation. A number of other recent initiatives geared to systematizing the gathering of the information provided within clinical presenting problems have also set out to incorporate elements of functional analysis (Burger et al., 2024, Davison, 2018; Fisher, 2015, Scholten et al., 2022). The emphasis in Hallam's model is on the transparent and systematic gathering of data that is represented atheoretically within a flow diagram, with the aim of underpinning the explanations (and, therefore, causal reasoning) necessary to intervene effectively. Hallam and colleagues (Hallam et al., 2024) developed a rating scale of individual case formulation (ICF) based on the attempts of novices to master the system. The scale reflects a progression from evaluating the quality, clarity, and objectivity of observations to considering the causal relationships between observations and how these are explained, culminating with an appraisal of whether the formulation can be judged to be a sufficient basis for intervention. As such, the scale can be seen as proceeding along the "ladder of causality" Pearl (Pearl & Mackenzie, 2018) sets out to represent the progression from Rung 1, noticing that A and B appear together (data shows a pattern or correlation), to Rung 2, see that if B changes, A changes too (cause and effect through intervention), to Rung 3, asking what would happen to A if B were different, even if B cannot be changed in reality (using counterfactual,

“what if” thinking), which is the level required to envisage the right treatment for a specific person to address their specific problem. Hallam et al. found the reliability of ratings of the formulation diagrams created by novices was equal or better than the previous efforts cited in this regard, but that there was still considerable scope for improvement. The approach has not yet been evaluated among more experienced practitioners.

Among other functions, a formulation should serve as a guide to the further information necessary to characterize the features of the individual case and to provide a basis for tracking progress over the course of treatment. There have been approaches suggested for anchoring EMA assessment within the traditional clinical approach to formulation (e.g., Burger et al., 2021) and anticipation that network derived findings can augment formulation (Roefs et al., 2022). The present study was undertaken to work out the logistics of implementing formulation-guided EMA assessment within a sample of posttraumatic stress patients studied intensively over the span of a few weeks. PTSD is a promising context for exploring this, given the availability of well-worked-out general models that can serve as a shared point of reference for network models and therapist-derived formulations. A broadened perspective was adopted that included all aspects of the relevant clinical theoretical model (here, Ehlers and Clark’s [2000] PTSD treatment model) rather than being confined to the idiosyncratic notion of “symptom dynamics” at the core of the network approach as it has largely been advanced. We sought to take into account the full array of pertinent information (treatment model, clinical formulation, patient view of their problem, and cross-sectional and lagged empirical relationships between variables operationalizing these models), to work out methods for analyzing these relationships and, where possible, to test the implied theoretical

predictions. An individualized battery was constructed for each participant, comprising a combination of standardized and tailored assessment items that were administered multiple times daily over the course of the assessment period. Different approaches to network estimation were compared to clinician formulations constructed by their psychotherapy provider against the background of Hallam's ICF approach. This aimed to address the essential question of whether empirically derived clinical representations, such as those provided by network analysis, approximate clinician-derived formulations. We hypothesized (1) that patient, therapist, and empirical models of the presenting problem would resemble each other and (2) that implications of the theoretical model could be tested with the empirical data, with the prediction of greatest interest being the lead-lag relationship between cognitive and affect variables (high emotionality and a sense of severe current threat), the central expectation of the CBT model.

Method

Participants

Five participants (see Table 1 for demographic and clinical details) were recruited from UK National Health Service (NHS) departments providing specialist treatment for PTSD. Inclusion criteria were age 18 or older with a primary diagnosis of PTSD, at any stage of therapy for PTSD within an NHS service and having a clinician willing to submit a cognitive formulation for them. The sample was all female, three of whom were Caucasian, age range 21 to 51 years (m=33.20, SD=11.03). Two participants met criteria for at least one comorbid diagnosis, which included depression (n=1), generalised anxiety disorder (n=1) and borderline personality disorder (n=1). The was no minimum required level of therapist experience, which ranged from newly

qualified to highly experienced (see Table 1). However, all therapists were newly introduced to the formulation model used in the study.

Table 1

Participant Demographics

Participant	Age	Gender	Ethnicity	Diagnoses	Clinician – Years of Post-Qualification Experience
101	35	Female	White British	PTSD	20
102	21	Female	White British	PTSD, Depression, Generalised Anxiety Disorder	15
103	37	Female	White British	PTSD, Depression, Borderline Personality Disorder	3.5
104	51	Female	Indian	PTSD	7.5
105	22	Female	Middle Eastern: Arab	PTSD	1

NHS ethical approval was obtained from South Central Berkshire B Research Ethics Committee of the NHS Health Research Authority. Following ethical approval and agreement of the NHS service to participate, clinicians approached patients fitting recruitment criteria. Potential participants were then contacted by a researcher, who explained the study and took informed consent if they wished to proceed. All participants completed consent forms and were free to withdraw at any point.

Procedures

The study included four phases: a) recruitment of clinicians and patients, b) completion of formulation training and submission of patient formulation and predicted symptom relationships table by clinician, completion of baseline measures and predicted symptom relationships table by patient, c) 14-day EMA sampling period by patient, and d) completion of end of study measures by the patient. In recognition of the intensive data collection required by this study, participants were compensated for their time, up to a limit of £6/day.

Table 2. Selected items per participant

Participant	Memory items	Negative appraisals	Threat items	Strategies	Moderators
101	Memory Flashback / Nightmare	People can't be trusted You can never know who will harm you The event happened because of the sort of person I am I feel isolated and set apart from others	Upset / strong physical reaction Superalert, watchful or on guard Jumpy / easily startled	I dwell on how I used to be before the event I dwell on what other people have done to me I drink alcohol, take medication or use drugs I have had an alcoholic drink today	I am tired I am with friends
102	Memory Flashback	I am inadequate I have to be on guard all the time You can never know who will harm you I feel like an object, not like a person	Upset / strong physical reaction Superalert, watchful or on guard Jumpy / easily startled	I dwell on what I should have done differently I numb my feelings I try to push memories out of my mind I work hard at keeping busy with other things	I am alone I have eaten a normal amount of food for my last meal
103	Disturbing & unwanted	There is something wrong with me as a	Upset / strong physical reactions	I try hard to control my	It is hard to hold a

	memories Nightmares and flashbacks	person I can't rely on myself I will not be able to control my anger and will do something terrible People can't be trusted	Superalert, watchful or on guard Jumpy / easily startled I got no sleep in the last 24 hours	emotions I numb my feelings I drink alcohol, take medications or use drugs	thought in my mind for so long I feel I can't remember things accurately
104	Disturbing & unwanted memories Nightmares and flashbacks	I have to be on guard all the time You can never know who will harm you I will never be able to feel normal emotions again I wish I were invisible I feel mentally robbed	Upset / strong physical reactions Superalert, watchful or on guard Jumpy / easily startled	I dwell on what other people have done to me I go over what happened again and again I drift off into a world of my own	
105	Disturbing & unwanted memories Nightmares and flashbacks	There is something about me that made the event happen I will never be able to feel normal emotions again People can't be trusted	Upset / strong physical reactions Superalert, watchful or on guard Jumpy / easily startled I have slept less than 3 hours	I work hard at keeping busy with other things I worry that something similar will happen to me or my family I detach myself from the memories	I am in an academic lecture / seminar My sister is out socialising

Measures

Three self-report measures used within NHS services providing treatment for PTSD were completed by participants to establish symptom severity at baseline and provide a basis for their individualised survey. They repeated the same measures at the end of the 14-day EMA period.

PTSD Checklist for DSM-5 (PCL-5; Weathers et al, 2013). This 20-item self-report measure assesses the severity of PTSD symptoms over the last month. Items

correspond to DSM-5 criteria for PTSD (APA, 2013) and are rated on a five-point scale (0=*Not at all* to 4=*Extremely*).

Post-Traumatic Cognitions Inventory (PTCI; Foa et al., 1999). This is a 33-item self-report measure evaluates strength of belief in trauma related cognitions. Its three sub-scales are Negative Cognitions About the Self, Negative Cognitions About the World, and Self-Blame. Items are rated on a 7-point Likert scale (1=*totally disagree*, 7=*totally agree*). It has been found to differentiate well between individuals with and without a diagnosis of PTSD (sensitivity=0.78. specificity=0.93; Foa et al., 1999).

Response to Intrusions Questionnaire (RIQ; Clohessy & Ehlers, 1999; Murray, Ehlers, & Mayou, 2002). This 19-item self-report questionnaire assesses maladaptive responses to intrusions following a traumatic experience, specifically evaluating rumination, suppression, and numbing which form subscales. Items are rated on a 4-point scale (0=Never, 3=Always). The RIQ has been shown to predict PTSD outcomes (e.g., Kleim et al., 2012; Wild et al., 2016).

Personal Survey. Each participant provided responses to a personalised survey during the EMA collection period, constructed with reference to the standard CBT treatment model of Ehlers and Clark (2000) using a combination of items from the measures described above and personalized contextual details relevant to the features of their case. The format for each participant's survey was standardised and asked about four types of variables: core PTSD symptoms, cognitive symptoms, affective symptoms and contextual moderators, and cognitive-behavioral strategies. The first question checked for the occurrence of a trigger for symptoms since the last questionnaire and asked the participant to describe it. Questions 2 through 6

were based on the PCL-5 and concerning whether key symptoms associated with PTSD (memories, reexperiencing, distress/physical reactions, hypervigilance, feeling jumpy) had been experienced since the last assessment. The same items were administered to all participants at each assessment. Question 7 asked participants to rate their agreement with the three thoughts or beliefs with which they had indicated the strongest agreement on their initial PTCI. Question 8 asked them to rate the frequency with which they responded to intrusions using the three maladaptive strategies they identified as using most often on the RIQ. If more than three items had maximum scores on either the PTCI or RIQ, the items included were those the participant judged were most characteristic of their experience of PTSD and how they responded to symptoms. Questions 9 and 10 were left open for participants to indicate the presence/absence at the time of responding of personally chosen moderators which they felt affected their responses to symptoms. All personalised surveys consisted of 10-12 questions designed to take 3-4 minutes or less to complete. The final version of each participant's personalised survey was agreed in a call between the participant and a researcher prior to the start of EMA sampling.

Predicted Relationships Table. This was administered to obtain a representation of their problem from each participant analogous to the formulation provided by the clinician. Participants completed a table which cross-listed the items from their personalised survey, using an X to indicate any two items they perceived to co-occur or be related in some way, in response to the prompt 'Please complete the table, below, by placing an X in any unshaded box where the two items in question are related to each other in your life'. Participants were free to indicate as many or as few relationships between items as they perceived existed. This method is similar

to that of previous research that obtained subjective patient estimations to derive Perceived Causal Networks (Klintwall et al., 2021), or the basis for generating priors for Bayesian network analysis (Burger, Epskamp, et al., 2022). Each clinician also completed the same table for their patient.

14-day EMA Sampling Period. Participant surveys were hosted on the Qualtrics online platform (Version 2021), which is compliant with GDPR regulations. Following creation of their personalised survey, data collection proceeded using the EMA approach (Stone & Shiffman, 1994). Participants received a personalised survey via email to their mobile phone five times a day for fourteen days starting the day after their next therapy session. Surveys arrived every three hours within a 12-hour window. The start and end times for this window was decided by the participant to best reflect their typical routine and experience of symptoms (e.g., from 9 a.m. – 9 p.m.). Questionnaires expired and could no longer be completed two hours after arrival to ensure spacing of responses. Surveys were time-stamped with the time of completion. The response rate was very high in our sample, with an average of 90.2% (SD = 5.48%).

Clinician formulations**.** Clinicians whose patients had agreed to participate in the study completed an online training module in Individual Case Formulation (ICF; Hallam, 2013) for PTSD, which was the standard model for formulation used in the study. The same training procedure is described and evaluated in Hallam et al. (2024). ICF is a formulation approach founded on principles of functional analysis, which aims to precisely specify causal relationships, contextual moderators, and mediators. The ICF model for PTSD communicated in the training material incorporated Ehlers and Clark's (2000) cognitive model of PTSD. As such, it provided a common basis for comparison with network models, which sought to identify all

the factors contributing to and maintaining dysfunction. Following training, each clinician uploaded an anonymised cognitive case formulation for their patient following the ICF format. Clinicians completed the expected relationship table following constructing an ICF formulation.

Data Analysis

Each participant completed a personalised short questionnaire five times daily for fourteen days (except the first participant, who requested to continue to collect data for twenty-one days). This produced up to seventy data points per question provided there was no missing data. Vrijen et al., (2018) have recommended 20-50 observations per subject in order to estimate edges in a directed network, although simulation studies have shown that in directed networks, sensitivity is low even for large sample sizes that are unrealistic to achieve in a clinical setting (Mansueto et al., 2022). Our sampling rate (five surveys per day) and the duration of the data collection period (14 days) were selected as likely to produce sufficient observations to allow estimation of network models, allowing for missing data, balanced against the need to ensure that data collection was not overly burdensome. The protocol was piloted by a researcher ahead of the study start.

To facilitate analysis, single items were grouped into five Ehlers and Clark model (2020) elements: negative appraisals (PTCI), trauma memory (PCL-5 re-experiencing), matching triggers, sense of current threat (PCL-5 arousal/reactivity), and maladaptive strategies (RIQ). The specific items selected by each participant are detailed in Table 2.

Our primary network analyses applied three algorithms:

1. Mixed Graphical Models (MGM). We opted for MGMs over Graphical Gaussian Models (GGM) to estimate cross-sectional networks due to the mixture of dichotomous nodes (e.g., the presence vs absence of a trigger variable within the assessment window) and continuous nodes (Altenbuchinger et al., 2020). To account for the hierarchical data structure, continuous variables were mean-centered within participants to enable estimation of intra-individual fixed effects (Epskamp, Waldorp, et al., 2018). Edge weights quantified relationship strength, with edge thickness used graphically to indicate strength of relationship and color (blue for positive and red for negative) representing sign. Regularization was achieved using LASSO with an Extended Bayesian Information Criterion (EBIC) to select the optimal model by balancing goodness-of-fit and model complexity. Gamma was set at 0.25, which is considered a conservative parameter for discovery, balancing sensitivity and specificity (Beck & Jackson, 2020; Burger, Isvoranu, et al., 2022). MGMs were computed using the *bootnet* package (Epskamp, Borsboom, et al., 2018), which calls the *mgm* function from the *mgm* package (Haslbeck & Waldorp, 2020).

2. Multilevel Vector Auto-Regression (VAR). This followed the approach described by Epskamp et al. (2018) to generate two networks: (a) an undirected contemporaneous network, capturing symptom relationships within the same measurement window, and (b) a directed temporal network for prediction across three-hour time windows. Variables were first de-trended for time effects (i.e., spontaneous symptom change) to meet VAR stationarity assumptions. Data missingness was confirmed as completely at random (MCAR) using Little's MCAR test (Little, 1988). In addition, a logistic regression predicting missingness from Day, Beep, ID, and their interactions yielded no significant predictors, further supporting

the MCAR assumption. VAR was carried out using the *mlGraphicalVAR* function from the *graphicalVAR* package (Epskamp, 2017).

3. Directed Acyclic Graphs (DAGs)**.** DAGs were computed with the Restricted Maximisation (RSMAX2) algorithm using the *bnlearn* package (Scutari, 2010), combining constraint-based and score-based algorithms to improve the robustness of the inferred causal structure while mitigating overfitting and false dependencies. We mean-centered each participant's data to isolate within-person effects and account for hierarchical dependencies, a common data preprocessing practice for distinguishing within- from between-person relationships in repeated measures designs (Wang et al., 2019), although specific recommendations for DAGS with hierarchical data do not currently exist. These DAGs were constructed to examine the likely directionality in cross-sectional networks. To ensure robustness, they were bootstrapped 1000 times per person, and the combined average strength and directionality percentage was evaluated.

Network comparisons. Overall, each participant had four data-based network representations (MGM, VAR contemporaneous, VAR temporal, and DAG) and two subjective variable association matrices (clinician formulation and patient perception). For each network representation, we computed strength centrality and ranked the strongest edges. Individual MGM and VAR networks were compared to clinician formulations and patient-estimated matrices[1]. The edge strength in subjective matrices was not quantified, as the clinicians and patients only provided potential edge binary absence or presence.

[1] VAR networks (contemporaneous and temporal) were collapsed into a single network, as the clinician and patient's variable matrices did not differentiate between temporal and contemporaneous predictions.

To examine interpersonal variability in individual networks, we calculated Spearman correlations between corresponding edges for each participant pair (i.e., comparing all possible participant combinations, for example: 101 vs. 102, 101 vs. 103, 102 vs. 103, etc.). This approach quantifies the degree of similarity (or difference) between network structures across individuals. We report the average effect size and the percentage of significant correlations while adjusting for multiple comparisons with the False Discovery Rate (FDR) adjustment (Benjamini & Hochberg, 1995). We then used the Network Comparison Test (NCT; Claudia van Borkulo, 2022) to evaluate differences in network structure and global strength across pairs of participants. However, since NCT is designed for cross-sectional networks, we applied it only to the MGM networks.

We assessed the agreement between network algorithms and subjective estimations using classification metrics derived from confusion matrices, calculated separately for each network algorithm[2] and reference. Binary clinicians and patients' matrices served as the point of reference for calculating sensitivity (true positive rate), specificity (true negative rate), precision (positive predictive value), F1 score (balancing precision and sensitivity), and accuracy (overall correct classification rate). Metrics were evaluated across γ tuning parameters of 0, 0.10, 0.25, 0.50, 0.75, and 0.99. We then estimated mixed-effects linear models comparing classification measures across different tuning parameters and network algorithms (MGM vs VAR; Comparison I) and different references (clinician vs. patient; Comparison II), to compare the network algorithms' performance in recovering the clinical formulations and self-perceived matrices of association. As

[2] VAR contemporaneous and temporal networks were collapsed for these analyses as well

the goal was to assess performance variations rather than establish the superiority of any model, we did not correct for multiple comparisons.

Results

All participants provided sufficient data to estimate non-empty networks for most algorithms. The average density (percent non-zero edges out of all possible edges) was 0.32±0.13 for MGM networks, 0.00±0.00 (empty networks) for temporal VAR, excluding self-loops, and 0.34±0.15 for contemporaneous VAR. Corresponding average edge weights were 0.55±0.09, 0.00±0.00, and 0.14±0.02, respectively. Density did not differ between MGM and VAR contemporaneous networks ($t_{(4)}$=0.53, p=.62), but edge weights were significantly higher in MGM than in the contemporaneous VAR network ($t_{(4)}$=11.55, p=.0003, $p_{adjusted}$=.0006). Figure 1 depicts the individual networks across algorithms, clinicians' formulations and patients' perceived relationship matrices with reference to the diagram depicting the relationship between the elements of Ehlers and Clark's (2000) PTSD model.

Clinician formulations and patient perceived associations were significantly denser than data-derived networks (Clinician vs. MGM, $t_{(4)}$=3.56, p=.024, p_{adj}=.031, Clinician vs VAR contemporaneous, $t_{(4)}$=3.67, p=.021, p_{adj}=.031; Patient vs. MGM, $t_{(4)}$=3.57, p=.023, p_{adj}=.031, Patient vs. VAR contemporaneous, $t_{(4)}$=2.99, p=.040, p_{adj}=.040). The clinician and patient derived matrices were almost fully-saturated, with an average density of 0.70±0.19 for clinicians, and 0.78±0.18 for patients.

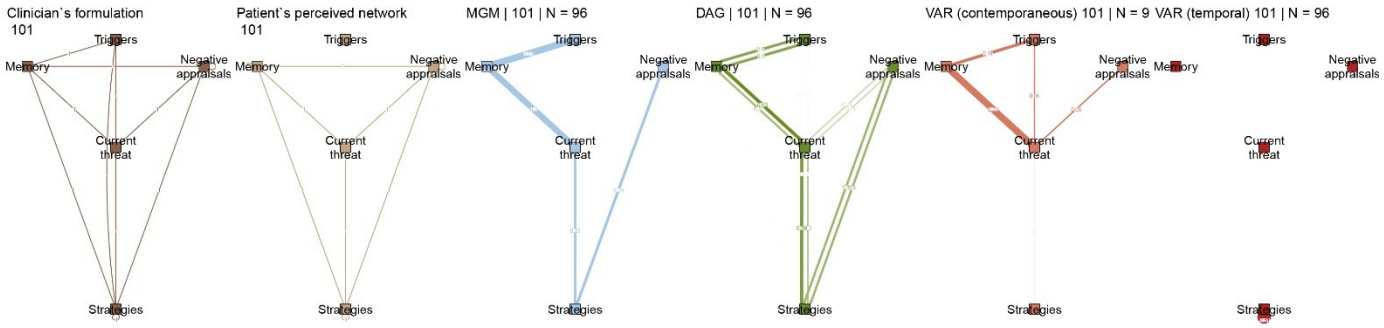
1.1. Network variability between participants

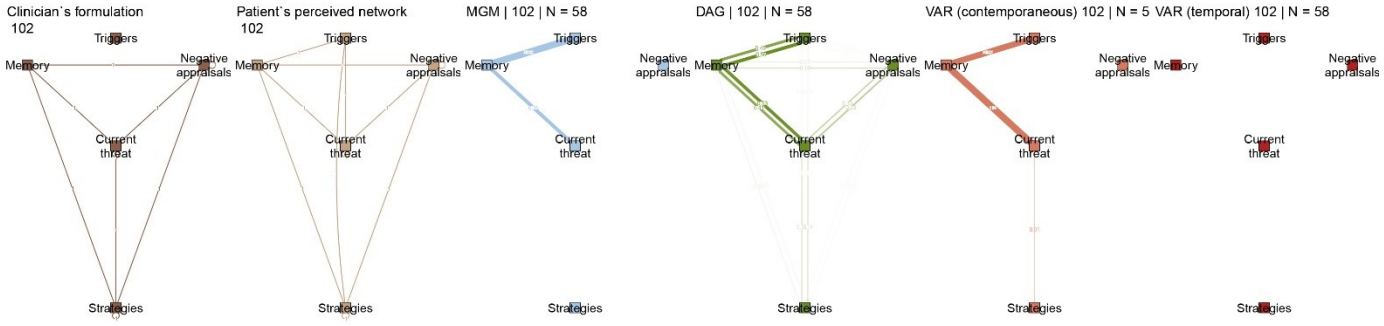
MGM and DAG showed moderate inter-individual correlations. For MGM, the average Spearman coefficient was 0.47 ± 0.19 (Median=0.44, range: 0.25–0.79); for DAG, 0.48 ± 0.15 (Median=0.48, range: 0.28–0.70). DAG networks yielded more

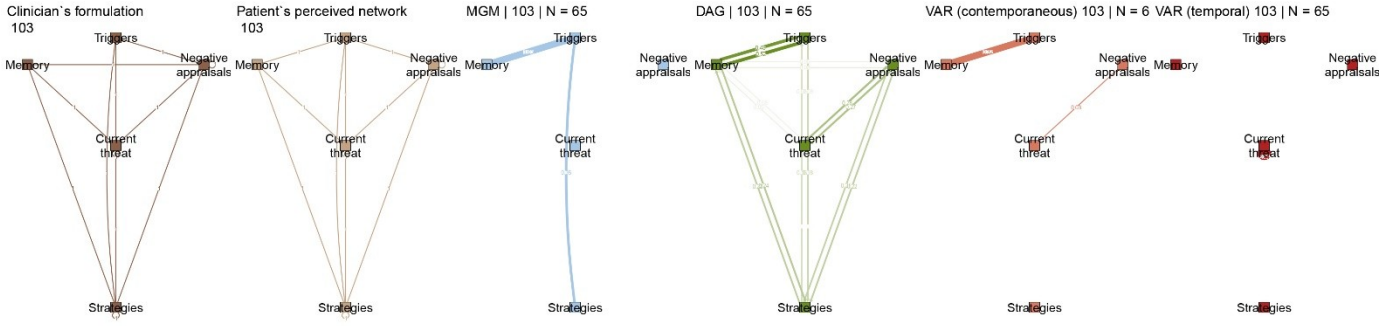
significant correlations (50% of pairwise comparisons) than MGM (20%), likely due to their greater number of directed edges. The Network Comparison Test showed high similarity across networks in both global strength and structure. Only 20% of comparisons (101 vs. 103 and 101 vs. 104) differed significantly in structure, with no significant differences in global strength. In contrast, VAR networks showed little inter-participant similarity. Average correlations were 0.09±0.33 (median=0.10, range: -0.39 to 0.55) for contemporaneous networks, and 0.00 ± 0.15 (median=-0.06, range: -0.15 to 0.27) for temporal networks (including self-loops). None of these were statistically significant.

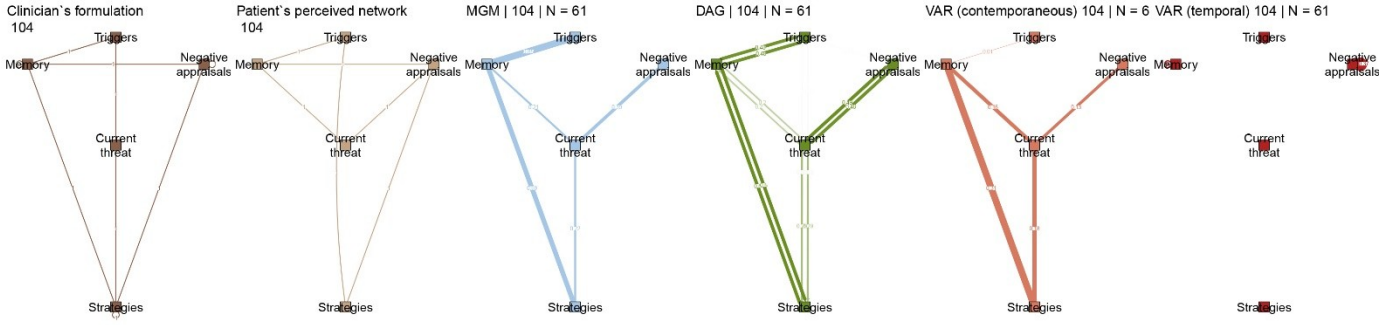
1.2. Correspondence of network estimations with Clinician Formulations and Patient Estimations

The edge-detection classification results across tuning parameters are summarized in Table 3 and depicted in Figure 2.


Clinician`s formulation 101
Patient`s perceived network 101
MGM | 101 | N = 96
DAG | 101 | N = 96
VAR (contemporaneous) 101 | N = 9
VAR (temporal) 101 | N = 96
Triggers
Memory
Negative appraisals
Current threat
Strategies


Clinician`s formulation 102
Patient`s perceived network 102
MGM | 102 | N = 58
DAG | 102 | N = 58
VAR (contemporaneous) 102 | N = 5
VAR (temporal) 102 | N = 58
Triggers
Memory
Negative appraisals
Current threat
Strategies


Clinician`s formulation 103
Patient`s perceived network 103
MGM | 103 | N = 65
DAG | 103 | N = 65
VAR (contemporaneous) 103 | N = 6
VAR (temporal) 103 | N = 65
Triggers
Memory
Negative appraisals
Current threat
Strategies


Clinician`s formulation 104
Patient`s perceived network 104
MGM | 104 | N = 61
DAG | 104 | N = 61
VAR (contemporaneous) 104 | N = 6
VAR (temporal) 104 | N = 61
Triggers
Memory
Negative appraisals
Current threat
Strategies

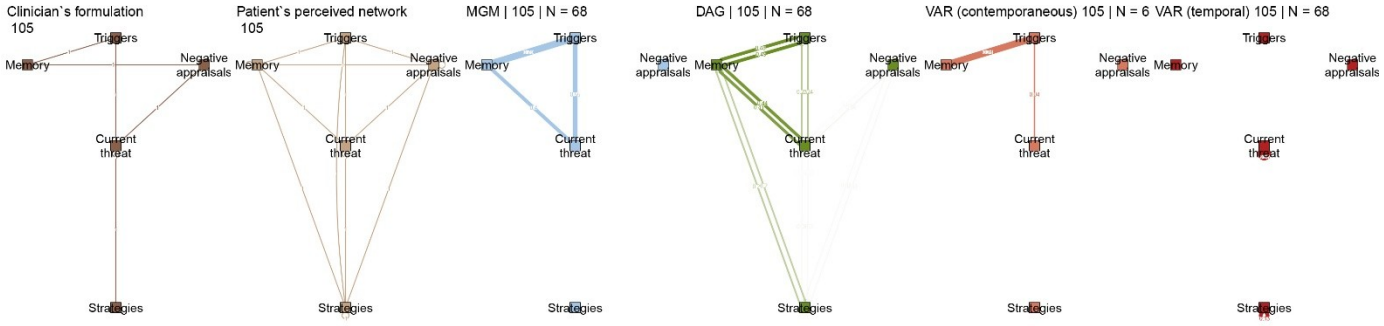

Clinician`s formulation 105
Patient`s perceived network 105
MGM | 105 | N = 68
DAG | 105 | N = 68
VAR (contemporaneous) 105 | N = 6
VAR (temporal) 105 | N = 68
Triggers
Memory
Negative appraisals
Current threat
Strategies

Figure 1. Individual networks, clinician formulations and patient perceived networks

Comparison I – algorithm performance. VAR outperformed MGM in reflecting formulations and perceived associations, under lenient tuning, showing higher overall accuracy (B = 0.07, $p = .027$), sensitivity (B = 0.16, $p < .001$), and F1 score (B = 0.13, $p < .001$), with marginally better specificity (B = 0.02, $p = .05$). However, as tuning became stricter, VAR's performance dropped significantly in accuracy (B = -0.14, $p = .019$), sensitivity (B = -0.31, $p < .001$), and F1 score (B = -0.29, $p < .001$). Precision remained stable across models and tuning levels.

Comparison II – reference matrices. MGM aligned more closely with patient-constructed variable association matrices than with clinician formulation matrices. Specifically, MGM achieved higher precision (B = 0.19, $p = .016$), specificity (B = 0.24, $p = .027$), and F1 score (B = 0.06, $p = .019$) with patient ratings as the reference, compared to when clinicians' formulations served as reference (or the "ground truth"). VAR's performance did not differ significantly by reference matrix.

Overall, MGM captured patients' subjective experiences more precisely, whereas VAR showed initial gains in sensitivity and accuracy but lacked robustness under more stringent tuning parameters.

2. Congruence with the Ehlers and Clark Cognitive Model

The degree of similarity between individual participant network estimations was arguably sufficient to support pooling these to form group level networks. Figure 3 presents the group-level MGM, VAR, and DAG networks. The following sections evaluate whether empirical patterns align with core predictions from the cognitive model.

Table 3. Classification (edge-detection) metrics across tuning parameters for the MGM and

Reference / Measure	Reference: Clinician		Reference: Patient	
	MGM	VAR	MGM	VAR
Accuracy	0.41 (0.17)	0.45 (0.17)	0.41 (0.11)	0.40 (0.14)
F1 Score	0.42 (0.16)	0.45 (0.18)	0.46 (0.09)	0.44 (0.17)
Precision	0.64 (0.22)	0.69 (0.26)	0.86 (0.18)	0.80 (0.23)
Sensitivity	0.31 (0.16)	0.35 (0.20)	0.34 (0.12)	0.34 (0.18)
Specificity	0.61 (0.35)	0.61 (0.38)	0.81 (0.21)	0.71 (0.28)

VAR networks, with reference to the clinician formulations and patient estimations.

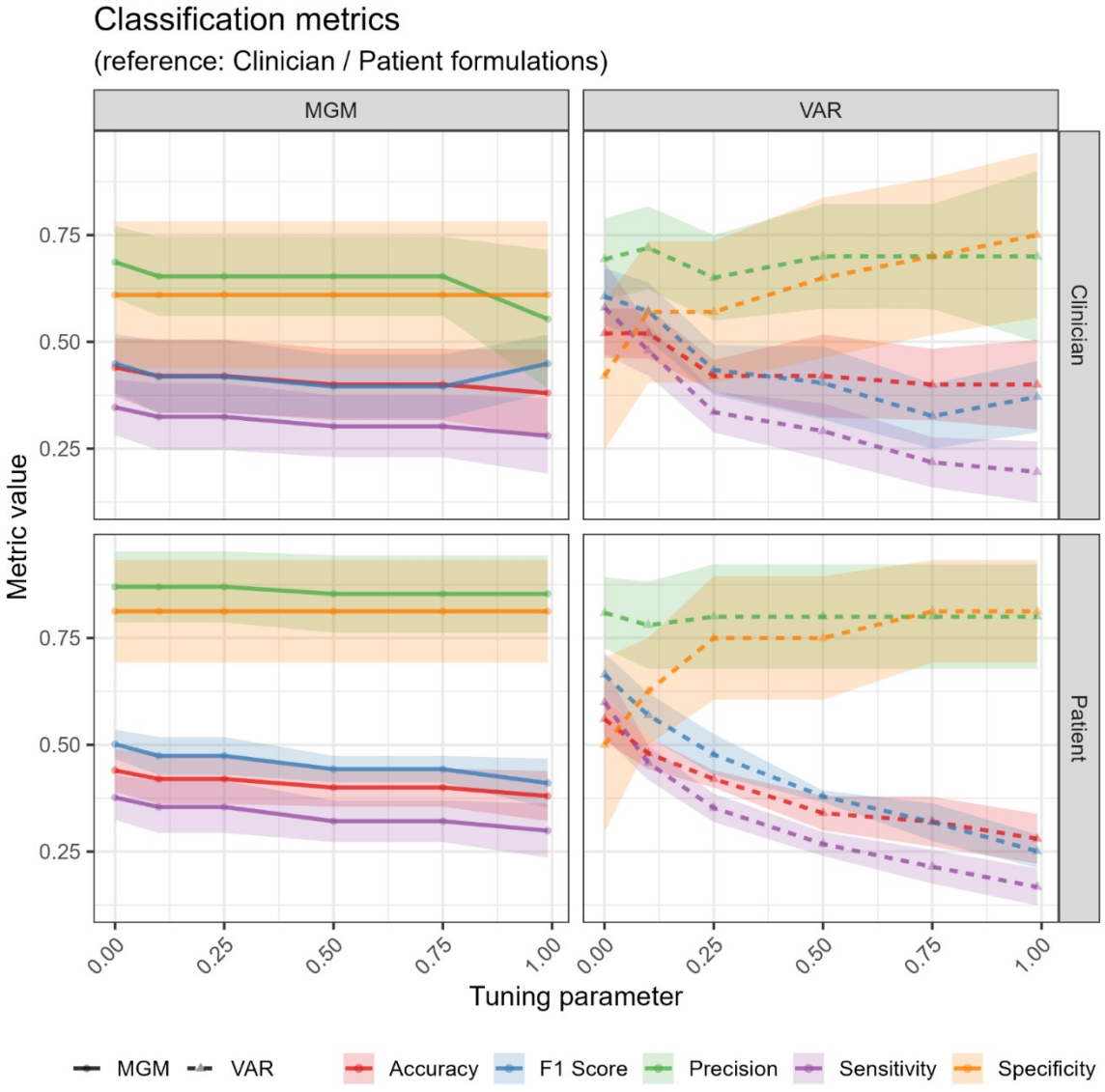


Figure 2. Classification (edge detection) metrics over tuning parameters.

2.1.

Contrary to the model's predictions, none of the networks included robust reciprocal links between trauma memory and negative appraisals. The only exception was a negligible temporal effect in the multilevel VAR network, where trauma memory slightly preceded negative appraisals (edge weight = 0.013). At the individual level, just 16.7%[3] of trauma memory–appraisal edges, across participants and network algorithms, were non-zero, with small and inconsistent effects (M±SD=0.02±0.04, Md=0, range=0.00–0.14).

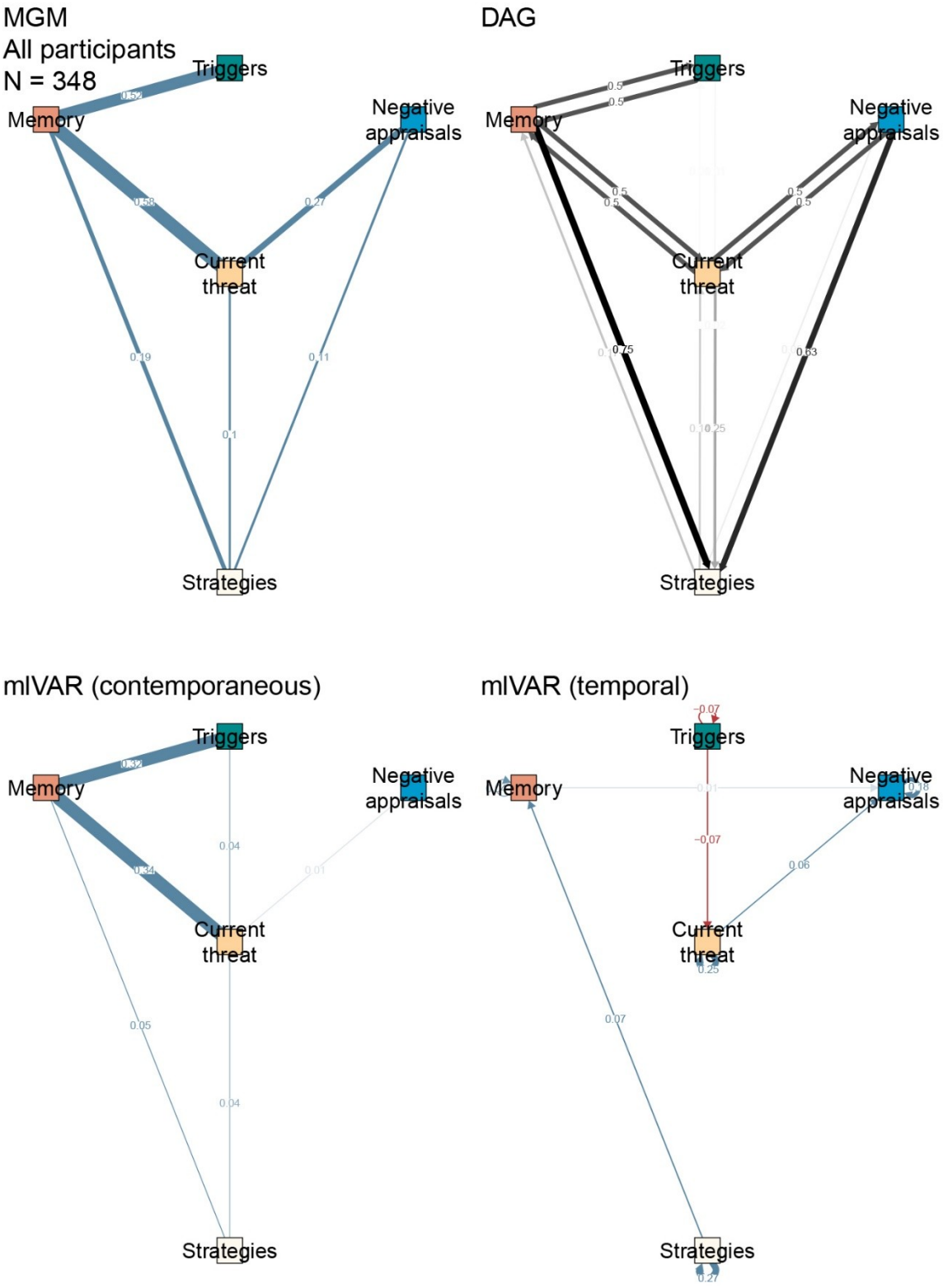
2.2.

Negative Appraisals give rise to Sense of Current Threat

The group-level MGM network showed a strong link from negative appraisals to current threat (edge weight =0.27), the third strongest edge overall. However, the bootstrapped DAG model assigned equal probability (0.5) to both directions. VAR networks showed a weak temporal effect and a negligible contemporaneous association (weight=0.011). Notably, the directionality in VAR ran counter to theoretical expectations, with current threat predicting appraisals over time (w=0.061) but no reciprocal edge from appraisals to current threat. At the individual level, 45% of appraisal–threat edges were non-zero across participants and algorithms, but effects were modest and variable (M=0.10±0.15; median=0.00; range=0.00-0.48).

[3]

Figure 3. MGM, ml-VAR, and DAG networks, illustrating the relationships between variables in the Cognitive Model.



2.3. Trauma Memory gives rise to Sense of Current Threat

Trauma memory and current threat were strongly linked in the group-level MGM (weight=0.58) and contemporaneous VAR networks (weight=0.34), representing the strongest edge in both models. DAG showed a weak directional preference from memory to threat, with 50.4% of bootstraps supporting that direction of flow. No temporal VAR effects were detected in either direction. At the individual level, 55% of trauma memory–threat edges across participants and network algorithms were non-zero, with relatively large and variable effects (M=0.19±0.24; Median=0.04; range=0.00-0.70).

2.4. Matching Triggers evoke Re-experiencing

Triggers and trauma memory were strongly linked in both group-level MGM (w=0.52) and contemporaneous VAR networks (w=0.32), ranking as the second strongest edge in each. DAG showed no directional preference, assigning equal probability to both directions. No temporal VAR effects emerged in either direction. At the individual level, 75% of trigger-memory edges across participants and algorithms were non-zero, with moderate to large effects (M=0.39±0.33; Median=0.45; range=0.00-1.03).

2.5. Perceived Threat Drives Coping Strategies, and Vice Versa

Current threat and coping strategies were modestly linked in both the group-level MGM (w=0.10) and contemporaneous VAR networks (w=0.035). DAG analyses suggested a directional preference from threat to coping (63.7% of non-zero

bootstraps), though this remained weak, and no temporal effects were observed. At the individual level, 45% of threat–coping edges were non-zero, with small and variable effects (M=0.07±0.11; median =0.00; range=0.00–0.30).

2.6. Strategies Prevent Change in Negative Appraisals and Trauma Memory

To test whether coping strategies inhibit change in appraisals and trauma memory, we combined Granger causality tests, using the *plm* package (Croissant & Millo, 2008; Diks & Panchenko, 2006), with descriptive correlation analyses to assess whether greater strategy use was associated with lower variability (i.e., less change) in negative appraisals and trauma memory, and associates with weaker symptom reduction. For each participant, we calculated variability in appraisals/memory (SD over the EMA period) and symptom reduction (beta coefficient from regressing appraisals/memory on time--negative values reflect improvement). These outcomes were then correlated with average strategy use across EMA, inversed so that negative correlations were consistent with strategies being associated with reduced variability and inversely related to symptom reduction[4].

Granger tests (lag-1 and, if relevant, lag-2) assessed whether past strategy use predicted future values of appraisals or memory beyond self-prediction, using FDR correction for multiple comparisons. At the group level, strategies predicted future negative appraisals (Z=2.67, p=.007), but not trauma memory (Z=-0.54, p=.58). Appraisals also predicted strategies (Z=2.20, p=.028), suggesting bi-directionality. No lag-2 effects were significant. In individual analyses, only one participant (101)

[4] As the correlation analyses used values collapsed across the EMA period, they included only five observations. Thus, we report only descriptive values and did not test for statistical significance.

showed significant strategy→appraisal prediction post-FDR, with another (105) showing a pre-correction effect (p=.027). No individuals showed effects for trauma memory. Overall, these results provide limited support that strategies predict negative appraisals over a 1-lag span (3 hours).

Descriptively, strategy use correlated with lower variability in appraisals (r=-0.51) and trauma memory (r=-0.39), and with weaker symptom reduction for both (appraisals: r=-0.65; memory: r=-0.32), suggesting that participants who use strategies frequently are prone to more persistent appraisals and trauma memories. This pattern suggests that frequent strategy use may dampen emotional flexibility and hinder symptom improvement (Figure 4).

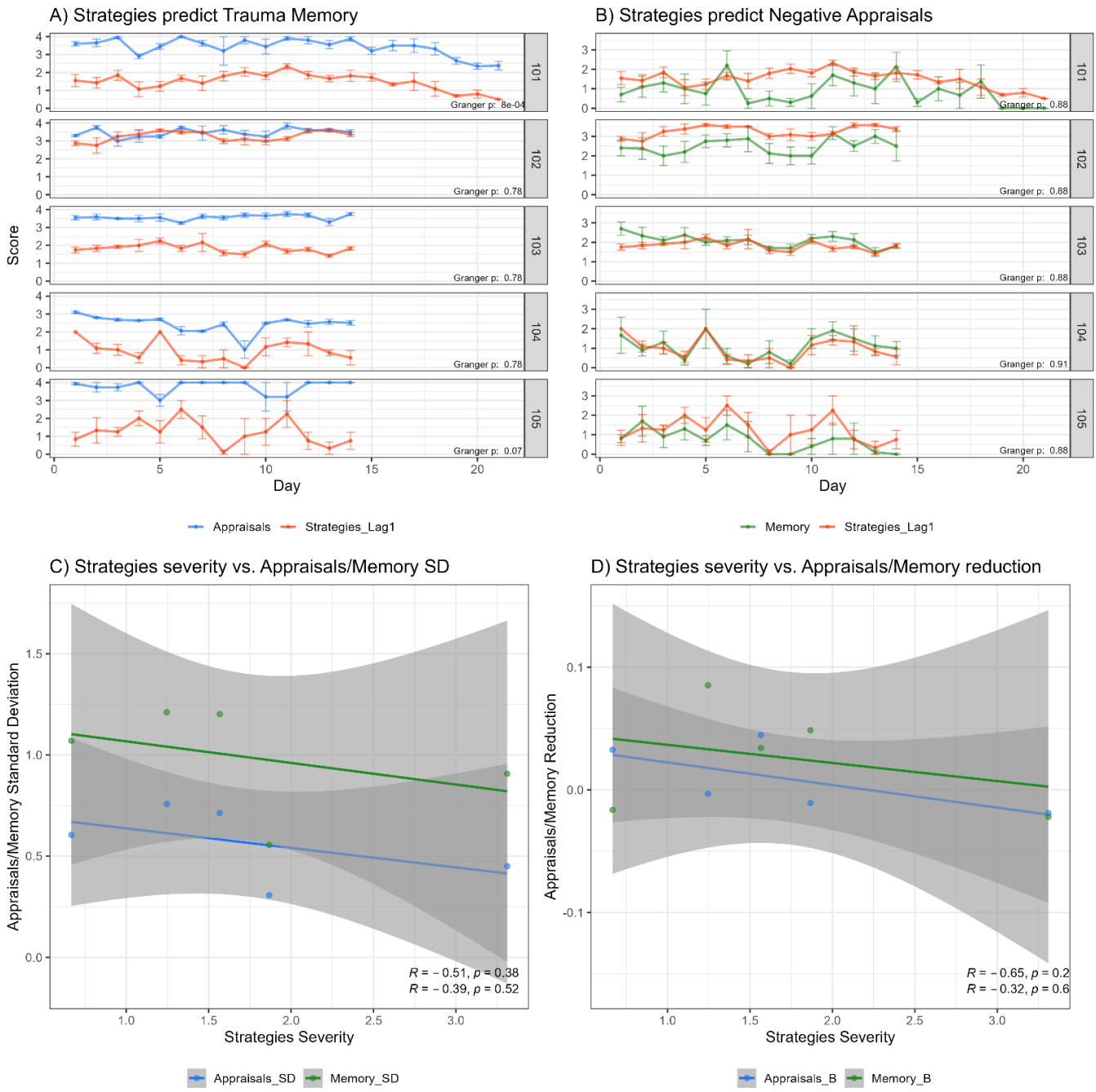


Figure 4. A time series of (lagged-1) strategies with Negative Appraisals [A] and Trauma Memory [B]. Values at the bottom-right corner are p-values from the Granger Causality Test. Lower panels show correlations between use of Maladaptive Strategies with Appraisals/Memory standard deviation [C] and (reversed) symptom reduction [D]. Negative correlation coefficients imply that using strategies reduces variance and hinders symptom reduction.

3.1.

Network Centrality

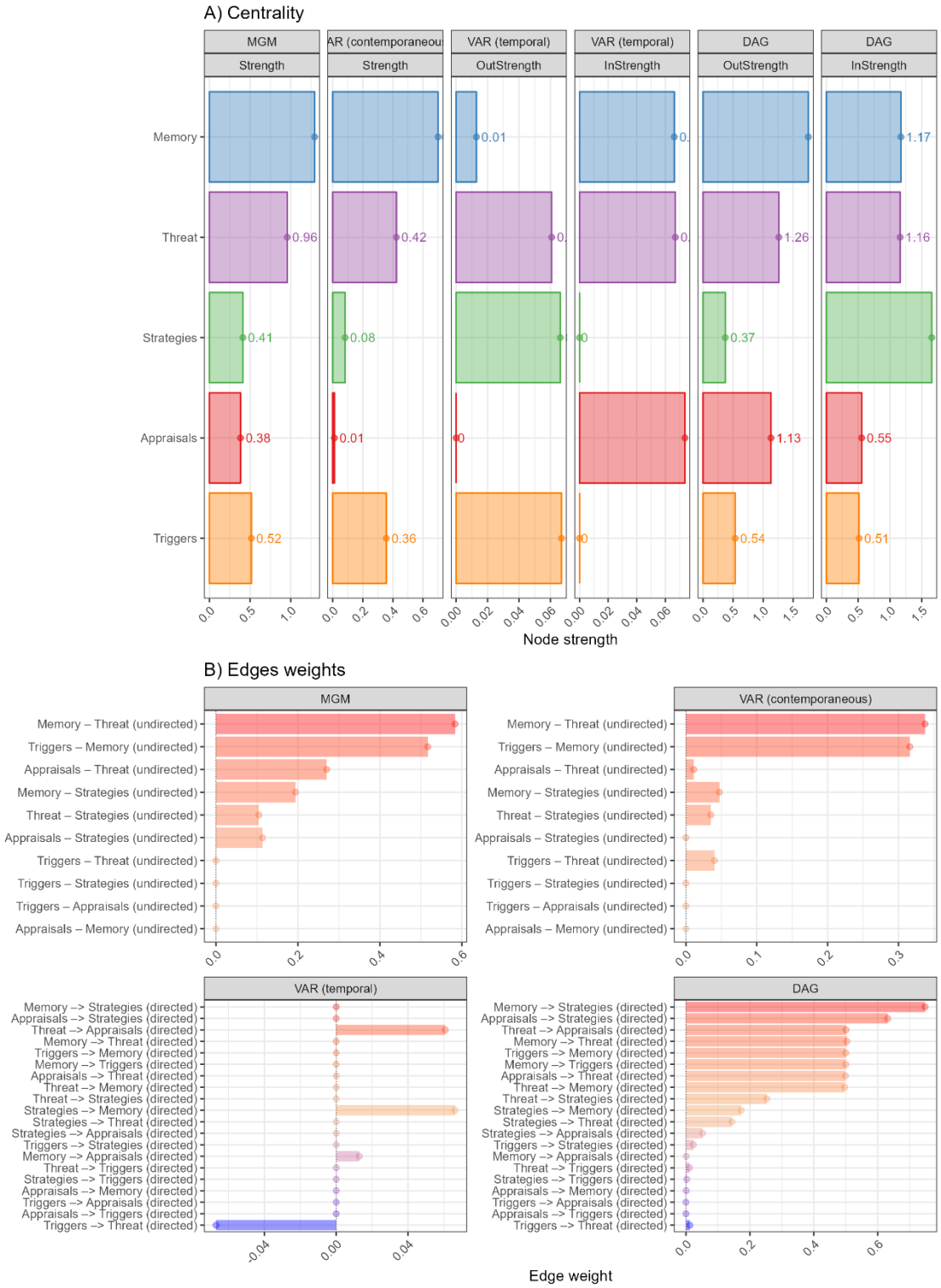
Figure 5A shows strength centrality across models. Trauma memory emerged as the most central node in MGM (1.29), DAG (out=1.75, in=1.17) and contemporaneous VAR networks (0.70), indicating strong momentary and short-term influence. However, in the temporal model, its influence was largely passive (in-strength=0.066; out-strength=0.013), suggesting it is shaped by, but does not shape, other symptoms over time. Sense of threat was also highly central (MGM=0.96; DAG out=1.26, in=1.16; VAR=0.42), with weak but relatively prominent temporal effects (out=0.061; in=0.067), implying a key role in symptom persistence. Triggers had the highest temporal out-strength (VAR=0.067), while strategies showed strong reactivity (DAG in=1.66) and temporal influence (VAR out=0.066). Negative appraisals were not major drivers but were most impacted by prior states (VAR in=0.074), consistent with their role as an outcome of internal processes.

3.2.

Edge Weights

Figure 3B shows that the strongest edges in the MGM and contemporaneous VAR networks linked memory with threat and triggers. DAG analyses showed no directional preference for these associations, assigning equal probabilities to both directions. When directionality did emerge in DAG, the most likely influences were from memory to strategies and appraisals to strategies. In the temporal VAR model, the strongest positive effects were from strategies to memory and threat to appraisals, while triggers negatively predicted future threat, suggesting a short-term dampening effect.

Figure 5. Strength centrality and edge weights across networks



Discussion

The network approach to psychopathology allows researchers to use real-life data to examine relationships between symptoms hypothesized to maintain mental health difficulties. In this study, we applied this approach to PTSD, focusing on cognitive and affective symptoms, which are considered central to symptom maintenance in the cognitive model of PTSD. We also compared the understanding of a participant's PTSD experience gained from network analysis with that obtained through clinical formulation. We hypothesized that individual networks would resemble therapist-derived formulations in terms of complexity and the relationships between problem components, and that network-derived links would align with clinical theory, specifically assuming a lead-lag relationship between cognitive and affective variables.

Contrary to this, network-derived formulations did not resemble therapist formulations in complexity or structure. Network accuracy averaged around 0.40 compared to both clinician and patient formulations, consistent with prior findings of a gap between data-driven models and clinical conceptualizations (Burger et al., 2022; Frumkin et al., 2021). While precision and specificity were relatively high, especially in patients' own estimations, sensitivity was low, suggesting network models often miss core formulation links. Several constraints limited support for our first hypothesis: network-derived formulations were more constrained than clinician-derived ones, both in number and type of variables. To ensure construct validity, we used standardized measures and required at least three variables per component (e.g., three cognitive items). Combined with limits on total variable count to reduce respondent burden, this reduced the idiographic and explanatory richness of network-based formulations.

This analysis assumed clinician and patient formulations reflect the true structure, but that assumption is debatable. Patients are not trained to formulate their difficulties, and clinicians may also fall short, as formulation is a complex skill requiring years of experience (Hua et al., 2025). Three of our five therapists had under 10 years of experience, with one just a year post-qualification, so caution is warranted in treating their formulations as ground truth. Similarly, patients are not expected to be adept at formulating their own difficulties. Indeed, most patients produced nearly fully saturated “all-connected” networks, suggesting difficulty in distinguishing distinct causal pathways. It is possible that repeating this exercise later in treatment, once both therapist and patient have developed a more refined understanding of the symptom interrelations, could yield a formulation closer to the true structure. In addition, future studies could explore whether networks can recover true structure via simulations and parameter recovery. Similarly, clinician accuracy might be tested through supervised formulation or conversation-based simulations, where a therapist attempts to reconstruct a known network structure from a simulated intake. While LLMs are increasingly used in mental health research, their role has so far focused on diagnosis and treatment modules (Hua et al., 2025). A promising next step is using LLMs to support formulation training, especially for early-career clinicians.

Our second hypothesis, regarding a lead-lag relationship between cognitive and affective variables, was not supported. While the group-level directed network included four cognitive-affective links, their effects were small, and idiographic networks showed only self-reinforcing loops rather than associations across domains. These findings suggest that directed networks were insufficient to support the hypothesized relationships. Contemporaneous networks and perceived

associations reported by participants and clinicians indicated stronger cognitive-affective connections than those identified in the directed networks. This discrepancy likely reflects the limitations of the estimated networks rather than the true absence of such links. The directed networks captured predictive relationships within a narrow three-hour window, and contemporaneous networks focused on symptom co-occurrence in even shorter intervals. As a result, relationships operating on longer or more variable timescales may have been missed. Continuous-time models (Abplanalp et al., 2024) may be better suited to detect these effects. More generally, it may ultimately be concluded that the role of formulation in hypothesising complex multi-component interactions that operate over longer periods of time cannot be expected to be captured by network model largely derived from bilateral associations. Alternatively, the intervals used for data collection may need to be derived from the formulation itself, reflecting the hypothesised temporal lags between components. Yet this raises, once again, the challenge of translating case formulations into statistical network models.

There was also considerable variability in our results. Moderate correlations between networks existed despite non-significant Network Comparison Test findings. While we aimed to balance personalization and generalizability by using idiosyncratic items mapped to Ehlers and Clark's components, uncontrolled factors such as therapist experience or patient's reflective capability may have shaped item selection. With three out of five therapists having under ten years of experience, individual differences likely influenced formulation quality. Moreover, our fixed operationalization of model elements, focus on emotional-cognitive mechanisms over symptoms, and tailored EMA prompts may have affected network structure.

One way to reduce discrepancies between the two approaches is to prioritize idiographic over standardized variables, as clinicians do. However, this requires caution to avoid selecting variables that simply reflect the same latent construct, which would create misleading associations (Birkeland et al., 2020). Involving both clinicians and participants in variable selection while relying on established theoretical frameworks may help improve ecological validity (Burger et al., 2022; Burger et al., 2021; Epskamp et al., 2018; Piccirillo & Rodebaugh, 2019). Burger et al., 2021; Epskamp et al., 2018; Piccirillo & Rodebaugh, 2019). Still, some differences may remain inherent: network-based formulations will likely remain more focused due to limits on data collection burden, and network methods struggle to capture stable or slowly unfolding processes (Robinaugh et al., 2020). Perceived Causal Networks (PCN) and longitudinal PCN approaches (Burger et al., 2024; Klintwall et al., 2021, 2023; Vogel et al., 2024) address some of these gaps by relying on participants' insights, but still face constraints in flexibility and scalability.

Despite limitations, network-derived formulations offered incremental validity over clinical ones by focusing exclusively on empirically estimated inter-variable relations rather than combining idiographic and theory-driven links. This feature may support fine-tuned symptom interaction mapping in clinical work by drawing attention to relationships between elements that have not been recognised, by either clinician or patient, in terms of their strength or association. Prior studies suggest network analysis can complement clinical formulation, for example, by informing individualized interventions in depression and social anxiety (David et al., 2018). Our findings support this view, even as they align with research that identified structural differences between clinician-rated and empirically estimated trauma symptom networks (Schumacher et al., 2021). In our case, discrepancies seemed to

reflect methodological trade-offs rather than poor clinical validity of the network model. Overall, our findings echo the argument that formalizing case conceptualization through techniques like network analysis may help bridge the gap between research and practice and promote more rigorous clinical reasoning (Burger et al., 2020; Burger et al., 2021).

The failure to find robust lead-lag relationships in temporal networks—only four links between cognitive and affective variables—likely reflects suboptimal variable selection and time intervals. As Burger et al. (2020) noted, empirical data must map precisely onto theoretical constructs to enable meaningful comparison. Standardized variables may have lacked relevance to individual participants, limiting our ability to capture the key symptom interactions proposed in the cognitive model of PTSD (Ehlers & Clark, 2000). Other explanations are also possible. For example, it could be that the primacy of cognitions in cognitive models is overemphasised, and rather, that people reason with emotion as well (e.g, Arntz, Rauner, & van den Hout, 1995). Also, the lag associations found in previous studies (Kleim et al., 2012, Wiedemann, et al., 2023) tended to be over weeks rather than hours and were examining relationships between a broad range of cognitive-behavioural variables and symptom measures, rather than the much more specific and idiosyncratic variables and narrower affective components (sense of threat) measured here. Those studies also potentially were limited because appraisals and cognitive/behavioural coping strategies are themselves PTSD symptoms in DSM 5 and so are prone to autocorrelation. This study potentially builds on this previous research by taking a much finer grained, specific and idiosyncratic approach to exploring the lag relationships between variables, and in doing so finds that they are potentially more complex, reciprocal, or unstable than the CBT theory suggests

Our study aimed to build on earlier work highlighting how network analysis can illuminate PTSD symptom dynamics. A systematic review (Birkeland et al., 2020) emphasized the centrality of recurrent trauma thoughts and negative emotion. Greene et al. (2018) found negative beliefs predicted avoidance, which in turn predicted emotions. Their 2020 follow-up found that PTSD clusters predicted emotions, but not vice versa. However, their NACM cluster grouped thoughts and feelings together. In Greene's networks, NACM predicted and was predicted by other affective states. Unlike Greene's peri-traumatic sample, our participants had confirmed PTSD, making cognitive-emotional links more clinically relevant. In our networks, trauma memory emerged as a central node, especially in relation to threat and triggers, reinforcing the view of PTSD as a disorder of memory (Brewin, 2011; van Marle, 2015).

Finally, our findings emphasize that variable selection and time lag are both critical for using network analysis to test cognitive theory. The discrepancy between perceived and observed cognitive-affective relationships proposes neither were sufficiently matched to the dynamics of the disorder. Selecting appropriate time lags is difficult given uncertainty about the speed of interactions. Prior studies have used intervals ranging from 90 minutes (Wichers et al., 2014) to 12–15 hours (Greene et al., 2018). As no standard exists, future designs should align time intervals with both theoretical predictions and individual presentation.

Limitations

This was a modest sample, justified by the study's specific aim: to operationalise a cognitive theory of PTSD, unlike previous larger network studies. Recruiting more

would have unnecessarily burdened participants with PTSD, so we limited the sample to what was essential for the foundational methodological aims of the research. Although each participant contributed at least 59 data points, findings cannot be widely generalised. Selection bias was also likely, as only individuals diagnosed with PTSD who could tolerate intensive data collection were included. This may have skewed the sample toward those with digital literacy and relatively stable symptoms. Participants also selected variables, which may have led them to avoid distressing symptoms or to choose those they understood well. Despite guidance, moment-to-moment interpretation of items may have varied (Chun, 2016).

Our analysis assumed data stationarity. While data were collected over just two weeks and all participants were mid-treatment, changes did occur. One participant, for example, showed a notable decline in PCL-5 score. We addressed this by detrending variables before VAR analysis, but this remains a limitation. Some data were missing, which we assumed to be missing completely at random, supported by Little's MCAR test. Still, if this assumption was incorrect, results may have been affected. Another limitation was that each participant had a different therapist, introducing variability in formulation quality. Therapist experience has been linked to better, more coherent formulations (Dudley et al., 2015), which may have influenced the clinician-derived formulations used for comparison.

Despite limitations, this was a proof-of-principle study piloting methods to examine dynamic symptom interactions and compare different formulation sources. Data came from patients in real-world clinical settings, enhancing ecological validity. As this was a novel approach, we developed a synthesis method grounded in existing literature, with transparent definitions and clear reporting to support replication.

Future research should build on the strengths and limitations of this work, addressing variability in both symptoms and formulation practices.

Implications

This study's findings have important clinical and scientific implications. Network-derived formulations can offer data-driven insights that complement, but do not replace, traditional clinical formulations. Our results, consistent with prior work (e.g., Kroeze et al., 2017), suggest that networks can highlight unique symptom interactions but remain too sparse to guide treatment decisions on their own (Greene et al., 2020). Instead, they are best used for exploring specific hypotheses developed during clinical formulation (David et al., 2018; Robinaugh et al., 2019) and potentially for highlighting novel or unrecognised functional relationships that can be explored further clinically. This might be especially pertinent in situations where the mechanisms of change suggested by the theory underlying the clinical formulation appear not to be operating as predicted, most notably where treatment is not progressing as planned or appears at some impasse. Networks might also highlight potential areas of complexity in interactions between variables in advance, that could be explored and then accounted for with an adapted treatment plan. This might suggest, for example, moving focus in early sessions to a variable that neither clinician or patient had accounted for as potentially important in the overall maintenance of the disorder, with moderators such as coping strategies, alcohol use, or over-work. Questions also remain about when to use network analysis and how long to collect data, especially since our two-week mid-treatment data window was too small to generalize confidently.

This study's scientific implications touch upon both hypotheses tested. Despite limited agreement between clinicians and data networks, our results confirmed that network analysis is a viable and informative method to investigate these relationships, serving as a proof-of-principle. A key challenge remains in choosing items and measurement intervals, as highlighted by prior research (Epskamp et al., 2018) and our own data.

Conclusion

This study tested two hypotheses: that network- and clinician-derived formulations would show comparable complexity and relationships, and that a lead-lag relationship would exist between cognitive and affective variables in PTSD. Neither was fully supported, though the study confirmed network analysis as a promising method for exploring these questions and highlighted ways to refine future research. Additionally, network analysis provided unique insights not possible through clinical formulation alone by using participants' empirical data to document idiographic symptom links. The study also suggested that clinical input could improve network analyses by helping define which symptoms to include. Lastly, it raised the need for theoretical work on how to capture symptoms that do not co-vary neatly with others, which current network approaches struggle to represent.

Our finding that network analysis could validate hypothesized symptom relationships in clinical contexts warrants further testing. Wichers et al. (2017) recommended randomized controlled trials to evaluate whether network insights improve clinical decision making and outcomes. We support this call. Given the data collection burden, demonstrating that network analysis provides sufficient benefit to patients is crucial for justifying its use in practice.

Contributors

All authors contributed to, reviewed, and approved the final manuscript.

Conceptualization: EB, NH, SEL, TD, AB, GB

Data curation: EB, NH, GB

Formal analysis: NH

Funding acquisition: GB

Investigation: EB, SEL, GB

Methodology: EB, NH, GB

Project administration: EB, GB

Resources: EB, SEL, GB

Software: NH

Supervision: GB

Validation: GB

Visualization: NH

Writing – original draft: EB, NH, GB

Writing – review & editing: SEL, TD, AB

Funding

Nimrod Hertz is supported by the Gates Cambridge Trust (#OPP1144). The funding source had no role in the study design, collection, analysis, or interpretation of data, the writing of the article, or decision to submit the article for publication.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Ethics statement

NHS ethical approval was obtained from South Central Berkshire B Research Ethics Committee of the NHS Health Research Authority.

Acknowledgments

We think the clients and clinicians who contributed their time and effort to this study.

References

Abplanalp, S. J., Reavis, E. A., Le, T. P., & Green, M. F. (2024). Applying continuous-time models to ecological momentary assessments: A practical introduction to the method and demonstration with clinical data. NPP—Digital Psychiatry and Neuroscience, 2(1), 2. https://doi.org/10.1038/s44277-024-00004-x

Altenbuchinger, M., Weihs, A., Quackenbush, J., Grabe, H. J., & Zacharias, H. U. (2020). Gaussian and Mixed Graphical Models as (multi-)omics data analysis tools. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms, 1863(6), 194418. https://doi.org/10.1016/j.bbagrm.2019.194418

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

Arntz, A., Rauner, M., & van den Hout, M. (1995). 'If I feel anxious, there must be danger': Ex-consequentia reasoning in inferring danger in anxiety disorders. *Behaviour Research and Therapy*, *33*(8), 917–925.

Bak, M., Drukker, M., Hasmi, L., & Os, J. van. (2016). An n=1 clinical network analysis of symptoms and treatment in psychosis. *PLOS ONE*, *11*(9), e0162811. https://doi.org/10.1371/journal.pone.0162811

Balaskas, A., Schueller, S. M., Cox, A. L., & Doherty, G. (2021). Ecological momentary interventions for mental health: A scoping review. PLOS ONE, 16(3), e0248152. https://doi.org/10.1371/journal.pone.0248152

Beck ED, Jackson JJ. Detecting Idiographic Personality Change. J Pers Assess. 2022 Jul-Aug;104(4):467-483. doi: 10.1080/00223891.2021.1984246.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B (Methodological), 57(1), 289–300.

Bieling, P. J., & Kuyken, W. (2003). Is Cognitive Case Formulation Science or Science Fiction? Clinical Psychology: Science and Practice, 10(1), 52–69. https://doi.org/10.1093/clipsy.10.1.52

Birkeland, M. S., Greene, T., & Spiller, T. R. (2020). The network approach to posttraumatic stress disorder: A systematic review. *European Journal of Psychotraumatology*, *11*(1), 1700614. https://doi.org/10.1080/20008198.2019.1700614

Birkeland, M. S., & Heir, T. (2017). Making connections: Exploring the centrality of posttraumatic stress symptoms and covariates after a

terrorist attack. *European Journal of Psychotraumatology*, *8*(sup3), 1333387. https://doi.org/10.1080/20008198.2017.1333387

Blevins, C.A., Weathers, F.W., Davis, M.T., Witte, T.K., Domino, J.L. (2015). The posttraumatic stress disorder checklist for DSM-5 (PCL-5): Development and initial psychometric evaluation. *Journal of Traumatic Stress, 28(6)*, 489-498. https://doi.org/10.1002/jts.22059

Boorsboom, D. (2008). Psychometric perspectives on diagnostic systems. *Journal of Clinical Psychology, 64(9)*, 1089-1108. https://doi.org/10.1002/jclp.20503

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

Brewin, C. R. (2011). The nature and significance of memory disturbance in posttraumatic stress disorder. Annual Review of Clinical Psychology, 7, 203–227. https://doi.org/10.1146/annurev-clinpsy-032210-104544

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, 104011. https://doi.org/10.1016/j.brat.2021.104011

Bringmann, L. F., Elmer, T., Epskamp, S., Krause, R. W., Schoch, D., Wichers, M., Wigman, J. T. W., & Snippe, E. (2019). What do centrality measures measure in psychological networks? Journal of Abnormal Psychology, 128(8), 892–903. https://doi.org/10.1037/abn0000446

Bringmann, L. F., Hamaker, E. L., Vigo, D. E., Aubert, A., Borsboom, D., & Tuerlinckx, F. (2017). Changing dynamics: Time-varying autoregressive models using generalized additive modeling. Psychological Methods, 22(3), 409–425. https://doi.org/10.1037/met0000085

Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., Borsboom, D., & Tuerlinckx, F. (2013). A network approach to psychopathology: New insights into clinical longitudinal data. *PLOS ONE*, *8*(4), e60188. https://doi.org/10.1371/journal.pone.0060188

Brown, G. P. (2022). Second Wave Assessment and Case Formulation. In W. O'Donohue & A. Masuda (Eds.), Behavior Therapy: First, Second, and Third Waves (pp. 153–172). Springer International Publishing. https://doi.org/10.1007/978-3-031-11677-3_7

Bryant, R. A., Creamer, M., O'Donnell, M., Forbes, D., McFarlane, A. C., Silove, D., & Hadzi-Pavlovic, D. (2017). Acute and chronic posttraumatic stress symptoms in the emergence of posttraumatic stress disorder: A network analysis. *JAMA Psychiatry*, *74*(2), 135–142. https://doi.org/10.1001/jamapsychiatry.2016.3470

Burger, J., Andikkhash, V., Jäger, N., Anderbro, T., Blanken, T. F., & Klintwall, L. (2024). A novel approach for constructing personalized networks from longitudinal perceived causal relations. Behaviour Research and Therapy, 173, 104456. https://doi.org/10.1016/j.brat.2023.104456

Burger, J., Epskamp, S., Veen, D. C. van der, Dablander, F., Schoevers, R. A., Fried, E. I., & Riese, H. (2021). A clinical PREMISE for personalized models: Towards a formal integration of case formulations and statistical networks. PsyArXiv. https://doi.org/10.31234/osf.io/bdrs7

Burger, J., Ralph-Nearman, C., & Levinson, C. A. (2022). Integrating clinician and patient case conceptualization with momentary assessment data to construct idiographic networks: Moving toward personalized treatment for eating disorders. Behaviour Research and Therapy, 159, 104221. https://doi.org/10.1016/j.brat.2022.104221

Burger, J., van der Veen, D. C., Robinaugh, D. J., Quax, R., Riese, H., Schoevers, R. A., & Epskamp, S. (2020). Bridging the gap between complexity science and clinical practice by formalizing idiographic theories: A computational model of functional analysis. *BMC Medicine*, *18*(1), 99. https://doi.org/10.1186/s12916-020-01558-1

Chun, C.A. (2016). The expression of posttraumatic stress symptoms in daily life: A review of experience sampling methodology and daily diary studies. *Journal of Psychopathology and Behavioural Assessment, 38*(3), 406-420. https://doi.org/10.1007/s10862-016-9540-3

Clohessy, S., & Ehlers, A. (1999). PTSD symptoms, response to intrusive memories and coping in ambulance service workers. *British Journal of Clinical Psychology, 38*(3), 251-65. https://doi.org/10.1348/014466599162836

Croissant, Y., & Millo, G. (2008). Panel data econometrics in R: The plm package. Journal of Statistical Software, 27, 1–43.

David, S.J., Marshall, A., Evanovich, E.K., & Mumma, G.H. (2018). Intraindividual dynamic network analysis – implications for clinical assessment. *Journal of Psychopathology and Behavioral Assessment, 40*, 235.248. https://doi.org/10.1007/s10862-017-9632-8

Davison, G. C. (2018). A Return to Functional Analysis, the Search for Mechanisms of Change, and the Nomothetic-Idiographic Issue in Psychosocial Interventions. Clinical Psychological Science, 2167702618794924. https://doi.org/10.1177/2167702618794924

Diks, C., & Panchenko, V. (2006). A new statistic and practical guidelines for nonparametric Granger causality testing. Journal of Economic Dynamics and Control, 30(9–10), 1647–1669.

Doron-LaMarca, S., Niles, B. L., King, D. W., King, L. A., Pless Kaiser, A., & Lyons, M. J. (2015). Temporal associations among chronic PTSD symptoms in U.S. combat veterans. *Journal of Traumatic Stress*, *28*(5), 410–417. https://doi.org/10.1002/jts.22039

Dudley, R., Ingham, B., Sowerby, K., & Freeston, M. (2015) The utility of case formulation in treatment decision making; the effect of experience and expertise. Journal of Behavior Therapy and Experimental Psychiatry, 48, 66-74. https://doi.org/10.1016/j.jbtep.2015.01.009

Eells, T.D., Kendjelic, E.M., & Lucas, C.P. (1998). What's in a case formulation?: Development and use of a content coding manual. *Journal of Psychotherapy Practice and Research, 7(2)*, 144-153.

Ehlers, A., & Clark, D.M. (2000). A cognitive model of posttraumatic stress disorder. *Behaviour Research and Therapy, 38*, 319-345. https://doi.org/10.1016/S0005-7967(99)00123-0

Epskamp, S. (2017). graphicalVAR: Graphical VAR for experience sampling data. Vyhledáno Na Https://Cran. r-Project. Org/Web/Packages/graphicalVAR/graphicalVAR. Pdf.

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. Behavior Research Methods, 50(1), 195–212. https://doi.org/10.3758/s13428-017-0862-1

Epskamp, S., van Borkulo, C. D., van der Veen, D. C., Servaas, M. N., Isvoranu, A.-M., Riese, H., & Cramer, A. O. J. (2018). Personalized network modelling in psychopathology: The importance of contemporaneous and temporal connections. *Clinical Psychological Science*, *6*(3), 416–427. https://doi.org/10.1177/2167702617744325

Epskamp, S., Waldorp, L. J., Mõttus, R., & Borsboom, D. (2018). The Gaussian Graphical Model in Cross-Sectional and Time-Series Data. Multivariate Behavioral Research, 53(4), 453–480. https://doi.org/10.1080/00273171.2018.1454823

Fisher, A. J. (2015). Toward a dynamic model of psychological assessment: Implications for personalized care. Journal of Consulting and Clinical Psychology, 83(4), 825–836. https://doi.org/10.1037/ccp0000026

Foa, E. B., Ehlers, A., Clark, D.M., Tolin, D.F., & Orsillo, S.M. (1999). The posttraumatic cognitions inventory (PTCI): Development and validation. *Psychological Assessment 11(3)*, 303-314.

Fried, E. I., van Borkulo, C. D., Cramer, A. O. J., 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. https://doi.org/10.1007/s00127-016-1319-z

Frumkin, M. R., Piccirillo, M. L., Beck, E. D., Grossman, J. T., & Rodebaugh, T. L. (2021). Feasibility and utility of idiographic models in the clinic: A pilot study. Psychotherapy Research: Journal of the Society for Psychotherapy Research, 31(4), 520–534. https://doi.org/10.1080/10503307.2020.1805133

Granger, C.W. (1969). Investigating causal relations by econometric models and cross-spectral models. *Econometrica: journal of the Econometric Society*, 424-438.

Greene, T., Gelkopf, M., Epskamp, S., & Fried, E. (2018). Dynamic networks of PTSD symptoms during conflict. *Psychological Medicine*, *48*(14), 2409–2417. https://doi.org/10.1017/S0033291718000351

Greene, T., Gelkopf, M., Fried, E. I., Robinaugh, D. J., & Lapid Pickman, L. (2020). Dynamic network analysis of negative emotions and DSM-5 posttraumatic stress disorder symptom clusters during conflict. *Journal of Traumatic Stress*, *33*(1), 72–83. https://doi.org/10.1002/jts.22433

Hallam, R.S. (2013). *Individual Case Formulation*. Elsevier.

Hallam, R., Brown, G., Turner, L., Blomfield, E., El Leithy, S., & Mendes, I. (2024). Development of a training programme in individual case formulation skills and a scale for assessing its effectiveness. Behavioural and Cognitive Psychotherapy, 52(2), 200–203. https://doi.org/10.1017/S1352465823000577

Haslbeck, J. M., & Waldorp, L. J. (2020). mgm: Estimating time-varying mixed graphical models in high-dimensional data. Journal of Statistical Software, 93, 1–46.

Health & Care Professions Council. (2015). *The standards of proficiency for practitioner psychologists.*

https://www.hcpc-uk.org/standards/standards-of-proficiency/practitioner-psychologists

Hernan, M. A., & Robins, J. M. (2024). Causal Inference: What If. CRC Press.

Hua, Y., Na, H., Li, Z., Liu, F., Fang, X., Clifton, D., & Torous, J. (2025). A scoping review of large language models for generative tasks in mental health care. Npj Digital Medicine, 8(1), 230. https://doi.org/10.1038/s41746-025-01611-4

Jacobson, N. C., & Chung, Y. J. (2020). Passive Sensing of Prediction of Moment-To-Moment Depressed Mood among Undergraduates with Clinical Levels of Depression Sample Using Smartphones. Sensors, 20(12), Article 12. https://doi.org/10.3390/s20123572

King, L., McKenzie-McHarg, K., Horsch, A. (2017). Teaching a cognitive model to predict posttraumatic stress disorder following childbirth, *BMC Pregnancy and Childbirth, 17*(32), 1-12. https://doi.org/10.1186/s12884-016-1194-3

Kleim, B., Ehlers, A., & Glucksman, E. (2012). Investigating cognitive pathways to psychopathology: Predicting depression and posttraumatic stress disorder from early responses after assault. Psychological Trauma: Theory, Research, Practice and Policy, 4(5), 527-537. https://doi.org/10.1037/a0027006

Klintwall, L., Bellander, M., & Cervin, M. (2021). Perceived Causal Problem Networks: Reliability, Central Problems, and Clinical Utility for Depression. Assessment. https://doi.org/10.1177/10731911211039281

Kramer, I., Simons, C. J. P., Wigman, J. T. W., Collip, D., Jacobs, N., Derom, C., Thiery, E., van Os, J., MyinGermeys, I., & Wichers, M. (2014). Time-lagged moment-to-moment interplay between negative affect and paranoia: New insights in the affective pathway to psychosis. *Schizophrenia Bulletin*, *40*(2), 278–286. https://doi.org/10.1093/schbul/sbs194

Kroeze, R., Veen, D. C. van der, Servaas, M. N., Bastiaansen, J. A., Voshaar, R. C. O., Borsboom, D., Ruhe, H. G., Schoevers, R. A., & Riese, H. (2017). Personalized feedback on symptom dynamics of psychopathology: A proof-of-principle study. *Journal for Person-Oriented Research*, *3*(1), 1–10. https://doi.org/10.17505/jpor.2017.01

Lauritzen, S.L. (1996). *Graphical models (Vol. 17).* Clarendon Press.

Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. Journal of the American Statistical

Association, 83(404), 1198–1202.

Mansueto, A. C., Wiers, R. W., van Weert, J., Schouten, B. C., & Epskamp, S. (2022). Investigating the feasibility of idiographic network models. Psychological Methods.

McNally, R. J., Heeren, A., & Robinaugh, D. J. (2017). A Bayesian network analysis of posttraumatic stress disorder symptoms in adults reporting childhood sexual abuse. *European Journal of Psychotraumatology*, *8*(sup3), 1341276. https://doi.org/10.1080/20008198.2017.1341276

McNally, R. J., Robinaugh, D. J., Wu, G. W. Y., Wang, L., Deserno, M. K., & Borsboom, D. (2015). Mental disorders as causal systems: A network approach to posttraumatic stress disorder. *Clinical Psychological Science*, *3*(6), 836–849. https://doi.org/10.1177/2167702614553230

Mumma, G. H., Marshall, A. J., & Mauer, C. (2018). Person-specific validation and testing of functional relations in cognitive-behavioural case formulation: Guidelines and options. Clinical Psychology & Psychotherapy, 25(5), 672–691. psyh. https://doi.org/10.1002/cpp.2298

Murray, J., Ehlers, A., & Mayou, R.A. (2002). Dissociation and post-traumatic stress disorder: Two prospective studies of road traffic accident survivors. *British Journal of Psychiatry, 180*, 363-368.

Pearl, J. (2009). Causal inference in statistics: An overview. Statistics Surveys, 3(none), 96–146. https://doi.org/10.1214/09-SS057

Pearl, J., & Mackenzie, D. (2018). The Book of Why: The New Science of Cause and Effect. Allen Lane.

Piccirillo, M. L., & Rodebaugh, T. L. (2019). Foundations of idiographic methods in psychology and applications for psychotherapy. *Clinical Psychology Review*, *71*, 90–100. https://doi.org/10.1016/j.cpr.2019.01.002

Reeves, J.W., & Fisher, A.J. (2020). An examination of idiographic networks of posttraumatic stress disorder symptoms. *Journal of Traumatic Stress, 33*, 84-95. https://doi.org/10.1002/jts.22491

Regambal, M.J., & Alden, L.E. (2012). The contribution of threat probability estimates to reexperiencing symptoms: A prospective analog study. *Journal of Behavior Therapy and Experimental Psychiatry, 43(3)*, 947-951. https://doi.org/10.1016/j.jbtep.2012.02.003

Robinaugh, D., Haslbeck, J., Waldorp l., Kossakowski, J., Fried, E., Millner, A., McNally, R.J., van Nes, E.H., Scheffer, M., Kendler, K.S., & Borsboom, D. (2019). Advancing the network theory of mental disorders: A

computational model of panic disorder. *Preprints.* Advance online publication. www.psyarxiv.com

Robinaugh, D.J., Hoekstra, R.H.A., Toer, 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/10.1017.S0033291719003404

Roefs, A., Fried, E. I., Kindt, M., Martijn, C., Elzinga, B., Evers, A. W. M., Wiers, R. W., Borsboom, D., & Jansen, A. (2022). A new science of mental disorders: Using personalised, transdiagnostic, dynamical systems to understand, model, diagnose and treat psychopathology. Behaviour Research and Therapy, 153, 104096. https://doi.org/10.1016/j.brat.2022.104096

Santangelo, P.S., Ebner-Priemer, U.W., and Trull, T.J. (2013). Experience sampling methods in clinical psychology. *The Oxford Handbook of Research Strategies for Clinical Psychology*, 188–210.

Scholten, S., Lischetzke, T., & Glombiewski, J. A. (2022). Integrating theory-based and data-driven methods to case conceptualization: A functional analysis approach with ecological momentary assessment. Psychotherapy Research, 32(1), 52–64. https://doi.org/10.1080/10503307.2021.1916639

Schreuder, M. J., Wigman, J. T. W., Groen, R. N., Weinans, E., Wichers, M., & Hartman, C. A. (2022). Anticipating the direction of symptom progression using critical slowing down: A proof-of-concept study. BMC Psychiatry, 22(1), 49. https://doi.org/10.1186/s12888-022-03686-9

Schumacher, L., Burger, J., Zoellner, F., Zindler, A., Epskamp, S., & Barthel, D. (2021). Using clinical expertise and empirical data in constructing networks of trauma symptoms in refugee youth. *European Journal of Psychotraumatology*, *12*(1), 1920200. https://doi.org/10.1080/20008198.2021.1920200

Schreuder, M. J., Wigman, J. T. W., Groen, R. N., Weinans, E., Wichers, M., & Hartman, C. A. (2022). Anticipating the direction of symptom progression using critical slowing down: A proof-of-concept study. *BMC Psychiatry*, *22*(1), 49. https://doi.org/10.1186/s12888-022-03686-9

Scutari, M. (2010). Learning Bayesian networks with the bnlearn R Package. Journal of Statistical Software, 35(3).

Smit, A., Helmich, M. A., Bringmann, L., Oldehinkel, A., Wichers, M., & Snippe, E. (2023). Critical slowing down in momentary affect as early warning

signal of impending transitions in depression. https://doi.org/10.31234/osf.io/u6vdg

Spiller, T. R., Schick, M., Schnyder, U., Bryant, R. A., Nickerson, A., & Morina, N. (2017). Symptoms of posttraumatic stress disorder in a clinical sample of refugees: A network analysis. *European Journal of Psychotraumatology*, *8*(sup3), 1318032. https://doi.org/10.1080/20008198.2017.1318032

Stone, A.A., & Shiffman, S. (1994) Ecological momentary assessment in behavioral medicine. *Annals of behavioral medicine, 16*, 199-202. https://doi.org/10.1093/abm/16.3.199

U.S. Department of Veteran Affairs. (2021, November 10). *PTSD Checklist for DSM-5 (PCL-5).* https://www.ptsd.va.gov/professional/assessment/adult-sr/ptsd-checklist.asp

van Borkulo, C.M. (2022). Type Package Title Statistical Comparison of Two Networks Based on Three Invariance Measures Version 2.2.1. https://doi.org/10.13140/RG.2.2.29455.38569

Van der Krieke, L., Emerencia, A.C., Bos, E.H., Rosmalen, J.G., Riese, H., Aiello, M., Sytema, S., & de Jonge, P. (2015). Ecological momentary assessments and automated time series analysis to promote tailored health care: A proof-of-principle study. *JMIR research proposals, 4(3)*, e100. https://doi.org/10.2196/resprot.4000

van Marle, H. (2015). PTSD as a memory disorder. European Journal of Psychotraumatology, 6, 10.3402/ejpt.v6.27633. https://doi.org/10.3402/ejpt.v6.27633

Vandenbroucke, J. P., Broadbent, A., & Pearce, N. (2016). Causality and causal inference in epidemiology: The need for a pluralistic approach. International Journal of Epidemiology, 45(6), 1776–1786. https://doi.org/10.1093/ije/dyv341

Vogel, F., Blanken, T., Burger, J., Reichert, J., Scholten, S., & Klintwall, L. (2024). How Perceived Causal Networks (PECAN) can complement case conceptualization, diagnostic classification and data-based networks: An introduction to a method for constructing personalized networks.

Vrijen, C., Hartman, C.A., van Roekel, E., de Jonge, P., & Oldehinkel, A.J. (2018). Spread the joy: How high and low bias for happy facial emotions translate into different daily life affect dynamics. *Complexity, Article ID 2674523*, 1-15. https://doi.org/10.1155/2018/2674523

Wang, L., Zhang, Q., Maxwell, S. E., & Bergeman, C. S. (2019). On Standardizing Within-Person Effects: Potential Problems of Global Standardization. Multivariate Behavioral Research, 54(3), 382–403. https://doi.org/10.1080/00273171.2018.1532280

Weathers, F.W., Litz, B.T., Keane, T.M., Palmieri, P.A., Marx, B.P., & Schnurr, P.P. (2013). The PTSD Checklist for DSM-5 (PCL-5). Scale available from the National Center for PTSD at www.ptsd.va.gov

Wells, S.Y., Morland, L.A., Torres, E.M., Kloezeman, K., Mackintosh, M.-A., & Aarons, G.A. (2019). The development of a brief version of the posttraumatic cognitions inventory (PTCI-9). *Assessment, 26(9)*, 193-208. https://doi.org/10.1177/1073191116685401

Wichers, M. (2014). The dynamic nature of depression: A new micro-level perspective of mental disorder that meets current challenges. *Psychological Medicine*, *44*(7), 1349–1360. https://doi.org/10.1017/S0033291713001979

Wichers, M., Wigman, J.T.W., Bringmann, L.F., & de Jonge, P. (2017). Mental disorders as networks: some cautionary reflections on a promising approach. *Social Psychiatry and Psychiatric Epidemiology, 52*(2), 143-145.

Wichers, M., Wigman, J.T.W., & Myin-Germeys, I. (2015). Micro-level affect dynamics in psychopathology viewed from complex dynamical system theory. *Emotion Review, 7(4)*, 362-367. https://doi.org/10.1177/1754073915590623

Wiedemann M, Janecka M, Wild J, Warnock-Parkes E, Stott R, Grey N, Clark DM, Ehlers A. Changes in cognitive processes and coping strategies precede changes in symptoms during cognitive therapy for posttraumatic stress disorder. Behav Res Ther. 2023 Oct;169:104407. doi: 10.1016/j.brat.2023.104407.Wild, B., Eichler, M., Friederich, H.C., Hartmann, M., Zipfel, S., & Herzog, W. (2010). A graphical vector autoregressive modelling approach to the analysis of electronic diary data. *BMC medical research methodology, 10(1)*, 1-13. https://doi.org/10.1186/1471-2288-10-28

Wild, J., Warnock-Parkes, E., Grey, N., Stott, R., Wiedemann, M., Canvin, L., Rankin, H., Shepherd, E., Forkert, A., Clark, D.M., & Ehlers, A. (2016). Internet-delivered cognitive therapy for PTSD: a development pilot series. *European Journal of Psychotraumatology, 7*: 31019. https://doi.org/10.3402.ejpt.v7.31019