The University Of Sheffield.

Evidence of the Idiographic Effect of Cognitive Analytic Therapy and Analysis of its Primary Outcome Measure

Georgia Mangion

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Clinical Psychology

The University of Sheffield
Faculty of Science
Clinical Psychology Unit, Department of Psychology

Submission Date: December 2020

This Page is left intentionally blank

Declaration

I declare that this thesis has been submitted for the Doctorate in Clinical Psychology at the University of Sheffield. This thesis has not been submitted for the purpose of obtaining any other qualification or to any other institution.

Structure and Word Counts

Part I: Literature Review

Excluding references and tables: 7,953

Including references and tables: 12,320

Part II: Research Report

Excluding references and tables: 7,999

Including references and tables: 9,778

Total Word Counts

Excluding references and tables: 15,952

Including references and tables: 22,098

Lay Summary

Cognitive analytic therapy (CAT) is a popular psychotherapy used by clinicians for a variety of mental health presentations. CAT has an emerging evidence base and a theoretically grounded outcome measure called the Personality Structure Questionnaire. A major aspect of the CAT evidence base consists of single-case experimental design (SCED) studies. As CAT adopts a non-manualised approach, then such studies are well-matched in terms of evaluations, as SCEDs are often utilised when interventions are tailored to individuals. SCEDs have unique methodological features including the presence of different study phases, with discrete baseline and intervention phases, and repeated measurement of problems of interest throughout. The target problem approach of CAT is therefore again well-matched to SCEDs, as target problems are measured by idiographic measures in single-case studies.

Part I of this thesis is a systematic review exploring the effectiveness of CAT via the single-case evidence base. The results of 24 studies, including 26 patients, were identified and synthesised. Both published papers and studies available in the grey literature were reviewed. The quality of studies was assessed, with differences observed between published and unpublished studies. CAT was found to be effective for some diagnoses, but outcomes were less promising for other diagnoses such as anxiety and depression. Published studies were more likely to report effective CAT interventions; possible sources of bias are explored.

Part II of this thesis is a network analysis of the Personality Structure Questionnaire (PSQ). The PSQ is an 8-item measure of identity disturbance and is central to assessment and outcome in CAT. A large sample (n = 1549) of secondary PSQ data was utilised to generate an overall network of identity disturbance and to

compare networks of identity disturbance between the UK and Italy, adults and adolescents, clinical and community, and those with complex diagnoses and other diagnoses. There was a significant difference in networks of identity disturbance between the UK and Italy. No differences were identified in networks of identity disturbance between age, sample type or diagnosis. The results provide insight into identity disturbance and provide further evidence supporting the routine use of the PSQ in clinical settings.

Together, these studies contribute to the evidence base for CAT and its primary outcome measure. The findings suggest that the effectiveness of CAT for patients' target problems is dependent on their primary diagnosis and the qualification of the therapist. The networks of identity disturbance found allow clinicians using the PSQ to identify possible treatment targets for identity integration.

Acknowledgements

Firstly, I would like to thank my research supervisors, Steve Kellett and Mel Simmonds-Buckley, who have been hugely supportive in helping me turn a new project around after my first fell through due to Covid. It hasn't been an easy road; thank you for sticking with me. I could not have done this without your time, knowledge and support.

Thank you to everyone who provided data and collaborated on this research: Peter Taylor, Amy Degnan, Charlotte Humphrey, Annie Nehmad, Kate Freshwater, Raúl Berrios, Marisa Poggioli, Fiorani Stoffella, Samantha Hartley and Margaret Klimyuk. This wouldn't have been possible without you and I am incredibly grateful.

A special mention to my clinical supervisors and patients over the last three years; thank you for sharing your knowledge and experience and for trusting me with your stories. It has been a privilege to work with so many incredible, resilient people and has reminded me what all of this hard work is for.

To my parents, Pat and Peter. Thank you for always inspiring me and encouraging me to be the best I can be. There is no doubt in my mind that I have come this far because of you. I am so grateful for the support and love you have given me at every step of my journey.

To my Fiancé, James, for your unconditional support. Your belief that I could do this has never wavered, even when I haven't believed in myself. For that I will be eternally grateful.

Table of Contents

Part I

Literature Review

How effective is Cognitive Analytic Therapy (CAT) for target problems?

A systematic review of the single-case evidence base

Abstract

Objectives

Cognitive analytic therapy (CAT) is a popular integrative and transdiagnostic psychotherapy. In CAT, target problems are reformulated, recognised and then revised; this fits closely with the notion of idiographic outcome measures in single-case experimental designs. This systematic literature review had two aims. First, to identify and review the quality of studies that have used single-case methodologies to evaluate the effectiveness of CAT for target problems. Second, to provide a narrative synthesis of the effectiveness of CAT for target problems.

Method

A systematic review (PROSPERO registration CRD42020172589) of the literature was completed to identify CAT studies utilising single-case methodologies. Electronic searches were conducted in Scopus, PsycInfo and Medline databases. Grey literature was obtained by contacting CAT practitioners and accessing University course trainee reports. Inclusion and exclusion criteria were applied to screen papers for eligibility based on the PICO framework.

Results

A total of 24 studies were included consisting of 26 participants. CAT was more effective for some diagnoses, such as morbid jealousy and personality disorders. Outcomes were less positive for more common diagnoses including anxiety and depression. Published studies were more likely to present positive outcomes and were generally of higher methodological quality than unpublished studies.

Conclusions

The review demonstrated that CAT is effective for target problems for some diagnoses, but the evidence for more common diagnoses, such as anxiety and depression, was less

promising. Quality varied between studies; it is important for future researchers to follow quality guidelines and single-case reporting guidelines.

Practitioner Points

- The effectiveness of CAT seems to differ depending on diagnosis; this should be considered when determining the appropriateness of CAT for patients.
- Outcomes may be less reliable where idiographic measures are vague or difficult to quantify.
- Measures should be designed collaboratively to ensure they are meaningful and accessible, with frequency-based measures being utilised where possible.

Limitations

- Despite the large number of studies included, only a small number of participants per diagnosis were available.
- The majority of studies identified through grey literature searches were academic reports completed by Trainee Clinical Psychologists. This had implications in terms of therapist qualification and treatment durations; this may have impacted on intervention outcomes.

Keywords: Cognitive Analytic Therapy, Single Case Methodologies

Introduction

What is CAT?

Cognitive analytic therapy (CAT) was developed as a time-limited, integrative psychotherapy (Ryle, 1995). CAT integrates techniques which are derived from behavioural and cognitive models (Ryle, Kellett, Hepple & Calvert, 2014) and incorporates ideas from personal construct, object relations and social developmental theory (Taylor, Perry, Hutton, Seddon & Tan, 2015). Since its development, CAT has grown in popularity and is now widely used and applied to a range of presenting problems in the UK and overseas; namely Finland, Ireland, Spain, Italy, Australia, Greece and India (Ryle et al., 2014). The use of CAT has also broadened to different settings, including indirect use in multi-disciplinary teams. Cognitive analytic consultancy has shown promise in helping professionals to work in a more relationally-informed way, subsequently improving therapeutic relationships (Kellett et al., 2019).

Individual CAT is delivered in either 8, 16 or 24 session contracts with one follow-up session for the 8 and 16 session versions and four follow-ups for the 24-session version (Ryle, et al., 2014). CAT consists of three stages: reformulation, recognition and revision (Taylor, Jones, Huntley & Seddon, 2017). The reformulation stage of CAT involves exploring the influences of early experiences on the development of unhelpful patterns which maintain current difficulties (Carradice, 2004). This reformulation is explicitly shared using both a narrative reformulation letter and a diagram called a sequential diagrammatic reformulation (SDR); these set the agenda for therapy (Rayner, Thompson & Walsh, 2011). Integral to CAT is the understanding that actions and relationships are both causal and caused by others (Llewelyn, 2003); this range of self-self, self-other and other-self relationship patterns are described as reciprocal roles (Ryle & Fawkes, 2007). In the recognition stage of CAT, the goal is to

develop the client's awareness and observe the processes identified in the reformulation (Taylor et al., 2017). Subsequently, the therapist and client work collaboratively to identify alternative ways of relating, known as 'exits' from the unhelpful patterns identified (Taylor et al., 2017).

The evidence-base

Despite its prevalence in diverse settings with a range of presenting problems, a criticism of CAT is that its popularity in practice has sped ahead of the scientific evidence (Marriott & Kellett, 2009). The evidence for CAT is grounded in practice-based evidence and it has previously been identified that CAT did not meet the gold-standard for evidence-based practice (Llewelyn, 2003). It was therefore emphasised that further research, particularly in the form of randomised-controlled trials (RCTs), was needed (Margison, 2000). Since, RCTs and pilot RCTs have contributed to the evidence base, exploring the effectiveness of CAT with presentations including bipolar disorder (Evans, Kellett, Heyland, Hall & Majid, 2017), personality disorder (Clarke, Thomas & James, 2013) and borderline personality disorder (BPD) in adolescents (Chanen et al., 2008).

An initial review of the developing CAT evidence base, incorporating studies of various methodologies, found studies to be generally sound with 52% of high quality (Calvert & Kellett, 2014). A recent meta-analysis has expanded on this review, analysing 25 studies including 10 RCTs, as well as pre-post studies and case series, to provide a quantitative synthesis of the CAT evidence-base (Hallam, Simmonds-Buckley, Kellett, Greenhill & Jones, 2020). This meta-analysis found that CAT led to improvements in global functioning and interpersonal problems and reductions in depression symptoms in patients with a range of presenting problems. Whilst this

review is an important development in the evidence base, the use of standardised nomothetic measures to assess more general psychological functioning does not evaluate how useful CAT is for target problems (i.e. idiographic change). It has been argued that change in target problems specific to the patient is crucial in evaluating the effectiveness of CAT (Ryle & Kerr, 2002). Single-case experimental designs are an evaluation methodology centring on such target problems and thus, research exploring change resulting from CAT with single-case studies has been recommended (Hallam et al., 2020).

What are single-case methodologies?

Whilst RCTs have been identified as gold-standard, they have limitations including the artificial environments they are conducted in (Morley, 2018) and threats to external validity (Sanson-Fisher, Boevski, Green & D'este, 2007). Greater utilisation of the 'natural laboratory' of the clinic setting is therefore invaluable in psychological intervention research (Westen, Novotny & Thompson-Brenner, 2004). Single-case methodology is defined as a collection of methods used to evaluate the effectiveness of interventions in clinical practice based on analysis of change in individuals (Turpin, 2001). Single-case experimental designs (SCEDs) are a time-effective, cost-effective alternative to other methodologies, offering high internal and external validity (Rizvi & Nock, 2008). An advantage to single-case methodologies is that they provide opportunity to establish a cause and effect relationship between an intervention and behaviours targeted by the intervention (Tate, Perdices & Wakim, 2020).

Single-case methodologies are used when the problem of interest has a low base rate and large numbers of participants are unavailable, when interventions require tailoring to individuals and as a form of pilot work prior to larger-scale studies

(Shadish, 2014). Single-case methodologies can include more than one participant but their methodological characteristics differentiate them from other research designs (Lobo, Moeyaert, Baraldi Cunha & Babik, 2017). SCEDs differ significantly from anecdotal qualitative case studies which aim to richly describe and explore a case, but without the exploration of relationships between an independent variable and behaviour (Cohen, Feinstein, Masuda & Vowles, 2014).

The 'target problem' approach that CAT utilises provides a valuable opportunity for the effectiveness of interventions to be explored at an idiographic level (Kellett & Lees, 2019). Within single-case methodologies, specific problem(s) are identified and measured repeatedly throughout phases (Tate et al., 2020). The use of discrete phases is a feature of single-case methodologies that sets them apart from other methodologies but that also differs from routine clinical practice. All SCEDs consist of a baseline (phase A) and the introduction of an intervention (phase B) or series of interventions (Tate et al., 2020). During the baseline, data on the target problem is gathered before intervention commences (Rizvi & Nock, 2008). The presence of a baseline is important in evaluating treatment as it provides a clear picture of what patients' problems would be like without intervention (Kazdin, 1978). The use of different phases means that a participant serves as their own control, exploring change across time (Lobo et al., 2017). The most robust single-case methodologies use a withdrawal, or A/B/A/B, design; these have most often been applied with behavioural therapy. A/B/A/B designs involve the intervention being delivered (first phase B) and then being withdrawn (second phase A) before being reinstated (Kazdin, 1978). In such designs, interventions are deemed effective when the removal of treatment leads to outcomes returning to baseline levels. However, it is acknowledged that this is not necessarily true of psychotherapy

interventions as the aim is for patients to learn and implement methods of change post-therapy (Kellett, Gausden & Gaskell, 2020).

As SCEDs typically have either individual or a small number of participants, it is not possible to include formal analysis of moderating variables, as other experimental group design studies can (Maggin & Odom, 2014). Attempts have been made to improve the validity of single-case methodologies, including the development of hermeneutic single-case efficacy designs (HSCEDs). HSCEDs are a mixed quantitative and qualitative design which first aim to identify causal links between therapy and outcomes and then consider non-therapy explanations for change (Elliott, 2002). This methodology involves the development of a rich case record in which detailed evidence, not collected as part of a traditional SCED, is utilised (Spence, Kellett, Totterdell & Parry, 2019). A further 'adjudicated' form of HSCED has also been developed; this mimics a legal process, involving teams of researchers who argue for or against therapy as the explanation for change (Elliott, 2002). There has been a call for further adjudicated HSCEDs to explore the effectiveness of CAT at a single-case level (Calvert & Kellett, 2014).

Reviews of single-case research

Given that replication of successful outcomes is a limitation of single-case methodologies, it has been suggested that systematic reviews of single-case research are required to explore in which conditions the intervention is effective (Maggin & Odom, 2014). Additionally, it has been highlighted that a particular area of importance for the synthesis of single-case research is assessment of methodological quality (Maggin, O'Keefe & Johnson, 2011), as the absence of quality appraisal has been identified as a limitation of existing reviews (Fallon, Collier-Meek, Maggin, Sanetti & Johnson, 2015).

There is no known review of the evidence for CAT specific to single-case methodologies, thus highlighting a gap in the literature.

Objectives

The present review has the following objectives:

1) To identify and review the quality of studies that have used single-case methodologies to explore the effectiveness of CAT for target problems.
2) To provide a narrative synthesis of the effectiveness of CAT for the target problems of the patient.

Method

Protocol registration

A review protocol was published prospectively on the PROSPERO database (reference: CRD42020172589).

Search strategy

Following preliminary scoping of the literature, a comprehensive electronic search was conducted. Three electronic databases, SCOPUS, PsycInfo and Medline, were used. Literature was included from the date the databases began until 11th September 2020. Searches were limited to studies available in English. The search terms included the following keywords and Boolean operators:

"cognitive analytic therapy" AND "single case" OR "single Sample" OR "SCED" OR "case series" OR "n=1" OR "n-of-1" OR "idiographic" OR "reformulation"

Records were identified from an initial search and duplicates removed. Following this, studies were screened by title and abstract; any studies which immediately met exclusion criteria were removed. Subsequently, full text articles were retrieved and reviewed utilising inclusion and exclusion criteria. Forward and backward reference searching was completed on the papers that were identified as eligible in order to identify additional relevant studies. Attempts were also made to access studies in the grey literature. The University of Sheffield was contacted to obtain single-case reports utilising CAT that had been submitted as part of the Doctorate in Clinical Psychology programme. The University of Sheffield was chosen specifically as it is the only doctorate course for which a SCED assignment is a requirement. Lastly, key authors were approached in order to obtain unpublished studies.

Inclusion and exclusion criteria

Inclusion and exclusion criteria were guided by the PICO framework (Richardson, Wilson, Nishikawa & Hayward, 1995) and are displayed in Table 1.

Table 1. *Inclusion and Exclusion Criteria Informed by PICO Framework*

	Inclusion	Exclusion
Population	Patients with an identified target problem	Participants who were not patients, e.g. healthcare workers
Intervention	CAT, delivered directly to the patient	Any other psychological intervention or indirect CAT e.g. Cognitive Analytic Consultancy
Comparison	Participants act as their own comparator, with repeated measurement over distinct baseline and treatment phases	Studies where there is no comparator, i.e. no distinct baseline
Outcome	Idiographic outcomes utilised to measure target problems	Studies which do not measure target problems utilising an idiographic outcome
Study Design	A recognised quantitative single-case design	Any other research design e.g. RCT, case study, qualitative paper

Data extraction

A coding tool developed and piloted by the researcher (Appendix A) was used to extract the following key information from studies: authors, year, study design, patient demographics (age, gender), clinical characteristics (presentation, medication, treatment history), intervention setting, intervention characteristics (phase lengths), measurement characteristics (number and frequency of idiographic and nomothetic measures), analysis (nomothetic, idiographic, autocorrelation, graphing), competency assessment, therapist characteristics (qualification, gender, supervision) and outcomes. Extracted information was synthesised and a narrative summary presented.

Quality appraisal

Each study was rated using the Single-Case Experimental Design Scale (Tate et al., 2008). The SCED scale is an 11-item rating scale which has been found to provide a brief and valid evaluation of the methodological quality of single-case studies. The scale was developed based on weaknesses of SCEDs which threaten validity. Ten of the SCED scale items relate to methodological quality; there is one additional item, specification of clinical history, which does not contribute to the score. The scale is scored dichotomously, either present or absent, with one point given for each item explicitly evidenced in the report. Quality scores thus range from 0-10; higher scores indicate stronger methodological quality. No cut-off quality score is defined.

To assess interrater reliability, an independent second-rater (Trainee Clinical Psychologist) rated six (25%) of the included papers. A sample of papers was selected at random from both published and unpublished studies. The second-rater was blind to the author's quality scores. Cohen's kappa coefficient (Cohen, 1960) was calculated to assess inter-rater reliability. This demonstrated that there was a moderate level of

agreement between ratings ($k = 0.586$, $p = 0.002$). Discrepancies were due to differing interpretations of the SCED scale guidance and were resolved through discussion.

Data Synthesis

Data was synthesised using narrative synthesis. Due to the range of diagnoses represented in the literature, synthesis was broken down into two meta-categories of complex and other diagnoses. More specific diagnoses, such as depressive disorders or personality disorders, were then identified and data was synthesised accordingly.

When synthesising the measurement of target problems, the author categorised idiographic scales as measurement of thoughts, feelings or behaviours. Any scales that were unclear in terms of categorisation were discussed in a consensus meeting with two other researchers. Following the consensus meeting, it was decided that feelings would be further broken down into emotions and physiology, to ensure scales were appropriately and meaningfully categorised.

Results

Figure 1 presents the results of the search strategy. The searches initially identified 128 records. Title and abstract screening identified 29 studies for full-text review. Of those, 24 papers comprising 26 individual patients treated with CAT met the eligibility criteria. Study characteristics are presented in Table 2. Each study was allocated a number and these numbers will be used to cite studies in-text. Ten studies (42%) were published studies and fourteen studies (58%) were identified through the grey literature search; papers are categorised by publication status in Tables 2 and 3. Included studies were conducted between 2003 and 2020.

Quality Appraisal

Quality appraisal scores are summarised in Table 3 and full quality appraisal results are presented in Appendix B. Quality scores ranged from 5-9 (Mean = 6.38, *SD* = 1.47). All studies included in the review described patients' clinical histories, although this item does not contribute to the quality score. All studies met quality criteria for identifying target behaviours, providing an adequate baseline, sampling behaviour during treatment and providing a raw data record. All but one study (5) met quality criteria with regard to statistical analysis.

Quality was generally not met in terms of design, with all but two studies scoring 0 on this item due to their quasi-experimental A/B designs. Quality was inconsistent with regard to inter-rater reliability and independence of assessors, with only seven studies scoring on these items. Quality was also inconsistent with regard to replicability and generalisability; seven studies scored for replicability and eleven studies scored for generalisability.

Quality scores for published studies ranged from 6-9 (Mean = 7.40, *SD* = 1.26) whilst quality scores for unpublished studies ranged from 5-8 (Mean = 5.64, *SD* = 1.15). A Mann-Whitney U test found that quality was significantly higher in published studies ($U = 5.00$, $p > 0.001$). Specifically, published studies were more likely to score for replicability and generalisability. Additionally, only two unpublished studies scored for inter-rater reliability and independence of assessors; one of these studies (18) had been submitted for publication. Thus, published studies or studies intended for publication are more likely to fulfil quality criteria. Studies that identified CAT to be effective were generally of higher quality whilst the studies with the lowest quality scores generally reported ineffective interventions.

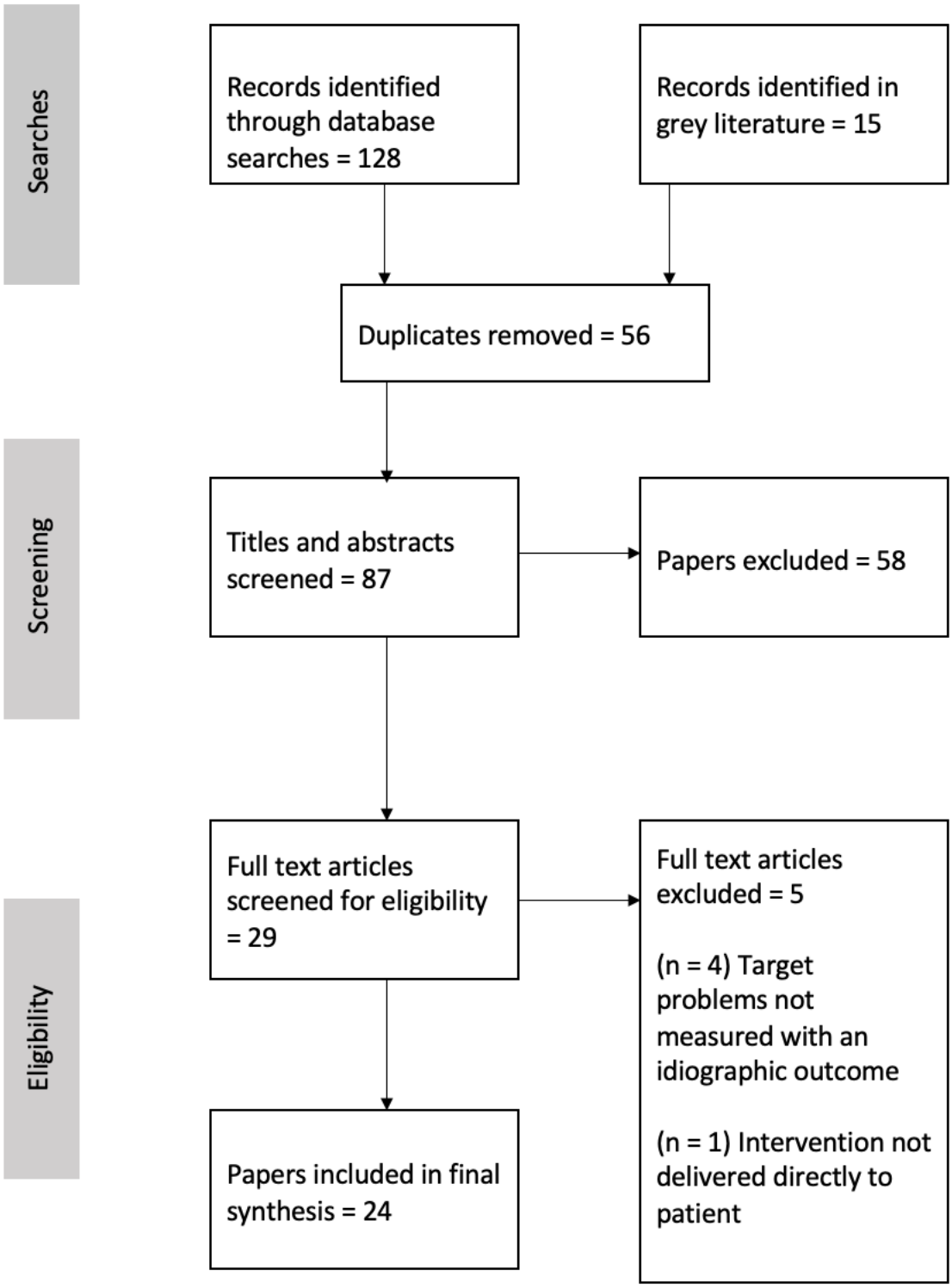


Figure 1. PRISMA diagram depicting literature search process and results

Table 2. *Study Characteristics of Included Studies*

	First Author, Year	Design	Patient (Gender, Age)	Setting	Therapist Experience	Diagnosis, Target Problem	Previous Treatment	Duration: Baseline, Intervention, Follow-Up
1.	Curling et al. (2018a)	Case series (*n* =3). A/B with follow-up	Female (54) Female (36) Male (58)	Routine secondary care, NHS	Consultant Clinical Psychologist, CAT therapist	Obsessive Morbid Jealousy: jealousy	1. Unresponsive to low-intensity CBT 2. Unresponsive to counselling and high-intensity CBT 3. None	1) A - 6 weeks, B - 42.5 weeks, FU - 20.5 weeks 2) A-7, B-13.5, FU-20.5 3) A-6, B-25, FU-13
2.	Curling et al. (2018b)	Adjudicated HSCED: A/B with follow-up	Female (38)	Routine care conditions, NHS	Consultant Clinical Psychologist, CAT therapist	Obsessive Morbid Jealousy	None	A – 27 days B – 90 days FU – 76 days
3.	Kellett (2005)	A/B with follow-up	Female (unknown age)	Not stated	Consultant Clinical Psychologist, CAT therapist	Dissociative Identity Disorder: dissociation	Long history including counselling, inpatient at behavioural unit, groups, psychotherapy.	A – 5 weeks B – 25 weeks FU – 24 weeks
4.	Kellett (2007)	A/B with follow-up	Female (21)	Not stated	Consultant Clinical Psychologist, CAT therapist	Histrionic Personality Disorder: HPD behaviours	Medication only - nonresponsive	A – 21 days B – 182 days FU – 154 days
5.	Kellett et al. (2020)	A/B/A/B	Female (46)	Outpatient psychotherapy service, NHS	Consultant Clinical Psychologist, CAT therapist	BPD: sensitivity to abandonment	Dropped out of counselling and CBT. Accessed BPD group.	A – 6 weeks B1 – 24 weeks B2 28 weeks FU – 25 weeks
6.	Kellett & Hardy (2014)	Mixed methods A/B with follow-up	Male (36)	Secondary care CMHT	Consultant Clinical Psychologist, CAT therapist	Paranoid Personality Disorder: paranoia, anxiety	None	A – 42 days B – 42 days FU – 140 days

7.	Kellett & Lees (2019)	A/B with follow-up	Female (29)	Secondary care CMHT	Consultant Clinical Psychologist, CAT therapist	Dependent Personality Disorder: dependency traits, anxiety, depression	Nonresponsive to medication and counselling intervention	A– 4 weeks B – 30 weeks FU – 27 weeks
8.	Kellett et al. (2017)	A/B with follow-up	Male (41)	NHS	Consultant Clinical Psychologist, CAT therapist	Hypersexuality Disorder: sex addiction, cruising, masturbation, pornography usage	Unknown	A – 21 days B – 147 days FU – 63 days
9.	Kellett & Totterdell (2013)	A/B with follow-up	Female (unknown age)	Adult mental health, NHS	Consultant Clinical Psychologist, CAT therapist	Morbid jealousy: experience of jealousy, partner hypervigilance, jealous disinhibition and anxiety	Unknown	A – 35 days B – 98 days FU – 84 days
10.	Spence et al. (2019)	Adjudicated HSCED: A/B with follow-up	Male (64)	Tertiary outpatient psychotherapy service, NHS	Clinical Psychologist, CAT therapist	Hoarding Disorder: chronic hoarding problems	Two previous CBT interventions: one terminated due to lack of progress, one drop-out	A – 27 days B – 273 days FU – 175 days
					Unpublished Studies			
11.	Collins (2011)	A/B	Female (37)	CMHT	Trainee Clinical Psychologist	BPD: flashbacks, panic, anger, sadness	Unknown	A – 29 days B – 34 days
12.	Crosby (2015)	A/B	Female (24)	Secondary care CMHT	Trainee Clinical Psychologist	GAD: perfectionism, worry, feelings of lack of purpose	Previous CBT with IAPT	A – 4 weeks B – 7 weeks

13.	Field (2003)	A/B	Female (48)	Not stated	Trainee Clinical Psychologist	OCD: compulsive behaviours, obsessional thoughts, anxiety, enjoyment of day	Previous input from Clinical Psychologist lasting two years	A – 4 weeks B – 9 weeks
14.	Hall (2013)	A/B with follow-up	Female (44)	CMHT	Trainee Clinical Psychologist	Bulimia Nervosa: feeling accepted, self-care, ok as me	Previous group CBT; effective short-term	A – 14 days B – 48 days FU – 58 days
15.	Harvey (2011)	A/B	Female (58)	Adult mental health service	Trainee Clinical Psychologist	Depression: depression, motivation, anxiety	Not stated	A – 28 days B – 84 days
16.	Heathcote (2019)	A/B	Female (39)	Secondary CMHT	Trainee Clinical Psychologist	BPD: sleep, emotion regulation, sense of self	DBT; successful in reducing self-harm	A – 14 days B – 43 days
17.	Kellett et al. (2020)	A/B/A/B with follow-up	Female (26)	Specialist psychotherapy service	CAT therapist	Bipolar II: self-critical, compassion, body image, mood, worry	Counselling	A – 44 days B – 168 days A – 82 days FU – 167 days
18.	Kellett & Stockton (2020)	Mixed methods A/B with follow-up	Male (50)	Tertiary specialist psychotherapy service	Consultant Clinical Psychologist, CAT therapist	Obsessive Morbid Jealousy: managing jealousy to reduce relationship impact	None	A – 21 days B– 70 days FU – 69 days

19.	Power (2019)	A/B with follow-up	Male (25)	Secondary care community therapies	Trainee Clinical Psychologist	Major Depressive Disorder: self-worth, connection with feelings/others	Not stated	A– 21 days B – 77 days FU – 6 weeks
20.	Ryder (2010)	A/B	Male (49)	Secondary care CMHT	Trainee Clinical Psychologist	Anxiety: anxiety symptoms, confidence in social situations	16 sessions of solution-focused therapy, 1 year graded exposure, relaxation group	A – 28 days B – 70 days
21.	Sherman (2014)	A/B	Male (25)	Secondary care CMHT, NHS	Trainee Clinical Psychologist	Depression: low mood, anger	Previous CBT; initially effective but benefits not maintained	A – 6 weeks B – 9 weeks
22.	Sims (2015)	A/B	Female (unknown age)	Secondary care community therapies	Trainee Clinical Psychologist	Depression: confidence in coping ability, self-criticism, activity	Previous emotional coping skills group, unsuccessful CBT from IAPT	A – 14 days B – 112 days
23.	Swan (2009)	A/B with follow-up	Female (36)	CMHT	Trainee Clinical Psychologist	Bulimia Nervosa: feeling not good enough, guilt, urge to vomit	6-week inpatient admission to eating disorder unit	A – 14 days B – 48 days FU – 3 months
24.	Walsh (2014)	A/B	Male (30s)	Secondary mental health	Trainee Clinical Psychologist	Depression: low motivation	Previous CBT, self-help, counselling and medication; all ineffective	A – 14 days B – 35 days

Table 3. *Methodology and Findings of Included Studies*

	First Author, Year	Idiographic Measures	Nomothetic Measures	Nomothetic Analysis	Idiographic Analysis	Outcome	Quality Rating
1.	Curling et al. (2018a)	4-6, daily: jealousy, watchfulness, trust, self-confidence, depression, secure, body image, rational, anxiety	PJQ, BDI-II, IIP-32, BSI. Assessment, termination, follow-up	Reliable & clinically significant change	ANCOVA, post-hoc pairwise comparison, Bonferroni corrections, PEM	Intervention effective for all three patients; reductions in morbid jealousy. Cases classed as not jealous by follow-up; partner violence extinguished.	7
2.	Curling et al. (2018b)	6, daily: jealousy intensity, compulsive observation, state-shifting, anxiety, self-esteem, behavioural balance	PJQ, BDI-II, IIP-32, BSI. Assessment, termination, follow-up	Reliable & clinically significant change	ANCOVA, post-hoc pairwise comparison, Bonferroni corrections, PEM	Intervention effective	9
3.	Kellett (2005)	7, daily: derealization, depersonalization, identity confusion, identity alteration, conversion, amnesia, identity shifts	BSI, BDI-II, IIP-32, DES-II, PSQ. Assessment, termination, follow-up	Reliable & clinically significant change	Visual analysis	Intervention effective: reduction in state and trait dissociation, improved integration of personality, improved depression and general psychiatric symptoms	6
4.	Kellett (2007)	5, daily: need to be noticed, focus on appearance, flirting, felt empty, felt like a child	BSI, BDI-II, IIP-32, PSQ, YSQ-SV. Assessment, termination, follow-up	Reliable & clinically significant change	ANOVA, t-tests	Intervention effective in reducing HPD symptoms, personality integration, depression. Sudden deterioration at termination	7
5.	Kellett et al. (2020)	7, daily: feeling abandoned, anxious, lonely, over-sensitive; self-hate,	BSL-23, BDI-II, BSI, IIP-32. Baseline and end of phases.	Reliable & clinically significant change	ANCOVA, Bonferroni correction, PEM, non-overlap of	Partially effective intervention. Observed deterioration from second treatment phase to follow-up across idiographic measures	8

		sense-of-self, distrusting of others	SIS (each session))		pairs, TAU-U		
6.	Kellett & Hardy (2014)	6, daily: suspiciousness, hypervigilance, dissociation, conspiracy, questioning, anxiety	BSI, BDI-II, IIP-32, PSQ. Assessment, termination, follow-up	Reliable & clinically significant change	Interrupted time series analysis	Intervention effective; six of seven target measures extinguished during treatment	6
7.	Kellett & Lees (2019)	6, daily: reassurance-seeking, avoidance of decisions, numbness, cognitive clarity, depression, anxiety	IDI, BDI-II, BSI, IIP-32. Assessment, termination, follow-up	Reliable & clinically significant change	ANCOVA, Bonferroni correction	Mixed outcome; not universally effective. Significant effect on primary idiographic measure (reassurance seeking) and nomothetic measure (self-confidence)	7
8.	Kellett et al. (2017)	7, daily: cruising, pornography, masturbation, sexual intrusions, interpersonal connectivity, anxiety, self-worth	SCS, BDI-II, BSI, IIP-32. Assessment, termination, follow-up	Reliable & clinically significant change	Mediation analysis, ANCOVA	Intervention effective. Cruising and pornography use extinguished. Primary nomothetic measure met recovery criteria	6
9.	Kellett & Totterdell (2013)	Patient, 5, daily: jealousy, anxiety, self-esteem, disinhibition, hypervigilance. Partner, 2, daily: jealousy, controlled	BSI (global severity index, BDI-II, IIP-32, RJQ, PJQ. Assessment, termination, follow-up	Reliable & clinically significant change	Time-series intervention analysis	Intervention effective for patient with reductions in jealousy and hypervigilance. Partner did not perceive change. CBT comparison treatment for second patient not effective	9
10.	Spence et al. (2019)	5, daily-three times daily: acquisition, stealing, discarding, fantasy proneness, anxiety	SIR, BDI-II, BSI, SIS, Clutter Image Rating. Assessment, termination, follow-up	Reliable & clinically significant change. Independent samples t-test	ANCOVA, PEM, logistical regression	Judges concluded CAT did not enable change due to overall outcomes. Improvements in target problems were observed.	9

			Unpublished Studies				
11.	Collins (2011)	4, daily: feeling chilled, peaceful; number of flashbacks, panic intensity	CORE-OM, IES-R. Pre/post assessment, pre/ post intervention. SRS (each session)	Reliable & clinically significant change	PEM, PND, PAND. Non-parametric Mann-Whitney U test.	Intervention not effective for target problems nor nomothetic outcomes. Positive outcome in terms of experience of therapy and therapeutic relationship	5
12.	Crosby (2015)	3, daily: feel I must prove myself/be perfect, worried, feel I don't have purpose	CORE-10 (pre/post baseline, weekly in phase B). IIP-32 (pre/post baseline, post-intervention)	Reliable & clinically significant change	PND, PEM, PAND	Intervention not effective for target problems. Improvement in interpersonal relationships and level of distress	5
13.	Field (2003)	3, daily: handwashing, checking, thoughts about handwashing	MOI, Padua Inventory, HADS, CORE. Assessment, end of intervention	Clinical change	Mann Whitney U test, Spearman's rho correlations	Intervention effective	5
14.	Hall (2013)	3, daily: feeling ok to be me, feeling accepted, caring for me	CORE-OM, IIP-32 (baseline, termination, follow-up). CORE-10 (weekly)	Reliable & clinically significant change	PND, PAND, PEM	Intervention effective for target problems and global distress. Intervention ineffective for interpersonal difficulties	5
15.	Harvey (2011)	3, daily: depression, motivation, anxiety	CORE-OM, assessment, post-assessment, termination, follow-up	Reliable & clinically significant change	PND, PAND, PEM, Mann-Whitney U test	Intervention not effective. Despite some improvement in global distress, patient remained clinically depressed	5
16.	Heathcote (2019)	3, daily: sleep, sense of self, emotion regulation	IIP-32, weekly. CORE-OM, pre/post baseline, termination	Reliable & clinically significant change	PND, PAND, PEM	Intervention not effective for target problems. Improved relational functioning and global distress.	5

17.	Kellett et al. (2020)	5, daily: self-criticism, compassion, body image, depression, worrying	PHQ-9, Mania Rating Scale. Each session.	Reliable & clinically significant change	PND, PAND, PEM	Positive outcome initially with positive life events. Adverse event in follow-up period (hypomanic episode)	8
18.	Kellett & Stockton (2020)	7, daily: jealousy, anxiety, trust, over-thinking, functioning as a couple, intrusive images, checking	PJQ, BDI-II, BSI, IIP-32. Assessment, termination and follow-up	Reliable & clinically significant change	ANCOVA, post-hoc pairwise comparison, Bonferroni correction, PND, PEM	Intervention effective. Idiographic jealousy measures reduced and improvements maintained over follow-up. Jealousy intensity reduced from moderate to mild.	8
19.	Power (2019)	3, daily: self-worth/confidence, feeling emotions, connection with others	CORE-10 (weekly). CORE-OM, IIP-32 (pre/post baseline, termination, follow-up)	Reliable & clinically significant change	PND, PAND, PEM	Mixed outcome; intervention effective in reducing psychological distress and enhancing self-worth and connection with feelings. Lack of evidence for connection to others	7
20.	Ryder (2010)	4, daily: tingling sensations, confidence, leaving house, speaking on the phone	CORE-OM, BDI-II, GAD-7 (pre/post baseline, termination). CORE-10 (weekly)	Reliable & clinically significant change	PND, PAND, PEM, Mann-Whitney U test	Intervention generally not effective; trend towards improvement for target problems but not statistically significant. Improvement in GAD-7 scores.	5
21.	Sherman (2014)	2, daily: low mood, anger	CORE-OM, IIP-32. Pre/post baseline and termination	Reliable & clinically significant change	PND, PAND, PEM	Intervention not effective. Some improvements were observed but these were during baseline and thus not attributable to CAT	5
22.	Sims (2015)	3, daily: confidence in coping, self-criticism, meaningful activity	CORE-OM (pre/post baseline, termination). IIP-32 (pre/post intervention). CORE-10 (weekly)	Reliable & clinically significant change	PND, PAND, PEM	Intervention not effective; no improvement on nomothetic measures or idiographic measures of self-criticism or confidence. Change was observed in engagement in activity	5

23.	Swan (2009)	4, daily: feeling not good enough, guilt, vomit urge, positivity about progress living with bulimia	BDI-II, BSI, IIP-32 (pre/post baseline, termination, follow-up). CORE-10 (weekly)	Reliable & clinically significant change	PND, Mann-Whitney U test	Mixed outcome. Intervention successful in reducing core pain, increasing positivity and improved psychological functioning. Unsuccessful for reducing urge to vomit	6
24.	Walsh (2014)	5, daily: self-care, concentration, attempted work, contacted friends, left house	CORE-OM (weekly). IIP-32 (pre/post baseline, termination)	Reliable & clinically significant change	PND, PEM, PEM-T	Intervention not effective; patient did not attempt his primary goals of returning to university and completing his work	5

Overview of included studies

Patient characteristics

The participants (n = 26) consisted of nine males and seventeen females. Patients were adults aged 21-64 with a mean age of 40[1] (SD = 12.27). Six patients had a primary diagnosis of morbid jealousy, six had a primary diagnosis of a personality disorder and six presented with depressive disorders. Three patients presented with anxiety disorders and two with eating disorders. Dissociative identity disorder (DID), hoarding disorder and hypersexuality disorder were the primary diagnosis for one patient each. Nine participants were reported to be taking medication at the time of intervention. Medication was most commonly anti-depressant medication but also included anxiolytics, tranquilisers and anti-psychotics. Information on medication was not universally reported; ten studies did not explicitly state whether the patient was medicated at the time of therapy.

Setting

Where intervention setting was stated, all studies were completed in outpatient services including secondary care community mental health teams (CMHTs), community therapy teams and specialist psychotherapy services. Seventeen participants had received previous treatment including one-to-one therapies such as cognitive behavioural therapy (CBT), group interventions and inpatient admissions. One participant had been previously treated with medication only. Four participants were reported to have had no previous intervention and for five participants, treatment history was not reported.

[1] From a sample of n = 22 participants for which exact ages were provided

CAT intervention

Intervention durations[2] ranged from 8-30 sessions for interventions with one treatment phase. Of the two studies which utilised A/B/A/B designs, one patient (5) received 21 sessions followed by a further 24, whilst another patient (17) received eighteen sessions followed by a further six. Published studies predominantly utilised 16 or 24-session CAT protocols; five patients were treated with 16-session CAT and six patients with 24-session CAT. There was one exception within the published literature (10) in which 30 sessions were delivered. Unpublished studies, however, did not follow standard CAT protocols with the exception of two which utilised 16-session CAT (15, 22) and two which utilised brief 8-session CAT (18, 24). The remaining unpublished studies had a variety of durations between 9-18 sessions. Thus, patients represented in the unpublished literature generally received fewer sessions.

Therapist characteristics

Nine of the included studies utilised female therapists whilst fifteen utilised male therapists. Twelve therapists were Trainee Clinical Psychologists. In twelve studies, the therapist was a Consultant Clinical Psychologist and accredited CAT psychotherapist. Within the published literature, the same therapist and author delivered interventions across several studies. In the unpublished literature, each intervention was completed by a different therapist.

In five studies, it was detailed that competency assessment was completed utilising the Competence in CAT (CCAT) measure (Bennett & Parry, 2004); four of these studies were in the published literature. CCAT scores of 20 or above indicate that

[2] Follow-up sessions not included where specified

CAT has been competently delivered (Bennett & Parry, 2004). Of the studies that utilised the CCAT measure, three (10, 18, 19) assessed competency in one session, one (5) assessed competency in two sessions and one (7) assessed competency in three sessions. All sessions that were reviewed using the CCAT measure were above the cut-off of 20; scores ranged from 25-32 with a mean score of 28.75 (*SD* = 3.15).

Twelve studies referenced supervision; one study detailed monthly supervision, seven studies detailed weekly supervision and four studies referenced supervision without reporting frequency. The remaining twelve studies did not refer to supervision.

Design

The majority of studies (83%) were A/B designs. Two studies (2, 10) were adjudicated HSCEDs and two (5, 17) were A/B/A/B designs. Baselines ranged from 14 days to seven weeks[3] (Mean = 29 days[4], *SD* = 10.84). Interventions ranged from 34 days to 42.5 weeks (Mean = 114 days, *SD* = 72.89). Fifteen studies had follow-ups and their length ranged from 6 weeks to 27 weeks (Mean = 119 days, *SD* = 49.23). All published studies contained a follow-up. Nine of the studies identified in the grey literature utilised A/B designs without follow-up.

Idiographic measurement

All studies utilised idiographic measures of target problems, completed daily. The number of scales for each patient ranged from 2-7 with a mean of 4.69 (*SD* = 1.57). The majority of studies utilised idiographic measures completed by the patient only. However, one study (9) utilised an additional idiographic partner scale. Summaries of

[3] Duration reported in days or weeks as per study reporting

[4] Where durations were reported in weeks, these were converted to days to calculate descriptive statistics

idiographic scales by study are displayed in Table 3. Of the idiographic scales, 83.6% measured intensity whilst 15.6% measured frequency. Inadequate detail was provided for one scale. Idiographic scales most commonly measured feelings (51; 41.8%) of which 40 (32.8%) related to emotion and 11 (9%) related to physiology. Thoughts were measured by 41 (33.6%) of scales whilst 30 (24.6%) scales measured behaviour.

Nomothetic measurement

The most frequently used nomothetic measure was the Inventory of Interpersonal Problems 32 (IIP32; Barkham, Hardy & Startup, 1996) which was used in eighteen studies. The IIP-32 is a 32-item measure which assesses interpersonal functioning. The second most commonly used measure was the Beck Depression Inventory (BDI-II; Beck, Steer & Brown, 1996) which was used in thirteen studies. The BDI-II is a 21-item measure of depression symptomology. The Brief Symptom Inventory (BSI; Derogatis, 1993) was also frequently used, with twelve studies utilising this 53-item measure which assesses nine primary symptom dimensions across three global indices: global severity index, positive symptom distress index and positive symptom total. Two measures of psychological distress were commonly used. The Clinical Outcome for Routine Evaluation Outcome Measure (CORE-OM; Barkham et al., 1998), a 34-item measure of distress, was utilised in nine studies. The shortened, 10-item version, the CORE-10 (Barkham et al., 2013), was used in six studies.

Diagnosis-specific measures were also utilised. Of note is the Prestwich Jealousy Questionnaire (PJQ; Beckett, Tarrier, Intili & Beech, 1992) which was used in four studies where the diagnosis was morbid jealousy. The PJQ is a 60-item measure consisting of items relating to the behavioural, cognitive and affective aspects of jealousy. In three studies, where the patient had a diagnosis of a personality disorder or

DID, the Personality Structure Questionnaire (PSQ; Pollock, Broadbent, Clarke, Dorrian & Ryle, 2001) was administered. The PSQ, an 8-item measure of identity disturbance, has been identified to be a measure central to assessment in CAT (Ryle & Kerr, 2002).

The majority of studies utilised nomothetic measures at each phase: baseline, termination and follow-up. Eight studies utilised additional weekly measures. Three studies completed measures at each session; in two studies (5, 11) these were session rating measures. In one study (17), measures of depression, the PHQ-9 and the Mania Rating Scale, were completed on a sessional basis.

Analysis

In 22 of the included studies, autocorrelation analysis was completed. This tests for serial dependency which is when repeated measures of a participant lead to an increased likelihood of errors which can result in the illusion of a more successful intervention (Todman & Dugard, 2001). All included studies provided graphing of data by study phase and all but one study (5) supplemented visual analysis with statistical analysis. The most frequent analysis of idiographic data was percentage of data exceeding the median (PEM; Ma, 2006) which was used in sixteen studies. Percentage of non-overlapping data (PND; Scruggs, Mastropieri & Casto, 1987) was used in thirteen studies and percentage of all non-overlapping data (PAND; Parker, Hagan-Burke & Vannest, 2007) in ten studies. Other frequently used analyses were ANCOVA (seven studies) and Mann-Whitney U tests (five studies). With the exception of one study (13), all studies analysed nomothetic data using reliable and clinically significant change.

Effectiveness of CAT for target problems

In order to synthesise studies, findings will be presented by diagnosis, using two meta-categories of complex[5] diagnoses and other diagnoses. Outcomes for each individual diagnosis represented in the sample are described.

Complex diagnoses

Personality disorders. Six studies explored the effectiveness of CAT for the target problems of patients with personality disorder diagnoses. Target problems were most frequently measured using thought and emotion-based scales and were infrequently measured using physiology-based scales. Three patients had diagnoses of BPD, one histrionic personality disorder (HPD), one dependent personality disorder and one paranoid personality disorder. CAT was not found to be an effective intervention for target problems in patients with BPD. Whilst one study (5) showed initial improvement, there was a decline between second treatment phase and follow-up. The other BPD studies (11, 16) were not successful for target problems. There were, however, improvements in nomothetic outcomes including therapeutic alliance (11), relational functioning and global distress (16).

CAT was demonstrated to be an effective intervention for a patient with HPD (4); there was an improvement in HPD symptoms, personality integration and depression although there was a sudden deterioration at termination of treatment, suggesting the patient found ending difficult. CAT was also shown to be an effective intervention for a patient with paranoid personality disorder (6); six of seven target problems extinguished over the treatment course. Outcomes for dependent personality

[5] Complex diagnoses include personality disorders, psychosis and eating disorders

disorder were mixed; whilst the intervention was not universally effective, it was effective for the primary target problem.

Dissociative identity disorder. One paper explored the effectiveness of CAT for a patient with a diagnosis of DID. Target problems were primarily measured using physiology-based scales. CAT was an effective intervention for this patient; this was in the context of several previous unsuccessful interventions. CAT led to a reduction in dissociation and improvements in personality integration, depression and general psychiatric symptoms.

Eating disorders. Two studies explored the effectiveness of CAT for the target problems of patients with eating disorders; both patients had diagnoses of Bulimia Nervosa. Idiographic scales most commonly measured thoughts, whilst there were no behaviour-based scales. The outcomes of both studies were mixed. One study (14) found CAT to be effective for global distress and target problems around sense of self and feeling accepted. However, there was no improvement in interpersonal relationships. In the second study (23), CAT was effective in reducing core pain, increasing positivity and improving psychological functioning. However, one target problem, urge to vomit, did not improve.

Other diagnoses

Depression. Six studies explored the effectiveness of CAT with patients with depressive disorders. Target problems were most commonly measured with thought and emotion-based scales. One patient within this subsample had a diagnosis of bipolar disorder; positive outcomes were initially observed in this study but there was an adverse event during the follow-up period when the patient experienced a hypomanic episode.

Generally, outcomes for patients with other depressive disorders were less promising. Four studies (15, 21, 22, 24) found that CAT was ineffective for target problems. One study (19) had a mixed outcome; CAT was effective in improving psychological distress and target problems of self-worth and connection with feelings, but there was a lack of improvement in the target problem of connection to others.

Anxiety. Two studies included patients with generalised anxiety and one with obsessive compulsive disorder (OCD). Target problems were most commonly defined as thoughts or behaviours. Both interventions for patients with generalised anxiety were found to be ineffective for target problems although one intervention (12) led to improvements in interpersonal relationships and distress. The OCD study (13) was concluded to be largely effective, with improvements in some target problems (handwashing and checking) as well as nomothetic outcomes. However, the intervention was ineffective for the target problem of anxiety.

Jealousy. Four studies provided interventions for morbid jealousy. One study was a case series and therefore a total of six participants with morbid jealousy were included. Idiographic scales most commonly measured thoughts and emotions, although several scales also measured behaviours. CAT was a universally effective intervention for all six patients across the four studies. Outcomes included reductions in jealousy and extinction of partner violence. One study (9) compared two interventions, with one patient receiving CAT and one patient (not represented in this review) receiving CBT. CAT was found to be effective whilst CBT was ineffective.

Hypersexuality disorder. One study found that CAT was an effective intervention for a patient with hypersexuality disorder. Target problems were most

commonly measured by behaviour-based scales within this study. The patient met recovery criteria on the primary nomothetic measure and their target problems of cruising and pornography use extinguished, with evidence of continued improvement at follow-up.

Hoarding disorder. One study explored the efficacy of CAT for hoarding disorder, exploring predominantly behaviour-based target problems. Judges in this adjudicated HCSED concluded that CAT had not enabled change. It was argued that CAT was not effective and that the therapeutic relationship was more beneficial than the intervention; lack of change on nomothetic hoarding measures was particularly influential in this conclusion. However, it is important to note that despite the overall conclusion of ineffectiveness, significant improvements for four target problems were observed. Thus, whilst not universally effective, CAT was effective for target problems.

Discussion

This review aimed to collate, synthesise and critically appraise the evidence base of single-case studies exploring the effectiveness of CAT for patients' target problems. The PRISMA checklist (Appendix C) was used to ensure rigour in the reporting of the review. Both published and unpublished studies were reviewed. The inclusion of grey literature in systematic reviews has been said to be beneficial in overcoming publication bias, as published studies tend to show larger effects of health interventions than studies identified in grey literature (Hopewell, McDonald, Clarke & Egger, 2007). Thus, the inclusion of grey literature is a strength of this review.

Main findings

There were a range of quality scores for the included studies. Studies that showed ineffective CAT were generally of poorer quality than those reporting effective interventions. However, it is important to consider confounding variables; studies reporting effective interventions were generally published studies and thus quality differences may be due to publication as opposed to causal relationships between quality and effectiveness. Important differences between published and unpublished studies will be explored.

Synthesised findings on effectiveness were grouped by diagnosis. The classification of target problems, and thus classification of idiographic scales, differed between diagnoses. Findings around effectiveness of CAT for the target problems of patients with personality disorder diagnoses were mixed. Findings were promising for patients with histrionic, dependent and paranoid personality disorder diagnoses, but outcomes for target problems were less positive for patients with BPD diagnoses. CAT has been previously identified to be an effective intervention for adults diagnosed with BPD (Kellett, Bennett, Ryle & Thake, 2013). However, it should be noted that this study found improvements in nomothetic outcomes. Similarly, BPD studies included in the review had positive findings in terms of nomothetic outcomes. Idiographic measures, constructed in relation to individuals' specific target problems, however, did not demonstrate effectiveness. Thus, whilst the studies in this review corroborate findings in terms of generalised outcomes, the effectiveness of CAT for target problems with this clinical population is unsupported. Conversely, whilst CAT was not universally effective across all outcomes for patients with eating disorders, both studies presented improvements in target problems.

Encouraging evidence emerged with regard to the effectiveness of CAT for patients with morbid jealousy. Interventions were universally effective for target problems of all patients. One study (9) was a matched SCED which compared outcomes of CAT and CBT for morbid jealousy; CAT was found to be effective whilst CBT was ineffective. This suggests that CAT-specific therapeutic techniques were effective, as opposed to a general treatment effect. Given that credible evaluations of intervention for morbid jealousy are rare (Kellett & Totterdell, 2013) the evidence that has emerged from this review makes an important contribution to the evidence base and warrants larger-scale research.

The majority of studies exploring CAT with patients with depressive disorders or anxiety found interventions to be ineffective for target problems. There were some exceptions; one study found initial positive outcomes with an adverse event in the follow-up period, one study found improvement in some, but not all, target problems and one intervention (for OCD) was effective. The studies which had mixed outcomes had follow-ups and one employed an A/B/A/B design. Thus, it is possible that the lack of effectiveness in A/B studies may have been influenced by design, with the absence of follow-up providing less opportunity for change to be observed. It is of interest that the OCD study utilised frequency-based measures that were easily quantifiable, measuring behaviours such as number of handwashes. These measures are likely to have been more concrete and thus more sensitive to change than other intensity-based scales utilised in depression and anxiety studies.

The findings regarding depression contradict existing evidence that CAT is an effective intervention for depression (Dunn, Golynkina, Ryle & Watson, 1997). However, it is important to note that this evidence emerged from a study utilising 16-

session CAT with an extended follow-up; therefore, intervention characteristics differed significantly from the reviewed studies which were typically shorter in duration without follow-up. However, there is recent evidence that brief 8-session CAT may be appropriate in treating depression (Kellett et al., 2018) that is unsupported by the reviewed studies. No known CAT studies specific to anxiety disorders have been published and thus there is no comparator in the evidence-base for anxiety.

It is important to note that authors may have determined effectiveness differently between studies. Some studies reported overall mixed or ineffective outcomes despite improvements in target problems. Thus, it is important that readers consider outcomes across measures, in addition to overall conclusions. Additionally, the presence of a follow-up may have influenced conclusions around effectiveness. For example, one paper (5) reported a partially effective intervention due to an observed deterioration between the second treatment phase and follow-up. In comparison, another paper (13) reported an effective intervention in a study with no follow-up. Had this study contained a follow-up, outcomes may not have been maintained. Hence, it is important that design characteristics of studies are considered when interpreting results.

The current review includes studies which explore the effectiveness of CAT with diagnoses which have not previously been represented in the evidence base. As described, SCEDs are often used in such circumstances; when a problem of interest has a low base rate and so larger numbers of participants are difficult to locate (Shadish, 2014). New contributions to the CAT evidence base included hypersexuality disorder, for which CAT was effective, and hoarding disorder, for which CAT was deemed to be ineffective overall but was effective for target problems. It is a strength of the present review that a wide range of diagnoses are captured. It must be noted that there are

primary diagnoses not represented in this sample, including psychosis which was explored during a recent case series that did not meet the inclusion criteria for review (Taylor et al., 2018).

Publication

Unpublished studies in the sample were more likely to report ineffective interventions, whilst published studies generally, but not exclusively, reported effectiveness. It may be argued that this is an example of publication bias. However, there are also other factors which offer potential explanations. Firstly, diagnoses generally differed between published and unpublished literature with more complex diagnoses and morbid jealousy being present in published studies; these are diagnoses for which interventions were generally effective. CAT is typically used for more complex diagnoses (Ryle et al., 2014) and as such, positive outcomes may be a result of clinical presentation as opposed to publication. Furthermore, published studies were more frequently conducted by qualified CAT practitioners and this may affect outcomes.

Interventions were generally shorter and without follow-up in unpublished studies. This is likely to be attributable to the grey literature predominantly consisting of SCED reports completed by Trainee Clinical Psychologists as part of clinical training. The nature of clinical training is such that time restrictions impact on treatment durations. It may be that there was less opportunity for change to emerge and be observed at follow-up. Indeed, evidence demonstrates that the more data points that are utilised in SCEDs, the less vulnerability there is to random occurrences influencing results (Rizvi & Nock, 2008). The studies without follow-up are also concerning in terms of fidelity to CAT, given that follow-up is a key feature of the model (Ryle et al., 2014).

Design

The majority of studies included in the review were A/B designs. An A/B design is seen as a quasi-experimental design, as opposed to a true SCED, as there is no use of manipulation (i.e. A/B/A/B) (Kazdin, 1978). In order to ensure outcomes are valid and reliable, experimental manipulation, such as treatment withdrawal or a treatment comparison, is required (Kellett, et al., 2020). Thus, findings should be interpreted with caution due to the large number of quasi-experimental designs used.

All but one study in the sample completed statistical analysis of idiographic data. The use of statistical analysis adds scientific rigour and demonstrates the fundamental difference between SCEDs and case studies which richly describe a case (Cohen et al., 2014). Whilst SCEDs historically depended on visual analysis alone (Parker et al., 2007), the failure to use statistics has been an obstacle to SCEDs being accepted by the wider research community (Shadish, 2014b). Thus, the use of statistical analysis in the majority of studies is promising.

It must be noted, however, that more traditional single-case analyses, such as non-overlap methods including PND, PAND and PEM, have significant limitations (Shadish, 2014b). As such, research has seen a shift to other methods such as Tau effect sizes which display more robustness in controlling for baseline trends (Brossart, Laird & Armstrong, 2018) and regression models which can account for trend and autocorrelation (Shadish, 2014b). Such methods were utilised in some of the more recent publications, demonstrating progression in analysis methods. There were patterns in terms of type of analysis and outcomes, with studies using methods such as ANCOVA more likely to find positive outcomes than studies using nonoverlap

statistics. However, this pattern must be heavily caveated in that published studies generally used more complex analysis whilst unpublished studies more commonly used nonoverlap methods. Thus, it may be that differences in outcome are linked to the use of grey literature as opposed to analysis methods. Given the limitations of single-case analysis methods, it has been highlighted as important for statistical analysis to be interpreted alongside visual analysis (Parker et al., 2007). Indeed, all studies provided graphing of data by phase.

Competency assessment

Of the studies reviewed, only five reported that competency assessment had been completed. In these studies, competency was assessed using the CCAT measure (Bennett & Parry, 2004) which is designed for use with audio-recordings of sessions and scores competence across 10 domains of therapeutic practice. The use of competency assessment ensures that therapeutic principles are adhered to and thus improves the robustness of findings. It is therefore a limitation that competency assessment was not completed in the majority of studies.

Critique

As described, the majority of the studies identified through the grey literature search were single-case reports completed by Trainee Clinical Psychologists. Whilst level of experience was not documented within each paper, some reports described being completed as early as the first year of training. For therapists early in their careers, fidelity to CAT may have been more challenging. As such, this may create an illusion of CAT being less effective for certain diagnoses, such as depression and anxiety. However, it is possible that if interventions were delivered by qualified CAT therapists, outcomes may have differed. Indeed, a recent meta-analysis found evidence that larger

effects were observed when therapists were CAT qualified (Hallam et al., 2020). Hence, interventions are not directly comparable across published and unpublished literature, and thus across diagnoses; results should therefore be interpreted with caution.

There is a reliance on reported information in this review which may skew findings. For example, the limited evidence of competency assessment indicates that fidelity to CAT may have been limited. This brings into question whether outcomes were a result of CAT or interventions in which there was a lot of therapeutic drift. However, as competency assessment was not explicitly discussed in all studies, it is possible that it occurred without being reported. The reliance on reported information is also a limitation of quality appraisal. The SCED scale utilised requires an explicit statement for an item to receive a score of 1 (Tate et al., 2008). A score of 0 therefore means that an item is not explicitly described; it does not mean that it did not occur. Thus, reduced quality scores may be a result of reporting as opposed to methodology.

Indeed, the reporting standards of studies within the review varied. For example, there were studies in the sample that did not specify patient age and several studies which did not describe the intervention setting. The SCRIBE reporting guidelines (Tate et al., 2016), which provide authors with a 26-item checklist of information that should be reported to ensure clarity, accuracy and transparency, identifies this information as crucial. However, it is important to acknowledge that many of the included studies preceded these guidelines. Future CAT research employing single-case methodologies should utilise this checklist to ensure accurate and transparent reporting.

Reporting was also of concern with regard to medication; many studies did not describe whether the patient was medicated. Whilst this is not stipulated in the SCRIBE

guidelines, it remains important in interpreting outcomes. Given that it is unclear how many patients in the sample were receiving psychopharmacological treatment alongside therapy, it is not possible to infer whether outcomes were attributable to CAT alone. This further highlights the need for detailed reporting of clinical information when writing single-case research.

The majority of studies included in the review utilised idiographic scales measuring intensity as opposed to frequency. Whilst there are certain constructs that are more appropriately measured by intensity scales, frequency scales offer a more concrete, quantifiable measurement that is straightforward for patients to report and thus more accurate. Intensity scales, conversely, are more subjective and thus may be less reliable and less sensitive. It is possible that scale type influenced findings; future research should seek to ensure scales are quantifiable and sensitive to change.

Whilst the current review makes an important contribution to the evidence base for the effectiveness of CAT, the use of single-case methodologies does have implications in terms of reliability and validity. It is possible that external events, such as life stressors, may impact on the outcome of a SCED. This is particularly true of A/B designs where there is no reversal phase and thus less evidence that change is attributable to the intervention. However, it is important to consider the ethical implications for patients; whilst withdrawal designs, such as A/B/A/B designs, enhance scientific rigour, the withholding or withdrawing of treatment from individuals with mental health problems poses ethical concerns (Rizvi & Nock, 2008).

Limitations of review

Despite the large number of studies included in the review, the number of participants remains small and subsamples by diagnosis were smaller still. This is an identified limitation of the CAT evidence-base; the variety of presentations CAT is used for means that diagnosis-specific evidence is limited (Hallam et al., 2020). This has implications for generalisability; interventions that were effective or ineffective for individual participants may not have the same outcome for others. Indeed, some diagnoses were only represented by one patient in the sample and as such, findings cannot be generalised. Thus, whilst the review provides an important synthesis of the evidence available from SCEDs representing a variety of clinical presentations, further research is required to consolidate findings.

Clinical implications

A number of studies reported positive outcomes in terms of nomothetic measures but outcomes were not consistent across idiographic measures of target problems. The limitations of certain types of idiographic measures have been discussed. Clinicians measuring outcomes in clinical practice should consider this when designing measures with patients, ensuring that they are accessible and meaningful. Close attention should be paid to the power dynamics of data design and data collection (Stockton & Kellett, 2020) to ensure that the patient is not simply pleasing the therapist by reporting change. Clinicians should also consider the differences in effectiveness between diagnoses and hold this in mind when selecting appropriate interventions.

Single-case methodological implications

The present study provides a systematic review of single-case research which assesses methodological quality; this has been previously highlighted as an area of need (Maggin

et al., 2011). The findings around quality, particularly in unpublished literature where quality was lower, demonstrate the importance of researchers adhering to quality guidelines, such as the SCED scale (Tate et al., 2008). Future studies should ensure accuracy of reporting by consulting the SCRIBE statement (Tate et al., 2016). SCEDs assessing the effectiveness of CAT make a valuable contribution to the evidence-base; further single-case studies should be completed to explore different clinical presentations currently underrepresented in the literature.

The current CAT SCED evidence base is limited by design with regard to the overreliance on simple bi-phasic (A/B) designs. Further research should make use of withdrawal (A/B/A/B) and crossover (ABC) designs and try to ensure idiographic data is collected over a follow-up period. This will better index the durability of change in target problems. Additionally, collecting outcomes from a person other than the patient would increase the internal validity of studies.

Theoretical implications

The present review is an important first step in exploring the idiographic effectiveness of CAT with multiple clinical presentations. Initial promising findings emerged for the effectiveness of CAT with certain diagnoses including personality disorders, morbid jealousy and hypersexuality disorder. Larger scale group studies are now warranted to consolidate findings with specific diagnoses. Further research is also needed to consolidate findings around the single-case evidence-base. Research could expand on the present review by quantifying the effects of CAT for patients' target problems with a meta-analysis.

Despite the PSQ being central to the CAT assessment process (Ryle & Kerr, 2002), the measure was only used in three studies. This suggests a disparity between measures used in routine clinical practice and those selected for SCEDs. One explanation for the PSQ not being routinely selected may be that validation attempts have lagged behind clinical uptake (Berrios, Kellett, Fiorani & Poggioli, 2016). Thus, further research validating the PSQ is indicated to improve evidence for its use in clinical practice.

Conclusions

This systematic review provided a thorough review of the quality and findings of single-case literature exploring the effectiveness of CAT in treating patients' target problems. Study quality varied between published and unpublished studies, with published papers displaying higher quality. Published studies were also more likely to report positive outcomes than studies in the grey literature. The effectiveness of CAT appears to differ depending on the diagnosis of the patient. These findings are caveated, with limitations discussed. The review has implications for clinicians in terms of designing idiographic measures and selecting appropriate interventions depending on clinical presentation. Further, larger-scale research is required to consolidate findings.

References

Barkham, M., Bewick, B., Mullin, T., Gilbody, S., Connell, J., Cahill, J., . . . Evans, C. (2013). The CORE-10: a short measure of psychological distress for routine use in the psychological therapies. Counselling and Psychotherapy Research, 13, 3-13. https://doi.org/10.1080/14733145.2012.729069

Barkham, M., Evans, C., Margison, F., Mcgrath, G., Mellor-Clark, J., Milne, D., & Connell, J. (1998). The rationale for developing and implementing core outcome batteries for routine use in service settings and psychotherapy outcome research. *Journal of Mental Health, 7,* 35-47. Doi: 10.1080/09638239818328

Barkham, M., Hardy, G., & Startup, M. (1996). The IIP-32: a short version of the inventory of interpersonal problems. *British Journal of Clinical Psychology, 35,* 21-35. Doi: 10.1111/j.2044-8260.1996.tb01159.x.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II.* San Antonio, Texas: Psychological Corporation.

Beckett, R., Tarrier, N., Intili, R., & Beeck, A. (1992). *The Prestwich jealousy questionnaire.* Unpublished document: Manchester University.

Bennett, D., & Parry, G. (2004). A measure of psychotherapeutic competence derived from cognitive analytic therapy. *Psychotherapy Research, 14,* 176-192. Doi: 10.1093/ptr/kph016

Berrios, R., Kellett, S., Fiorani, C., & Poggioli, M. (2016). Assessment of identity disturbance: factor structure and validation of the personality structure

questionnaire in an Italian sample. *Psychological Assessment, 28,* 27-35. Doi: 10.1037/pas0000143

Brossart, D. F., Laird, V. C., & Armstrong, T. W. (2018). Interpreting Kendall's Tau and Tau-U for single-case experimental designs. *Cogent Psychology, 5*(1)*,* 1-26. Doi: 10.1080/23311908.2018.1518687

Calvert, R., & Kellett, S. (2014). Cognitive analytic therapy: a review of the outcome evidence base for treatment. *Psychologic and Psychotherapy: Theory, Research and Practice, 87,* 253-277. Doi: 10.1111/papt.12020

Carradice, A. (2004). Applying cognitive analytic therapy to guide indirect working. *Reformulation,* 23, 18-23.

Chanen, A. M., Jackson, H. J., Mccutcheon, L. K., Jovev, M., Dudgeon, P. Yuen, H. P., . . . Mcgorry, P. D. (2008). Early intervention for adolescents with borderline personality disorder using cognitive analytic therapy: randomized controlled trial. *British Journal of Psychiatry, 193,* 477-484. Doi: 10.1192/bjp.bp.107.048934

Clarke, S., Thomas, P., & James, K. (2013). Cognitive analytic therapy for personality disorder: randomized controlled trial. *The British Journal of Psychiatry, 202,* 129-134. Doi: 10.1192/bjp.bp.112.108670

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20,* 37-46. Doi: 10.1177/001316446002000104

Cohen, L. L., Feinstein, A., Masuda, A., & Vowles, K. E. (2014). Single-case research design in pediatric psychology: considerations regarding data analysis. *Journal of Pediatric Psychology, 39,* 124-137. Doi: 10.1093/jpepsy/jst065

Collins, C. (2011). *Effectiveness of cognitive analytic therapy with a woman with borderline personality disorder: a single case experimental design.* Unpublished manuscript.

Crosby, K. (2015). *The effectiveness of cognitive analytic therapy in treating generalized anxiety disorder: a single case design.* Unpublished manuscript.

Curling, L., Kellett, S., & Totterdell, P. (2018). Cognitive analytic therapy for obsessive morbid jealousy: a case series. *Journal of Psychotherapy Integration, 28*, 537-555. Doi: 10.1037/int0000122

Curling, L., Kellett, S., Totterdell, P., Parry, G., Hardy, G., & Berry, K. (2018). Treatment of obsessive morbid jealousy with cognitive analytic therapy: an adjudicated hermeneutic single-case efficacy design evaluation. *Psychology and Psychotherapy: Theory, Research and Practice, 91*, 95-116. Doi: 10.1111/papt.12151

Derogatis, L. R. (1993). *Brief symptom inventory (BSI): Administration, scoring and procedures manual (3rd edition).* Minneapolis: NSC Pearson

Dunn, M., Golynkina, K., Ryle, A., & Watson, J. P. (1997). A repeat audit of the cognitive analytic therapy clinic at Guy's hospital. *Psychiatric Bulletin, 21,* 165-168. Doi: 10.1192/pb.21.3.165

Elliott, R. (2010). Hermeneutic single-case efficacy design. *Psychotherapy Research, 12,* 1-12. Doi: 10.1080/713869614

Evans, M., Kellett, S., Heyland, S., Hall, J., & Majid, S. (2017). Cognitive analytic therapy for bipolar disorder: a pilot randomized controlled trial. *Clinical Psychology and Psychotherapy, 24,* 22-35. Doi: 10.1002/cpp.2065

Fallon, L. M., Collier-Meek, M. A., Maggin, D. M., Sanetti, L. M., & Johnson, A. H. (2015). Is performance feedback for educators an evidence-based practice? A systematic review and evaluation based on single-case research. *Exceptional Children, 81,* 227-246. Doi: 10.1177/0014402914551738

Field, E. (2003). *A single case experimental design assessing the efficacy of cognitive analytic therapy in treating obsessive compulsive disorder.* Unpublished manuscript.

Hall, J. (2013). *The use of a brief version of cognitive analytic therapy (bCAT) with a working aged woman with bulimia nervosa, alcohol problems, and a depressive episode: evaluating the effectiveness of bCAT.* Unpublished manuscript.

Hallam, C., Simmonds-Buckley, M., Kellett, S., Greenhill, B., & Jones, A. (2020). The acceptability, effectiveness, and durability of cognitive analytic therapy:

systematic review and meta-analysis. *Psychology and Psychotherapy: Theory, Research and practice, 12286.* Doi: 10.1111/papt.12286

Harvey, K. (2011). *Treating a complex presentation of depression with cognitive analytic therapy: a single case study.* Unpublished manuscript.

Heathcote, C. (2019). *Utilising a CAT-informed intervention with a client with emotion regulation and interpersonal difficulties.* Unpublished manuscript.

Hopewell, S., McDonald, S., Clarke, M., & Egger, M. (2007). Grey literature in meta-analyses of randomized trials of health care interventions. *Cochrane Database of Systematic Reviews 2007.* Doi: 10.1002/14651858.MR000010.pub3

Kazdin, A. E. (1978). Methodological and interpretive problems of single-case experimental designs. *Journal of Consulting and Clinical Psychology, 46,* 629-642. Doi: 10.1037/0022-006X.46.4.629

Kellett, S. (2005). The treatment of dissociative identity disorder with cognitive analytic therapy: experimental evidence of sudden gains. *Journal of Trauma and Dissociation, 6,* 55-81. Doi: 10.1300/j229v06n03_03

Kellett, S. (2007). A time series evaluation of the treatment of histrionic personality disorder with cognitive analytic therapy. *Psychology and Psychotherapy, 80,* 389-405. Doi: 10.1348/147608306x161421

Kellett, S., Alhadeff, L., Gaskell, C., & Simmonds-Buckley, M. (2020). *Evaluating the effectiveness of cognitive analytic therapy for bipolar disorder: co-production of a single case base experimental design.* Unpublished manuscript.

Kellett, S., Bennett, D., Ryle, A., & Thake, A. (2013). Cognitive analytic therapy for borderline personality disorder: therapist competence and therapeutic effectiveness in routine practice. *Clinical Psychology and Psychotherapy, 20,* 216-225. Doi: 10.1002/cpp.796

Kellett, S., Gausden, J., & Gaskell, C. (2020). The effectiveness of cognitive analytic therapy for borderline personality disorder: utilizing a withdrawal experimental design to improve sensitivity to abandonment. *Psychology and Psychotherapy: Theory, Research and Practice, 75*. Doi: 10.1111/papt.12278

Kellett, S., Ghag, J., Ackroyd, K., Freshwater, K., Finch, J., Freear, A. . . Simmonds-Buckley, M. (2019). Delivering cognitive analytic consultancy to community mental health teams: initial practice-based evidence from a multi-site evaluation. *Psychology and Psychotherapy, 93,* 429-455. Doi: 10.1111/papt.12221

Kellett, S., & Hardy, G. (2014). Treatment of paranoid personality disorder with cognitive analytic therapy: a mixed methods single case experimental design. *Clinical Psychology and Psychotherapy, 21,* 452-464. Doi: 10.1002/cpp.1845.

Kellett, S., & Lees, S. (2019). Quasi-experimental N=1 evaluation of the effectiveness of cognitive analytic therapy for dependent personality disorder. *Journal of Psychotherapy Integration, 30,* 458-475. Doi: 10.1037/int0000170

Kellett, S., Simmonds-Buckley, M., & Totterdell, P. (2017). Testing the effectiveness of cognitive analytic therapy for hypersexuality disorder: an intensive time-series evaluation. *Journal of Sex and Marital Therapy, 43,* 501-516. Doi: 10.1080/0092623X.2016.1208129

Kellett, S. & Stockton, D. (2020). *Can the green-eyed monster be tamed quickly? Treatment of obsessive morbid jealousy with brief cognitive analytic therapy.* Manuscript submitted for publication.

Kellett, S., Stockton, C., Marshall, H., Hall, J., Jennings, C., & Delgadillo, J. (2018). Efficacy of narrative reformulation during cognitive analytic therapy for depression: randomized dismantling trial. *Journal of Affective Disorders, 239,* 37-47. Doi: 10.1016/j.jad.2018.05.070

Kellett, S., & Totterdell, P. (2013). Taming the green-eyed monster: temporal responsivity to cognitive behavioural and cognitive analytic therapy for morbid jealousy. *Psychology and Psychotherapy, 86,* 52-69. Doi: 10.1111/j.2044-8341.2011.02045.x

Llewelyn, S. (2003). Cognitive analytic therapy: time and process. *Psychodynamic Practice, 9,* 501-520. Doi: 10.1080/13533330310001616759

Lobo, M. A., Moeyaert, M., Baraldi Cunha, A., & Babik, I. (2017). Single-case design, analysis, and quality assessment for intervention research. *Journal of Neurologic Physical Therapy, 41,* 187-197. Doi: 10.1097/NPT.0000000000000187

Ma, H. (2006). An alternative method for quantitative synthesis of single-subject researchers: percentage of data points exceeding the median. *Behaviour Modification*, 30, 598-617. Doi: 10.1177/0145445504272974

Maggin, D. M., & Odom, S. L. (2014). Evaluating single-case research data for systematic review: a commentary for the special issue. *Journal of School Psychology, 52,* 237-241. Doi: 10.1016/j.jsp.2014.01.002

Maggin, D. M., O'Keeffe, B. V., & Johnson, A. H. (2011). A quantitative synthesis of methodology in the meta-analysis of single-subject research for students with disabilities: 1985-2009. *Exceptionality, 19,* 109-135. Doi: 10.1080/09362835.2011.565725

Margison, F. (2000). Cognitive analytic therapy: a case study in treatment development. *British Journal of Medical Psychology, 73,* 145-150. Doi: 10.1348/000711200160354

Marriott, M., & Kellett, S. (2009). Evaluating a cognitive analytic therapy service; practice-based outcomes and comparisons with person-centred cognitive-behavioural therapies. *Psychology and Psychotherapy, 82,* 57-72. Doi: 10.1348/147608308x336100

Morley, S. (2018). *Single case methods in clinical psychology: A practical guide.* Abingdon, UK: Routledge.

Parker, R. I., Hagan-Burke, S., & Vannest, K. (2007). Percentage of all non-overlapping data (PAND): An alternative to PND. The Journal of Special Education, *40*, 194-204. Doi: 10.1177/00224669070400040101

Pollock, P. H., Broadbent, M., Clarke, S., Dorrian, A., & Ryle, A. (2001). The personality structure questionnaire (PSQ): a measure of the multiple self-states model of identity disturbance in cognitive analytic therapy. *Clinical Psychology and Psychotherapy, 8,* 59-72. doi:10.1002/cpp.250

Power, N. (2019). *Is 15 sessions of cognitive analytic therapy effective in the treatment of an individual with an emotionally unstable personality disorder presentation? A single case experimental design.* Unpublished manuscript.

Rayner, K., Thompson, A. R., & Walsh, S. (2011). Clients' experience of the process of change in cognitive analytic therapy. *Psychology and Psychotherapy: Theory, Research and Practice, 84,* 299-313. Doi: 10.1348/147608310X531164

Richardson, W. S., Wilson, M. C., Nishikawa, J., & Hayward, R. S. (1995). The well-built clinical question: a key to evidence-based decisions. *ACP Journal Club, 123,* 12-13. Doi: 10.7326/ACPJC-1995-123-3-A12

Rizvi, S. L., & Nock, M. K. (2008). Single-case experimental designs for the evaluation of treatments for self-injurious and suicidal behaviors. *Suicide and Life-Threatening Behavior, 38,* 498-510. doi:10.1521/suli.2008.38.5.498

Ryder, A. (2010). *Using a cognitive analytic therapy approach with an adult presenting with chronic anxiety and depression: a single case experimental study.* Unpublished manuscript.

Ryle, A. (1995). *Cognitive analytic therapy: Developments in theory and practice.* Chichester, UK: John Wiley.

Ryle, A., & Fawkes, L., (2007). Multiplicity of selves and others: cognitive analytic therapy. *Journal of Clinical Psychology, 63,* 165-174. Doi: 10.1002/jclp.20339

Ryle, A., Kellett, S., Hepple, J., & Calvert, R. (2014). Cognitive analytic therapy at 30. *Advances in Psychiatric Treatment, 20,* 258-268. Doi: 10.1192/apt.bp.113.011817

Ryle, A., & Kerr, I. (2002). *Introducing cognitive analytic therapy: principles and practice.* Chichester, UK: John Wiley.

Sanson-Fisher, R. W., Bonevski, B., Green, L. W., & D'este, C. (2007). Limitations of the randomized controlled trial in evaluating population-based health interventions. *American Journal of Preventive Medicine, 33,* 155-161. Doi: 10.1016/j.amepre.2007.04.007

Scruggs, T. E., Mastropieri, M.A., & Casto, G. (1987). The qualitative synthesis of single- subject research. Remedial and Special Education, 8, 24-33. Doi: 10.1177/074193258700800206

Shadish, W. R. (2014). Analysis and meta-analysis of single-case designs: an introduction. *Journal of School Psychology, 52,* 109-122. Doi: 10.1016/j.jsp.2013.11.009

Shadish, W. R. (2014b). Statistical analyses of single-case designs: the shape of things to come. *Current Directions in Psychological Science, 23,* 139-146. Doi: 10.1177/0963721414524773

Sherman, N. (2014). *The effectiveness of cognitive analytic therapy for a 25 year old male with depression.* Unpublished manuscript.

Sims, J. (2015). *Does 16 session cognitive analytic-informed therapy improve outcome measures for confidence in ability to cope, self-critical thinking, and engagement in meaningful activity? A single case experimental design.* Unpublished manuscript.

Spence, C., Kellett, S., Totterdell, P., & Parry, G. (2019). Can cognitive analytic therapy treat hoarding disorder? An adjudicated hermeneutic single-case efficacy design evaluation. *Clinical Psychology and Psychotherapy, 26*, 673-683. Doi: 10.1002/cpp.2390

Swan, H. (2009). *The use of reformulation in cognitive analytic therapy with a woman diagnosed with bulimia nervosa: evaluating the effectiveness of reformulation.* Unpublished manuscript.

Tate, R. L., Mcdonald, S., Perdices, M., Togher, L., Schultz, R., & Savage, S. (2008). Rating the methodological quality of single-subject designs and *n*-of-1 trials: introducing the single-case experimental design (SCED) scale. *Neuropsychological Rehabilitation, 18,* 385-401. Doi: 10.1080/09602010802009201

Tate, R. L., Perdices, M., Rosenkoetter, U., Shadish, W., Vohra, S., Barlow, D. H., . . . Wilson, B. (2016). The single-case reporting guideline in behavioural interventions (SCRIBE) 2016 statement. *Journal of Clinical Epidemiology, 73,* 142-152. Doi: 10.1016/j.jclinepi.2016.04.006

Tate, R. L., Perdices, M., & Wakim, D. (2020). Integrating routine clinical interventions with single-case methodology: parallels, differences and bridging strategies. *Brain Impairment, 21,* 99-109. Doi: 10.1017/BrImp.2019.8

Taylor, P. J., Jones, S., Huntley, C. D., & Seddon, C. (2017). What are the key elements of cognitive analytic therapy for psychosis? A delphi study. *Psychology and Psychotherapy: Theory, Research and Practice, 90,* 511-529. Doi: 10.1111/papt.12119

Taylor, P. J., Perry, A., Hutton, P., Tan, R., Fisher, N., Focone, C. . . . Seddon, C. (2018). Cognitive analytic therapy for psychosis: a case series. *Psychology and Psychotherapy, 92,* 359-378. Doi: 10.1111/papt.12183

Taylor, P. J., Perry, A., Hutton, P., Seddon, C., & Tan, R. (2015). Curiosity and the CAT: considering cognitive analytic therapy as an intervention for psychosis. *Psychosis, 7,* 276-278. Doi: 10.1080/17522439.2014.956785

Todman, J. B., & Dugard, P. (2001). *Single-case and small-n experimental designs: A practical guide to randomization tests.* New York: Psychology Press.

Turpin, G. (2001). Single case methodology and psychotherapy evaluation: From research to practice. In C. Mace, S. Moorey & B. Roberts (Eds.), *Evidence in the psychological therapies: a critical guide for practitioners* (pp. 91-113). New York: Brunner-Routledge.

Walsh, S. (2014). *A single case analysis of 8-session cognitive analytic therapy.* Unpublished manuscript.

Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin, 130,* 631-663. Doi: 10.1037/0033- 2909.130.4.631

Appendices

Appendix A – Data Extraction Tool

Authors	
Publication Year	
Study Design	
Patient Demographics (Age, Gender)	
Clinical Description of client	
Intervention Context	
Diagnosis and Target Problem	
Intervention characteristics (length baseline, length intervention, follow-up)	
Idiographic Measurement (frequency, no. of scales)	
Nomothetic Measurement (no. of measures, type, frequency)	
Type of nomothetic analysis	
Type of idiographic analysis	
Autocorrelation analysis	
Graphing of idiographic outcomes by phase of study	
Competency assessment	
Qualification of therapist	
Gender of therapist	
Supervision support	
Previous treatment history of client	
Medication	
Generalised outcomes	

Appendix B – Quality Appraisal Ratings

Study[6]	Clinical History?	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	Total Score
1												7
2												9
3												6
4												7
5												8
6												6
7												7
8												6
9												9
10												9
11												5
12												5
13												5
14												5
15												5
16												5
17												8
18												8
19												7
20												5
21												5
22												5
23												6
24												5

Key: Present Yes (1) No (0)

[6] Study ID as per Table 1 in-text

Appendix C – PRISMA Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	8
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	9
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	2, 9
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	9-10
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	9
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	9
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	10, 14
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	11

Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	11
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	n/a
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	11
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	12
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	11
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	12
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	14
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	15-22
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	n/a
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	18-22
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	n/a
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	13
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	28-31
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	32-42
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	37-40

Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	32-42
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	n/a

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

Part II

Research Report

A Network Analysis of the Personality Structure Questionnaire (PSQ)

Abstract

Objectives

Despite the relevance of identity disturbance to psychopathology, it is an under-researched area. The Personality Structure Questionnaire (PSQ) is a reliable and valid measure of identity disturbance, but further research into the measure has been recommended. This study sought to complete a network analysis of the PSQ in a multi-site, cross-national study.

Method

Secondary data (n = 1549) was utilised to explore the network structure of the PSQ and examine whether networks differed across four subsample comparisons: UK (n = 625) versus Italy (n = 521), adults (n = 521) versus adolescents (n = 254), clinical (n = 769) versus community (n = 780) and complex (n = 477) versus other diagnoses (n = 1073). *R* was utilised to estimate a global network and subsamples were compared using jointly estimated networks. Stability analyses were conducted to assess robustness of networks.

Results

Network estimation of the full sample showed that the most central items were PSQ3 (stable sense of self) and PSQ5 (changing mood) in terms of standardised node strength and predictability. Network structures significantly differed between the UK and Italy. Networks did not differ by age, sample type or diagnosis. Centrality of PSQ items was largely consistent between subsamples.

Conclusions

The study provides additional evidence of the psychometric foundations of the PSQ. The networks of identity disturbance presented potentially enable clinicians to target symptoms to enable personality integration. Further research is required to explore networks in and between other cultures.

Practitioner Points

1) The findings around the most central items of the PSQ may inform intervention for those presenting with identity disturbance. Specifically, interventions that target changing mood and aim to create a more stable sense of self are likely to be effective.
2) Clinicians can utilise these findings to enhance assessment of identity disturbance.
3) Given the brevity of the PSQ, it appears suitable for use in everyday clinical practice and on a session-by-session basis.

Limitations

- Only Western European countries with potential cultural similarities were compared in terms of nationality.
- Subsamples were not always comparable in terms of sample size.

Keywords: Identity Disturbance, Network Analysis, PSQ

Introduction

What is identity?

There are various theories of identity and there is no one definition of identity that is universally agreed upon. Terms such as identity, ego, the self and personality have often been used interchangeably, leading to differing interpretations (Jorgensen, 2006). Erikson, a major identity theorist, described that the term 'identity' has multiple different but closely connected meanings (Erikson, 1968). On one hand, identity is a feature of an individual that involves internal self-definition, whilst on the other hand, identity is constructed in a social context and is shaped by wider culture (Deaux, 2000). There is often a differentiation between social identity, which is derived from membership of social groups, culture or society, and personal identity, which is an individual's concept of themselves as a unique person with defining characteristics, traits and individual needs (Jorgensen, 2006).

The consolidation of identity is a central developmental task within adolescence (Erikson, 1968) and adolescents may display a normative 'identity crisis' as part of this process (Kernberg, 2006). Identity consolidation involves an individual experiencing themselves as consistent over time and contexts, displaying stable attitudes and values, and having long-term goals. In contrast, identity confusion involves incoherence or confusion about who one is (Westen, Betan & DeFife, 2011). Without an adequately formed identity, an individual's cognitive, affective, behavioural and interpersonal functioning may be impaired (Taylor & Goritsas, 1994). Thus, identity is important to psychological health (Jorgensen, 2006). Despite the relevance of identity to psychopathology, however, identity continues to receive less attention in comparison to many clinical constructs (Kaufman, Cundiff & Crowell, 2015).

Identity disturbance and its relationship to psychopathology

Identity disturbance is believed to be caused by the inability to develop and consolidate one's sense of identity (Erikson, 1968). Identity disturbance is characterised by an unstable sense of self that is superficial, lacking in complexity and polarised (Horz-Sagstetter et al., 2018), as well as the occurrence of rapid shifts between differentiated states of mind that is created and maintained through dissociation (Ryle & Kerr, 2002). Identity disturbance is identified as a key characteristic of personality disorders (Kernberg, 2006) and is one of the diagnostic criteria for Borderline Personality Disorder (BPD) (Wilkinson-Ryan & Westen, 2000). Identity disturbance in this clinical population has been said to reflect an inability to integrate positive and negative self-representations, leading to a shifting view of the self, rapidly shifting roles (e.g. dominant to submissive) and feelings of emptiness (Wilkinson-Ryan & Westen, 2000). The most pronounced form of identity disturbance is considered to be dissociative identity disorder (DID) (Modestin, Oberson & Erni, 1998). DID is a disorder defined by the presence of more than one personality state; individuals with DID exhibit multiple selves (Boysen & VanBergen, 2013).

Whilst identity disturbance is one of the diagnostic criteria for BPD, there is evidence it is not exclusive to this clinical population. Identity disturbance has been found to be higher amongst patients with varying diagnoses when compared with controls (Neacsiu, Herr, Fang, Rodriguez & Rosenthal, 2015). Further, this study found that emotional dysregulation was significantly related to identity disturbance, with anxiety and depression severity being significant predictors. It has recently been argued that identity disturbance, along with feelings of emptiness, are also prominent in schizophrenia; hence, it has been proposed that identity disturbance should be included

within definitions of the schizophrenia spectrum to ensure that BPD is not always the assumed diagnosis (Zandersen & Parnas, 2019).

Cognitive analytic therapy (CAT) has been identified as a promising area of psychotherapy in terms of its alternative conceptualisation of personality pathology and its effectiveness in treating and managing personality disorders (Pollock, Broadbent, Clarke, Dorrian & Ryle, 2001). Within CAT, there is an emphasis on reformulating problems relating to an individual's identity formation and coherence. The multiple self-states model (MSSM; Ryle, 1997) is a CAT-specific model which provides a reformulation of identity disturbances and guides therapy (Pollock et al., 2001). The MSSM conceptualises disturbances in personality from healthy identity development to extreme identity disturbance, including DID (Bedford, Davies & Tibbles, 2009).

Within CAT, there is a focus on reciprocal role procedures (RRPs) which are learned self-other interactions. The MSSM outlines three levels of increasing identity disturbance in the context of these RRPs (Pollock et al., 2001). At the first level, the nature of an individual's RRPs may be impacted on by past adversity or trauma, leading to restricted flexibility in interactions. At level two, an individual may have a fragmented experience of the self, which leads to the presence of two of more discrete self-states; self-states refer to a dominant RRP repeatedly displayed by an individual. At the third level, an individual has a deficient capacity to self-reflect or self-observe, causing them to act in a maladaptive way due to state-switching underpinned by dissociation (Pollock et al., 2001). This state-switching has been stated to overwhelm and confuse clinicians and lead to overly narrow interventions (Ryle, 1997). Whilst the MSSM approach to identity disturbance is of particular relevance to the current research, it must be acknowledged that the MSSM is specific to a CAT

conceptualisation of identity disturbance, with a focus on interpersonal interactions. Thus, it must be highlighted that it is a limitation of the MSSM that its CAT-specific approach does not incorporate other theories of identity and identity disturbance.

Measurement of identity disturbance

Despite its prominence in psychopathology, identity disturbance is under-researched (Adler, Chin, Kolisetty & Oltmanns, 2012). Adequate measures of identity and identity disturbance are important for both research and clinical practice. However, many available measures are not suited for clinical use; measures are too brief (e.g. a single question embedded in a wider clinical measure) or too lengthy, requiring clinical interviews (Kaufman et al., 2015). An example of the latter is the Identity Disturbance Questionnaire (IDQ); a questionnaire consisting of 35 indicators of identity disturbance, completed by an observer based on a clinical interview or observations (Wilkinson-Ryan & Westen, 2000). Other measures of identity disturbance are specific to adolescents, such as the Identity Disturbance Questionnaire – Adolescents (IDQ-A; Westen et al., 2011) and the Assessment of Identity Development in Adolescence (AIDA; Goth et al. 2012). The 30-item Self-Concept and Identity Measure (SCIM), which was developed with the limitations of other measures in mind, measures dimensions of identity, both healthy and disturbed (Kaufman et al., 2015).

The Personality Structure Questionnaire

The Personality Structure Questionnaire (PSQ; Pollock et al., 2001) was constructed after Ryle (1995) proposed the development of a measure of identity disturbance for clinical use. The PSQ is a brief 8-item measure of identity disturbance that is theoretically grounded in the MSSM of CAT (Pollock et al., 2001). Development of the PSQ was based on clinical observation of identity disturbance in relation to presence of

differing self-states, changeability in moods and behavioural loss of control (Pollock et al., 2001). Three validation studies of the PSQ have been conducted utilising adult samples. The first study (Pollock et al., 2001) utilised clinical and community samples and found that the test-rest ($r = 0.75$) and coefficient alphas ($\alpha = 0.59$ to 0.87) were satisfactory. The clinical sample scored consistently higher on the PSQ when compared with the community sample. The PSQ was also found to discriminate between diagnostic groups. Furthermore, an exploratory factor analysis demonstrated that the PSQ factored into a single scale.

In the second study, Bedford et al. (2009) explored the psychometric properties of the PSQ with a large clinical sample. Examination of inter-correlation coefficients identified all item correlations to be statistically significant and the test-retest ($p < 0.00$) and coefficient alphas ($\alpha = 0.87$) were satisfactory. Further, an exploratory factor analysis again identified a single unifactorial scale and found all factor loadings to be highly satisfactory. Participants' scores on the PSQ remained stable during the waiting list but reduced following intervention, demonstrating that the measure was sensitive to change.

In the third study, Berrios, Kellett, Fiorani & Poggioli, (2016) conducted a confirmatory factor analysis of the PSQ in an Italian sample. A three-factor structure (differing self-states, mood variability and behavioural loss of control) was identified. All factor loadings were significant and the factor structure was invariant regardless of gender or clinical diagnosis. The findings demonstrated that the PSQ has transcultural stability and reliability. The study also evidenced clinical validity of the PSQ in that meeting diagnostic criteria for a mental health disorder predicted total PSQ scores. The highest PSQ scores were found in those with personality disorders, in line with the

evidence base around identity disturbance in this clinical population. Thus, it is suggested that the PSQ may be useful in distinguishing individuals classified with particular diagnoses, with a score of 26-28 having been identified as an appropriate cut-off (Berrios et al., 2016).

The network approach

The network approach to psychopathology has received increasing recognition in recent years (Fried et al., 2018). The network approach understands mental health disorders as complex networks of interacting symptoms (Fried, Epskamp, Nesse, Tuerlinckx & Borsboom, 2016). The approach challenges more traditional approaches to research, such as the common cause approach, which searches for underlying causes of overt symptoms (Borsboom, 2017). Instead, a presentation such as depression can be explained as problems interacting in a dynamic system (Fried et al., 2018).

The integral idea behind the network approach is that one symptom arising can cause other symptoms to arise (Bringmann et al., 2013). For example, if an individual believes others can read their mind (delusions) they may experience extreme suspicion (paranoia) which may lead to the avoidance of others (social isolation). Thus, symptoms are causing and maintaining each other as opposed to being effects of a common cause (Borsboom, 2017). The insights offered by the network approach have led to the development of psychometric models called 'network models' (Bringmann et al., 2013). These are constellations of interacting symptoms that are clustered together (Borsboom & Cramer, 2013). Network models are largely data-driven and exploratory in nature (Fried et al., 2018). Within network models, symptoms are represented as 'nodes' and their connections, known as 'edges', represent their pairwise relationships (Richetin, Preti, Costantini & De Panfilis, 2017).

The network approach provides promising direction towards improving prevention and intervention in clinical settings by investigating which symptoms are more strongly connected or 'central' (Fried et al., 2017). Centrality, a key concept of the network approach, quantifies how closely interconnected a symptom is with other symptoms within a network (Fried et al., 2016). Nodes which are found to be central are identified as having the greatest influence on the development and maintenance of a disorder (Robinaugh, Millner & McNally, 2016). Network analyses of clinical measures enable researchers to determine which items on the measure are most central to the condition of interest due to their relationships with other items (Stochl et al., 2019). Identifying that some nodes are more central is beneficial in clinical practice as the identification of core features of a disorder can firstly support diagnostic processes and, secondly, enhance understanding of the development and maintenance of disorders by offering insight into which symptoms can trigger others (Richetin et al., 2017).

Despite the recent growth in the network analysis evidence base, many studies are limited by single samples or use of non-clinical samples, meaning that there is uncertainty about how networks are structured in clinical data and how networks replicate across datasets (Fried et al., 2018). As a result, concerns around the generalisability and stability of network models across samples have been raised (Forbes, Wright, Markon & Krueger, 2017) and further cross-sample network analyses have been called for (Fried et al., 2018).

Network models and psychiatric diagnoses

The network approach has been used to explore a range of diagnoses including depression (Fried et al., 2016), anxiety (Beard et al., 2016), psychosis (Isvoranu,

Guloksuz, Epskamp, van Os & Borsboom, 2019) and post-traumatic stress disorder (Fried et al., 2018). The network analysis literature specific to identity disturbance, however, is in its infancy. A wider network analysis investigating BPD identified that identity disturbance plays a crucial role in the structure of BPD-related psychopathology in both clinical and non-clinical samples (Richetin et al., 2017). However, a large-scale network analysis later found identity disturbance to be central in only a 'low BPD' group (Southward & Cheavens, 2018). The latter study utilised a measure of emotion regulation and it has therefore been argued that, in order to address the inconsistency between studies, future research utilising a validated measure of identity disturbance is required (Southward & Cheavens, 2018).

The present study

Whilst the literature around network approaches to psychopathology is growing, clear gaps in the evidence remain. The current study sought to apply the network approach to the PSQ in order to explore the structure of identity disturbance in more detail. A cross-sample, cross-national network analysis was used to improve generalisability. The study had the clinical objective of identifying networks of identity disturbance in order to highlight any particular nodes that might then be targets for intervention.

Aims

The present study aimed to:

1) Estimate the overall network structure of identity disturbance measured by the PSQ and explore item connectivity using measures of centrality.
2) Compare whether networks differ in terms of network structure and centrality metrics across four sets of subsamples:
    a) UK versus Italian subsamples.

b) Adult versus adolescent subsamples.

c) Clinical versus community subsamples.

d) Complex diagnosis versus other diagnosis subsamples.

3) Assess the accuracy and stability of the networks using bootstrapping methods.

Given that network approaches are exploratory and data-driven, a priori hypotheses were not proposed.

Method

Ethical considerations

Ethical approval was granted by the University of Sheffield's Department of Psychology research ethics committee (see Appendix A). Collaborators providing secondary data for this research sought ethical approval from their relevant ethical review boards at the time of data collection and consent for data to be used for secondary analysis was obtained. All data was anonymised prior to being received by the researcher.

Sample

The sample (n = 1553) consisted of secondary data from multiple sites. The six data sets are described in Appendix B. The final sample consisted of data from both adolescents, under 18 years, and adults, aged 18+. Participants were aged between 12-70 with a mean age of 28.52[7] (SD = 12.81).

There is no conclusive guidance on sample size, and thus statistical power, within network analysis (Thorlund & Mills, 2012). The required sample size is

[7] Calculated from a sample of n = 1343 where exact ages, as opposed to age ranges, were available.

dependent on the number of parameters that need to be estimated in the network based on the number of nodes [k]; parameters = k x (k-1/2 x k) (Epskamp, Borsboom & Fried, 2018). Thus, in this case the number of parameters estimated based on the eight-item PSQ is 36 (8 x 7/2 +8). The guidelines regarding how many participants are needed per parameter are not clear, however a rule of thumb of three participants per parameter has been suggested (Fried & Cramer, 2017). This criteria was met by the subsamples within this study.

Measure

The PSQ (Appendix C) is an eight-item measure in which self-rated responses are scored between 1-5 (Pollock et al., 2001). Total scores are therefore between 8-40. Higher scores on the PSQ indicate higher levels of identity disturbance.

The data consisted of participants' item-level responses for the eight items of the PSQ. Total PSQ scores for each participant were also calculated. A professionally translated Italian version of the PSQ was utilised by the primary researchers who provided Italian subsample data.

Missing data

Four participants in the overall dataset had missing values for one or more PSQ items and thus total PSQ scores. Network analysis is intolerant of missing data; the absence of a single node can lead to a significantly different understanding of the network and can result in misleading measurements, such as centrality (Borgatti, Carley & Krackhardt, 2006). Thus, only participants with complete data (n = 1549) were included in the analysis.

Data analysis

SPSS Statistics version 26.0.0.0 was used to obtain descriptive statistics and to complete independent-samples t-tests comparing total PSQ scores across the four comparison groups: country, age, sample type and diagnosis. All further analyses were conducted in R version 4.0.2 and RStudio 1.3.1056. The R package *qgraph* (Epskamp, Cramer, Waldorp, Schmittmann & Borsboom, 2012) was utilised to generate networks. A network consists of nodes and edges; in this case, nodes represent PSQ items and edges the statistical relationship between them. Thicker edges represent stronger relationships between nodes.

Firstly, a network was estimated utilising the full sample (n = 1549[8]). Secondly, jointly estimated networks were estimated to make comparisons across subsamples. Comparisons were made between the UK (n = 625) and Italy (n = 521), adults (n = 521) and adolescents (n = 254), clinical (n = 769) and community (n = 780), and complex diagnoses (n = 477) versus other diagnoses (n = 1073). Participants were defined as having a complex diagnosis if their primary presenting problem was recorded as self-harm, personality disorder, psychosis or an eating disorder[9]. The nature of the data was such that all adolescent participants were Italian. Thus, the age comparison utilised Italian adults only, whilst the nationality comparison excluded adolescents, in order to prevent age and nationality acting as confounding variables within comparisons.

Network estimation

A Gaussian Graphical Model (GGM; Lauritzen, 1996), a network in which edges represent partial correlations of ordinal or continuous data, was used to estimate the

[8] After removal of participants with missing data (n = 4).
[9] Eating disorders include obesity.

overall PSQ network in the pooled sample. Using partial correlations ensures that relationships between nodes are not confounded by relationships with other network variables; this enables unbiased centrality analyses (Stochl et al., 2019). GGMs are usually estimated using the graphical lasso, a method which utilises regularisation to avoid estimation of spurious edges. As a result, sparser, more interpretable networks are obtained (Friedman, Hastie & Tibshirani, 2008) in which covariance among nodes is explained with as few edges as necessary (Beard et al., 2016). If two nodes are connected by an edge in the resulting graph then they are statistically related after controlling for all other variables in the network; if no edges are present then they are conditionally independent (Fonseca-Pedrero et al., 2018).

The Fused Graphical Lasso (FGL; Danaher et al., 2014), an extension of GGM, was used to estimate networks across the subsamples using the R package *EstimateGroupNetwork* (Costantini & Epskamp, 2017). FGL allows the examination of similarities and differences across different samples and is utilised to produce a more accurate estimation of network structures than estimating networks individually (Danaher et al., 2014). FGL applies two penalty terms controlled by tuning parameters; firstly, a density penalty and secondly, a penalty on differences among corresponding edge weights in networks computed in different samples (Fried et al., 2018). K-fold cross-validation was utilised to select the tuning parameters for the penalty terms; this procedure means that the FGL neither masks differences nor inflates similarities between subsamples (Fried et al., 2018).

In all networks, Polychoric correlations were utilised to calculate edges. Polychoric correlations estimate associations between two variables that are theorised to be continuous and normally distributed but measured on ordinal scales (Beard et al.,

2016). Polychoric correlations have been shown to be unreliable in some circumstances (Epskamp & Fried, 2018). Following recommendations to ensure the appropriateness of Polychoric correlations, both Polychoric and Spearman's correlations were initially tested and compared in four steps: plotting the networks and visually inspecting, comparing the minimum and maximum edge-weights in each network, calculating the mean edge-weight in each network and correlating the Polychoric and Spearman edge-weights. These checks indicated that the Polychoric and Spearman's correlations were similar and thus it was concluded that Polychoric correlations were appropriate.

Network comparison

Networks were estimated and compared between UK and Italy[10], adult and adolescent[11], clinical and community, and complex and other diagnosis subsamples. Participants were represented in multiple networks (i.e. a single participant may be represented in Italian, adult, clinical and complex networks). Comparisons of networks where there are very different sample sizes are difficult to interpret as the level of regularisation is influenced by the sample size; with a smaller sample size, fewer edges are retained (Epskamp & Fried, 2018). One solution to enable more meaningful comparisons is to compare networks using a data-driven permutation test (van Borkulo et al., 2017). The R package *NetworkComparisonTest* (van Borkulo et al., 2017) was used to statistically compare each pair of networks using this method and to explore whether all edges were identical between them. Where there were significant differences between networks, post-hoc tests were utilised to investigate how many edges were significantly different. *NetworkComparionTest* was also utilised to explore whether global strength estimates, the sum of all absolute edge values, differed between networks. To assess whether

[10] Adult participants only.
[11] Italian participants only.

differences in sample size were influencing results, due to low power, a sensitivity analysis was conducted applying the Network Comparison Test (NCT) to subsamples with equal sizes. This comparison was achieved by subsampling the larger dataset to match the smaller one and repeating the NCT. If this produced different results to the uneven sample size comparisons, then sample size adjusted findings were reported.

Network inference

Centrality metrics were completed to explore which items were most integral to networks. High centrality nodes have strong connections to many other nodes, whilst low centrality nodes are peripheral with fewer and weaker connections (Robinaugh et al., 2016). Understanding the connectedness of nodes reflects how potentially clinically relevant a node may be (Rhemtulla et al., 2016). Previously, network analysis literature has investigated three measures of centrality: betweenness, closeness and node strength (Fried et al., 2018). The closeness of a node is the average distance from a node to all other nodes in the network (McNally et al., 2015) whilst betweenness is the number of times a node lies between two other nodes on their shortest connecting edge (Fried et al. 2018). However, more recent research has identified that betweenness and closeness are often not reliably estimated (Epskamp et al, 2017). Thus, analysis focused on node strength and predictability as the centrality indices. Node strength is the sum of each edge linked to the node and provides a relative measure of centrality (McNally et al., 2015). A node's predictability is an absolute measure of connectivity and represents the shared variance of each node with its neighbours (Haslbeck & Fried, 2017). Node predictability was calculated using the R package *mgm*.

Network stability

Factors including sample size, network size and network structure can impact on the robustness of network models; these factors must therefore be assessed (Borsboom, Robinaugh, Rhemtulla & Cramer, 2018). In order to ensure that results are robust and replicable, networks need to be assessed for stability (Epskamp et al., 2018). Stability estimation is not yet possible within jointly estimated networks (Fried et al., 2018) and therefore stability was assessed within individual networks. Multiple tests of stability were completed using case dropping bootstrapping methods in the R package *Bootnet* (Epskamp et al., 2018). The correlation stability (CS) coefficient was utilised to assess the stability of centrality indices when observing only portions of the data (Stochl et al., 2019).

Results

The demographic characteristics of the sample are summarised in Table 1.

Table 1. Demographic characteristics of the full sample

Variable	*N*	%
Group		
Clinical	772	49.7
Community	781	50.3
Gender		
Male	652	42.0
Female	887	57.1
Unknown	14	0.9
Age		
< 18	256	16.5
18-29	766	49.3
30-49	426	27.4
50-64	83	5.3
>65	18	1.2
Unknown	4	0.3
Nationality		
UK	627	40.4
Italy	777	50.0
USA	58	3.7
Australia	8	0.5
Canada	10	0.6
Other – Europe	30	1.9
Other – Asia	21	1.4
Other – Middle East	13	0.8
Other – Central America/Caribbean	4	0.3
Other – South America	3	0.2
Unknown	1	0.1
Diagnosis		
None	781	50.3
Depression	108	7.0
Self-harm	39	2.5
Anxiety, OCD and PTSD	79	5.1
Personality disorders	30	1.9
Psychosis	187	12.0
Eating disorders and obesity	221	14.2
Developmental disorders	58	3.7
Language disorders	4	0.3
Behavioural disorders	12	0.8
Physical health	10	0.6
Substance misuse and addiction	5	0.3
Other	13	0.8
Unknown	6	0.4

Descriptive statistics

Means for total and item-level PSQ scores by subsample are displayed in Table 2.

Table 2. Mean item level and total PSQ scores with standard deviations (SD) for each subsample

	Mean (Standard Deviation)							
Item	**UK**	**Italy**	**Adult**	**Adolescent**	**Clinical**	**Community**	**Complex**	**Other**
PSQ1	3.12 (1.17)	2.77 (1.17)	2.77 (1.17)	3.31 (1.45)	3.29 (1.23)	2.79 (1.09)	3.25 (1.24)	2.95 (1.15)
PSQ2	2.72 (1.18)	2.59 (1.18)	2.59 (1.18)	3.20 (1.30)	2.98 (1.31)	2.61 (1.10)	2.95 (1.30)	2.72 (1.18)
PSQ3	2.88 (1.16)	2.65 (1.18)	2.65 (1.18)	2.99 (1.28)	3.13 (1.27)	2.58 (1.07)	3.13 (1.28)	2.73 (1.15)
PSQ4	2.90 (1.13)	2.81 (1.18)	2.81 (1.18)	3.16 (1.16)	3.21 (1.21)	2.67 (1.08)	3.18 (1.13)	2.83 (1.14)
PSQ5	3.15 (1.23)	2.85 (1.21)	2.85 (1.21)	3.25 (1.32)	3.38 (1.33)	2.77 (1.11)	3.41 (1.32)	2.92 (1.21)
PSQ6	3.06 (1.28)	2.77 (1.25)	2.77 (1.25)	3.14 (1.37)	3.22 (1.34)	2.72 (1.21)	3.21 (1.34)	2.86 (1.26)
PSQ7	2.76 (1.19)	2.96 (1.11)	2.96 (1.11)	3.35 (1.24)	3.19 (1.25)	2.68 (1.08)	3.17 (1.25)	2.83 (1.16)
PS8Q	3.50 (1.09)	3.19 (1.23)	3.19 (1.23)	3.66 (1.20)	3.67 (1.18)	3.19 (1.15)	3.65 (1.21)	3.33 (1.17)
Total PSQ	24.08 (6.92)	22.58 (6.59)	22.58 (6.59)	26.06 (6.30)	26.06 (7.18)	22.01 (5.83)	25.94 (7.24)	23.16 (6.48)

There was a significant difference between UK and Italy in total PSQ scores, t (1144) = 3.74, $p < 0.05$. There were also significant differences between adult and adolescent total PSQ scores, t (773) = 6.99 , $p < 0.05$. and between clinical and community subsamples, t (1547) = 12.21, $p < 0.05$. Finally, there was a significant difference in total PSQ scores between participants with complex versus other diagnoses t (1547) = -7.50, $p < 0.05$.

Network Analysis

Aim 1: Global network estimation

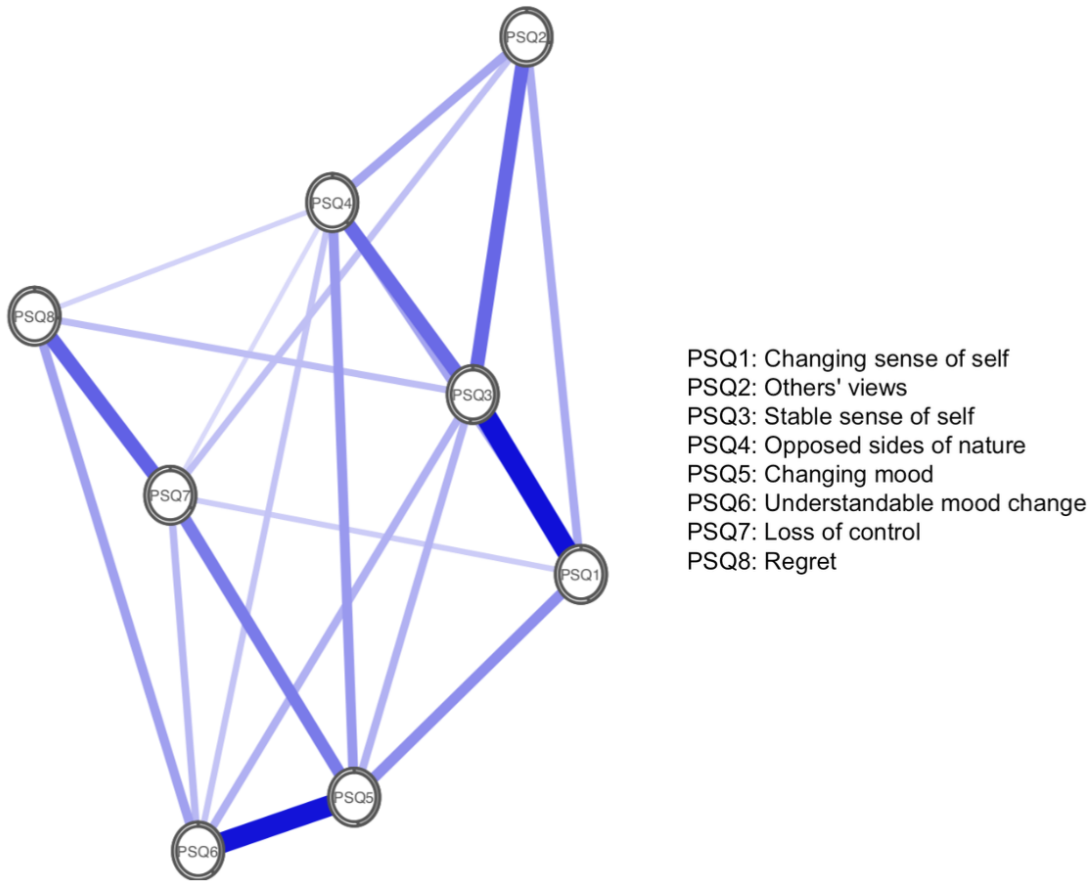
Figure 1a depicts the overall estimated network for the full sample (n = 1549). Each node represents an item of the PSQ whilst edges represent relationships between items, controlling for all other variables. All edges in the global network were positive; there were no negative associations between nodes. Thicker edges show stronger associations between nodes. The strongest edges were between PSQ1 and PSQ3 (changing sense of self and stable sense of self; 0.34[12]) and between PSQ5 and PSQ6 (changing mood and understandable mood change; 0.33). Moderate edges are shown between PSQ7 and PSQ8 (loss of control and regret; 0.22); PSQ3 and PSQ2 (stable sense of self and others' views; 0.21); and PSQ3 and PSQ4 (stable sense of self and opposed sides of nature; 0.21).

Figure 1b presents the centrality metrics for the overall PSQ network. The most central items in the full sample in terms of node strength, the sum of edges connected to a node, were PSQ3, (stable sense of self), and PSQ5 (changing mood). The least central items were PSQ8, (regret) and PSQ2 (others' views). Likewise, predictability, the

[12] Shortened summaries of PSQ items are utilised. For full PSQ items, see Appendix C.

variance of a node explained by its neighbours, was highest for PSQ3 and PSQ5 and lowest for PSQ2 and PSQ8. In sum, PSQ3 (stable sense of self) shared the strongest edge and had the highest node strength and predictability, indicating that it is the symptom most related to other symptoms in identity disturbance.

(a)



(b)

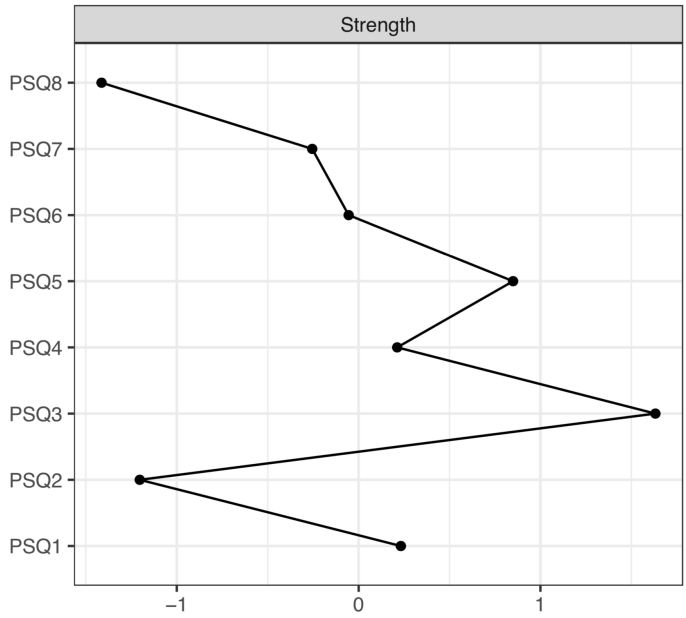


Figure 1. (a) Global network of identity disturbance across all datasets. Positive edges are represented by blue lines and negative edges by red lines. Shaded areas surrounding nodes represent node predictability. (b) Plot of node strength centrality.

Aim 2: Network comparison across subsamples

Comparison of PSQ networks across nationality (UK versus Italy)

Figure 2 depicts networks for participants from the UK and Italy. Both networks display a strong edge between PSQ5 (changing mood) and PSQ6 (understandable mood change). There is a strong edge between PSQ1 (changing sense of self) and PSQ3 (stable sense of self) in the UK network which is not observed in the Italian network. Additionally, there is a moderate edge between PSQ7 (loss of control) and PSQ8 (regret) in the UK network which is weaker in the Italian network.

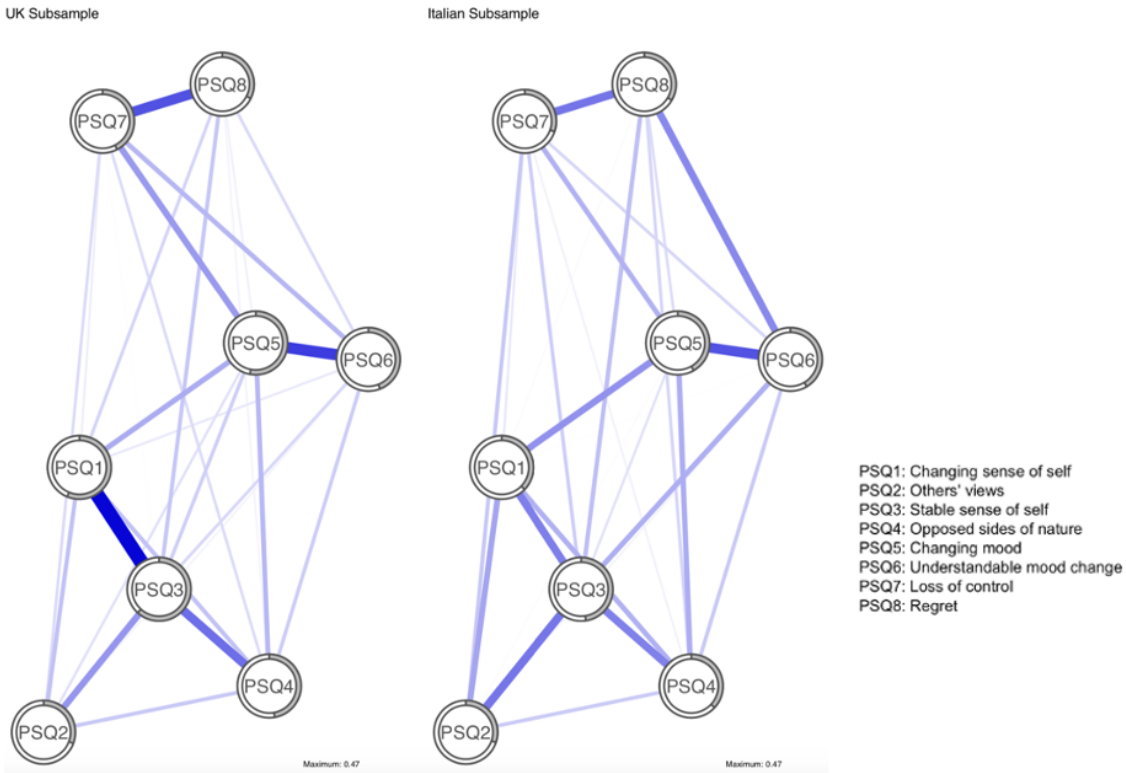


Figure 2. Jointly estimated network models of UK and Italian subsamples.

As displayed in figure 3, centrality was similar within the country network comparison. PSQ3 (stable sense of self) was the most central node in both the UK (1.65) and Italian (1.97) subsamples in terms of node strength. PSQ2 was the least

central node in both the UK (-1.39) and Italian (-1.26) subsamples. The UK subsample had the highest predictability of all subsamples at 0.46. That is, 46% of the variance of a node is explained by its neighbours. Predictability in the Italian subsample was 0.37.

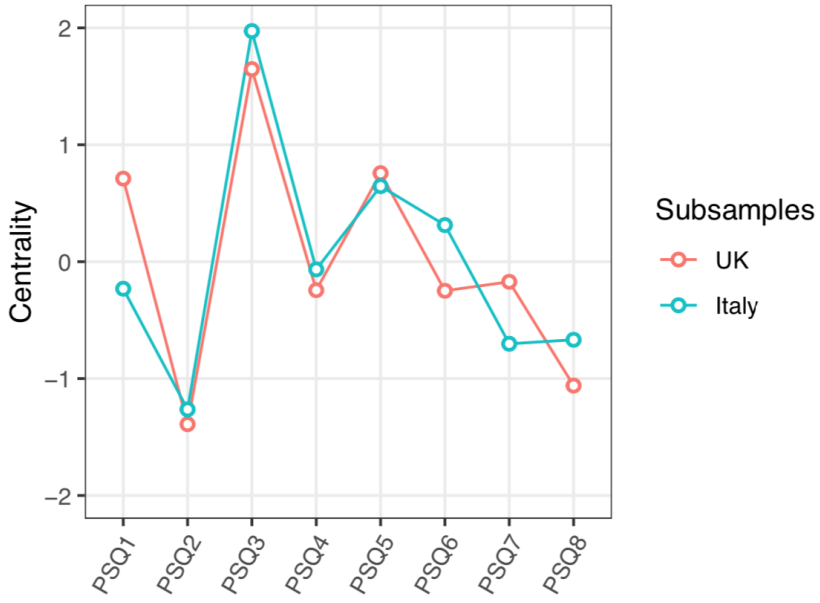


Figure 3. Comparison of centrality metrics between country subsamples.

The NCT showed that the UK and Italian subsamples were significantly different ($p < 0.05$). Post-hoc tests identified that 3 of 36 pairs of edges (8%) significantly differed between subsamples. These edges were PSQ1-PSQ3, PSQ6-PSQ8 and PSQ7-PSQ8. Global strength estimates did not significantly differ between the UK and Italy (UK = 3.37, Italy = 3.24, $p = 0.15$). Further comparison, adjusted for equal sample size, produced consistent results.

Comparison of PSQ networks across age (adult and adolescent)

Figure 4 depicts networks for adults and adolescents. Within both networks, there is a strong edge between PSQ5 (changing mood) and PSQ6 (understandable mood change). Both networks also display strong edges between PSQ2 (others' views) and PSQ3

(stable sense of self); this was particularly strong in the adolescent subsample. There were moderate edges between PSQ7 (loss of control) and PSQ8 (regret) as well as PSQ6 (understandable mood change) and PSQ8 (regret) in the adult subsample which were not present in the adolescent subsample.

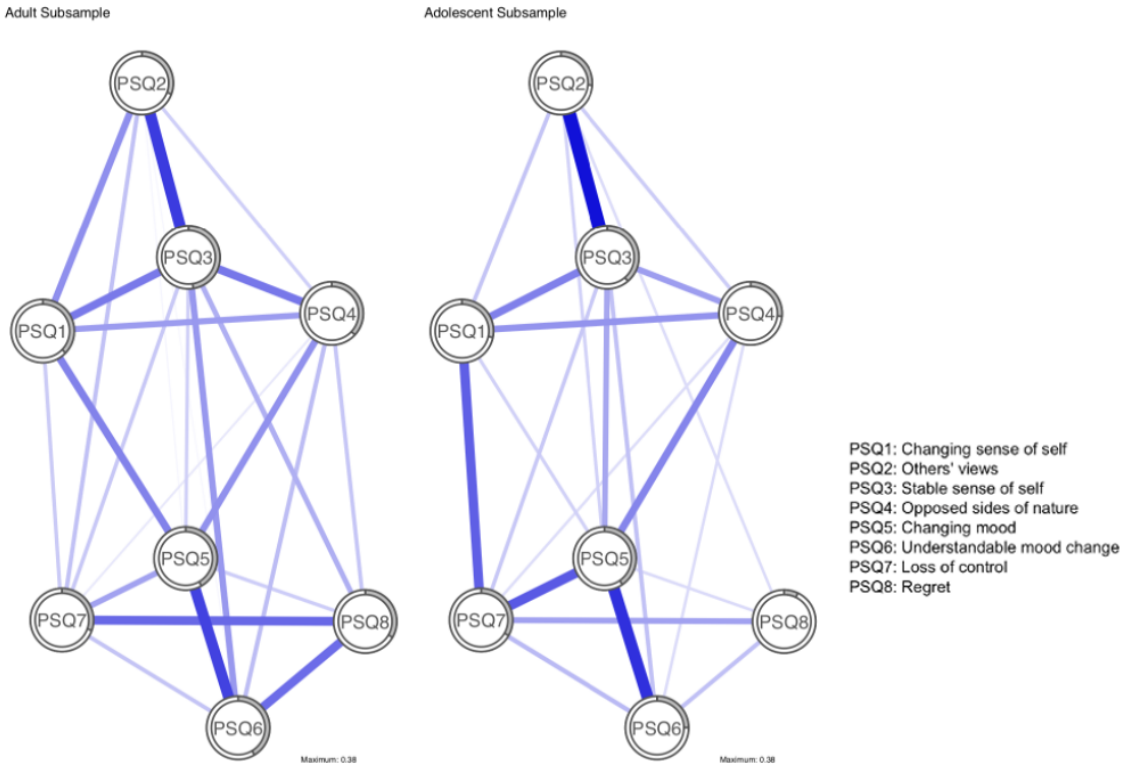


Figure 4. Jointly estimated network models of adult and adolescent subsamples.

Centrality for the adult and adolescent subsamples is depicted in figure 5. As displayed, there were differences in centrality between subsamples. In the adult subsample, PSQ3 (stable sense of self; 2.02) was most central, whereas PSQ5 (changing mood; 1.38) was most central in the adolescent subsample. The least central item in the adult subsample was PSQ2 (others' views; -1.20) whilst PSQ8 (regret; -1.77) was least central in the adolescent subsample. The adolescent subsample had the lowest predictability (0.28) of all of the subsamples. Predictability in the adult subsample was 0.37.

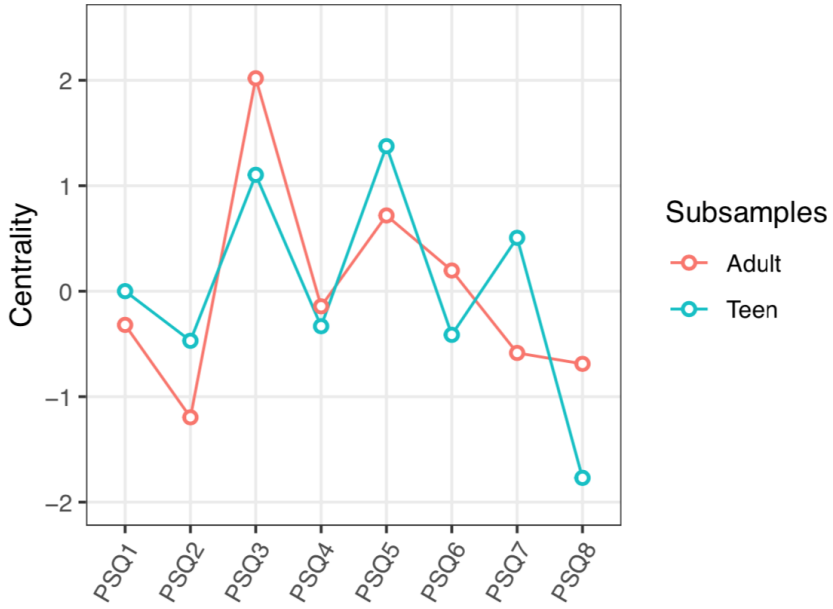


Figure 5. Comparison of centrality metrics between age subsamples.

The NCT showed that there was no statistically significant difference between the adult and adolescent networks ($p = 0.33$). Further analysis, adjusted for sample size, produced results consistent with the uneven sample size comparison. As the networks did not feature any significantly different edge weights, post-hoc analyses were not performed.

Comparison of PSQ networks across sample type (clinical versus community)

Figure 6 depicts networks for clinical and community participants. The networks are highly similar, with strong edges being observed in both networks between PSQ5 (changing mood) and PSQ6 (understandable mood change) as well as PSQ1 (changing sense of self) and PSQ3 (stable sense of self). Both networks displayed moderate edges between PSQ2 (others' views) and PSQ3 (stable sense of self) as well as PSQ7 (loss of control) and PSQ8 (regret).

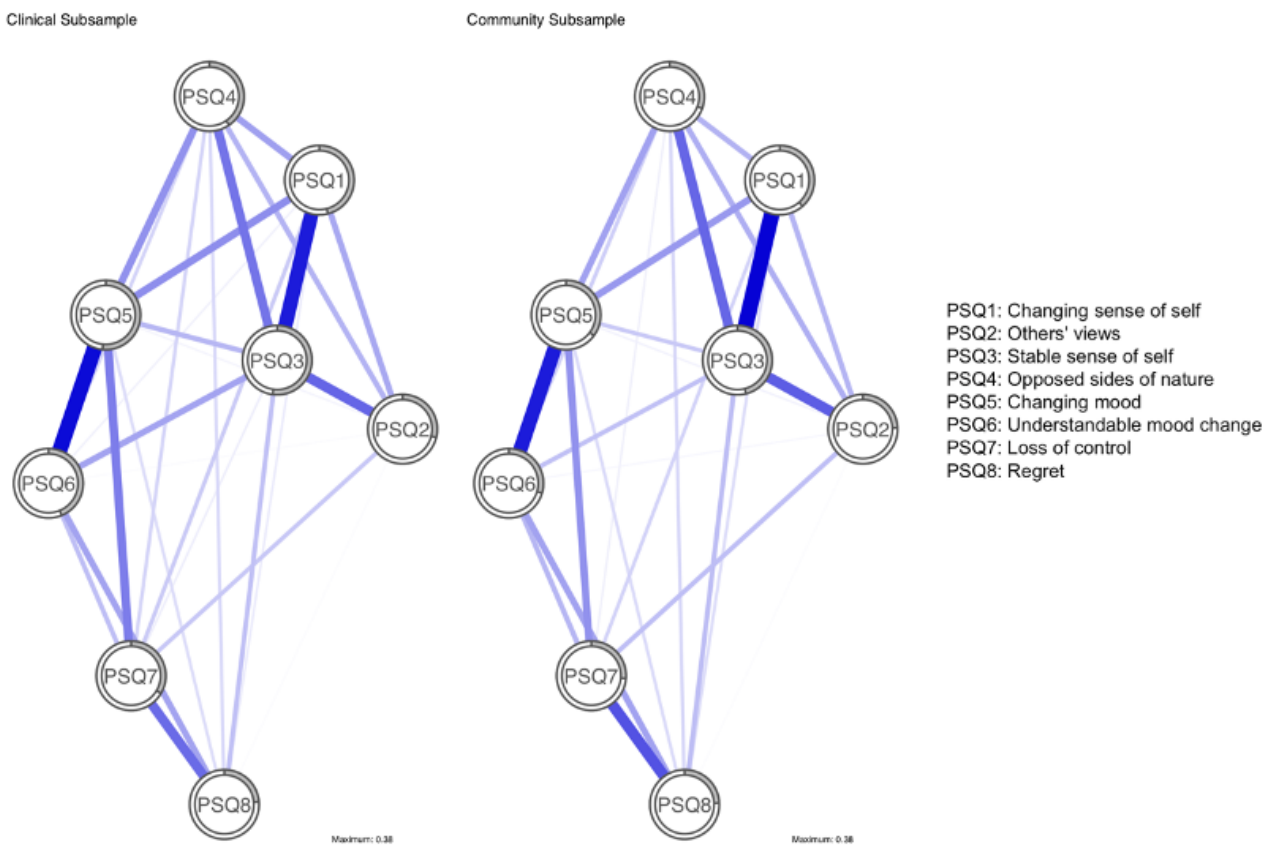


Figure 6. Jointly estimated network models of clinical and community subsamples.

As displayed in figure 7, centrality was highly consistent between the clinical and community subsamples. PSQ3 (stable sense of self) was the most central node for both clinical and community participants (1.52; 1.88 respectively) whilst PSQ2 (others' views) was the least central in clinical and community subsamples (-1.26; -1.19). Predictability was similar between subsamples at 0.40 in the clinical subsample and 0.32 in the community subsample.

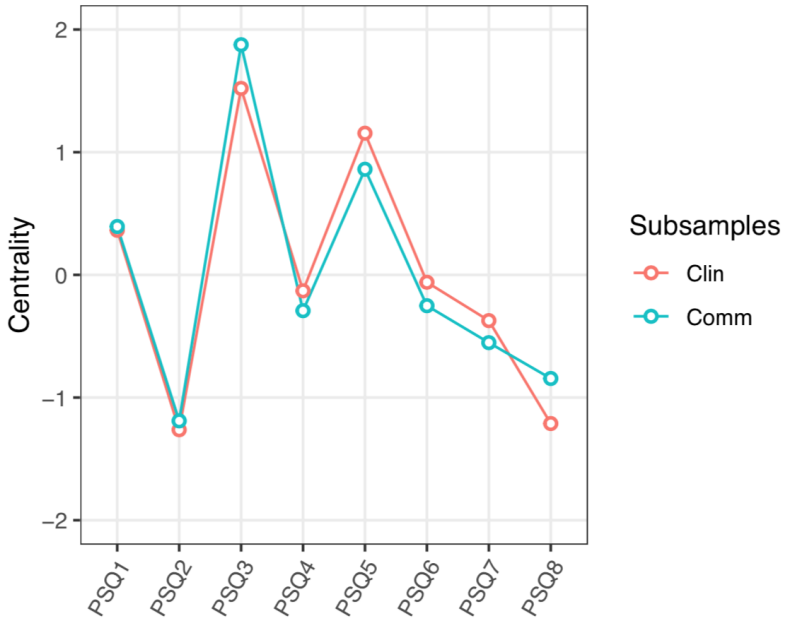


Figure 7. Comparison of centrality metrics between sample type.

The NCT showed that the community and clinical subsamples were not significantly different from each other ($p = 0.80$). Further analysis, adjusted for sample size, produced results consistent with the uneven sample size comparison. As the networks did not feature any significantly different edge weights, post-hoc analyses were not performed.

Comparison of PSQ networks across diagnoses (complex or other)

Figure 8 depicts networks for complex diagnoses versus other diagnoses. The networks are very similar, with strong edges being observed in both networks between PSQ1 (changing sense of self) and PSQ3 (others' views) as well as PSQ5 (changing mood) and PSQ6 (understandable mood change).

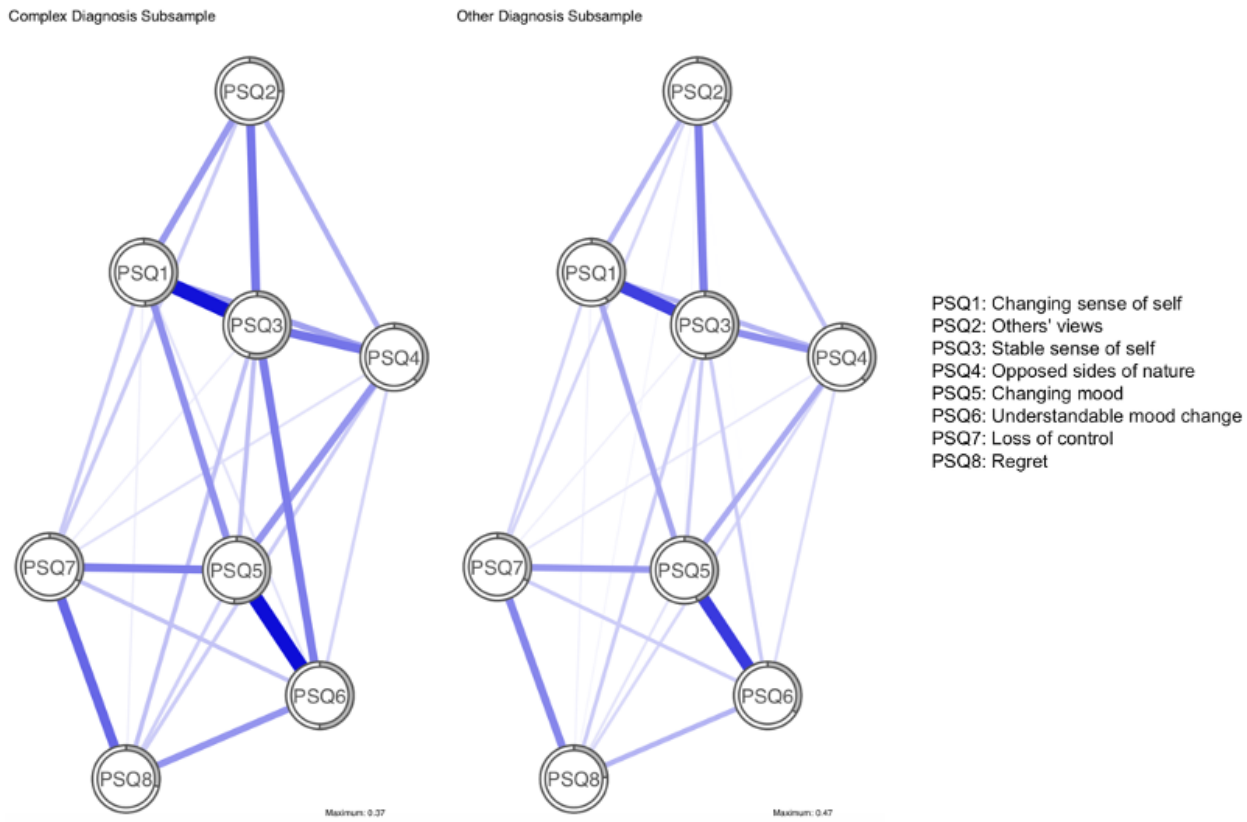


Figure 8. Jointly estimated network models of complex diagnoses and other diagnoses.

As displayed in figure 9, centrality was highly similar across diagnosis subsamples. PSQ3 (stable sense of self) was the most central item for both those with complex diagnoses (1.55) and those with other diagnoses (1.58). PSQ2 (others' views) was the least central item for both complex and other diagnosis subsamples (-1.45; -1.19 respectively). Predictability was very similar between subsamples at 0.41 in the complex subsample and 0.37 in the other diagnosis subsample.

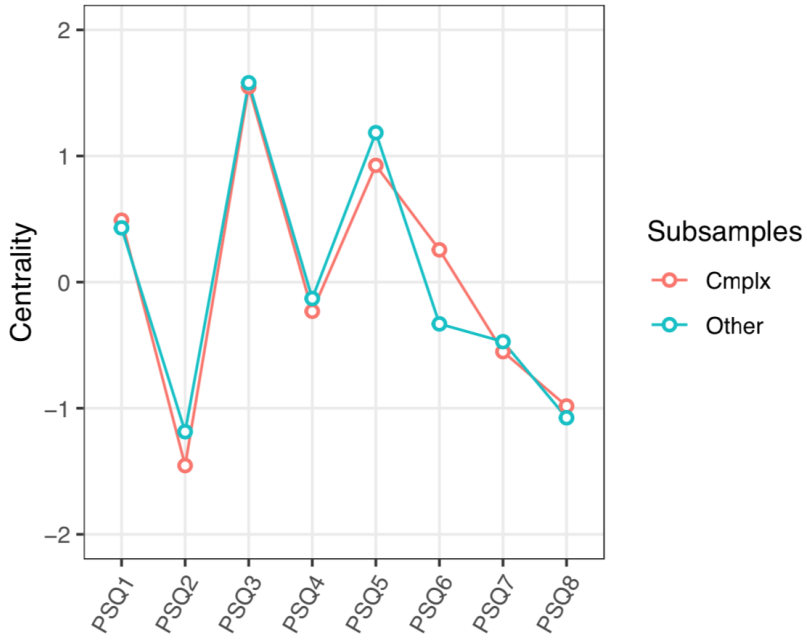


Figure 9. Comparison of centrality metrics between diagnosis subsamples.

The NCT showed that there was no significant difference between complex and other diagnosis subsample networks ($p = 0.39$). Further analysis, adjusted for sample size, produced results consistent with the uneven sample size comparison. As the networks did not feature any significantly different edge weights, post-hoc analyses were not performed.

Aim 3: Network stability

Guidance around the CS coefficient for strength centrality recommends that the coefficient should not drop below 0.25 and preferably be above 0.5 to assume that centrality indices are robust (Stochl et al., 2019). The UK, adult, clinical, complex and other diagnoses networks were stable (0.60, 0.75, 0.67, 0.67 and 0.67 respectively) at the preferable threshold. The Italian (0.28) and community networks (0.48) were stable at the acceptable threshold. The remaining network, adolescent, was below

recommendation (0.21). Therefore, the adolescent network and its centrality may be unreliable.

Summary of results

In sum, global network estimation identified that the most central nodes in the overall PSQ network were PSQ3 (stable sense of self) and PSQ5 (changing mood). Network comparisons identified a significant difference between UK and Italian networks. Networks did not differ by age, sample type (clinical or community) nor diagnosis. The majority of networks appeared to be adequately stable and robust, with the exception of the adolescent network which should be interpreted with caution.

Discussion

The present study used cross-sample and cross-national PSQ data to investigate the network structure of identity disturbance and to assess if networks differed across subsamples. T-tests found differences in total PSQ scores depending on nationality, age, sample type and diagnosis. Network analysis identified central PSQ items which were generally consistent between subsamples. Network comparisons identified that networks differed between UK and Italian subsamples. Networks did not differ between age, sample type or diagnoses.

Total PSQ scores

Total PSQ scores were significantly different between UK and Italy, adult and adolescent, clinical and community and complex and other diagnoses, with the UK, adolescent, clinical and complex groups having higher levels of identity disturbance than their comparators. This is in line with the wider evidence on identity disturbance. Firstly, patients with complex diagnoses included those with personality disorder in

which identity disturbance is identified as a key feature (Kernberg, 2006). Thus, the higher level of identity disturbance in the complex subsample is consistent with the evidence base. Additionally, the higher PSQ scores observed in the clinical subsample reinforce findings that the PSQ is a reliable tool in assisting diagnosis. Specifically, the mean total PSQ score in the clinical sample was consistent with evidence that a score of 26 and above is an appropriate diagnostic cut-off (Berrios et al., 2016). These initial results support existing evidence that the PSQ is a valid measure of identity disturbance and can be used to discriminate between differing diagnostic groups (Pollock et al., 2001). These preliminary findings encouraged the use of a network analysis.

Network analyses

The global network of PSQ items indicated that PSQ3 (stable sense of self) is the most integral symptom of identity disturbance. This was followed by PSQ5 (changing mood). This is consistent with literature that identifies lack of continuity in self-perception and affective instability to be prevalent in those presenting with identity disturbance (Walter et al., 2009). The findings around the centrality of stable sense of self and changing mood are consistent with the CAT approach to treatment of problems relating to identity, experienced as discontinuities in sense of self and changes in mood (Pollock et al., 2001). The MSSM model of CAT aims to treat patients with a diagnosis of BPD, for whom identity disturbance is prevalent, by mapping their range of states and switches between these states (Kellett, Bennett, Ryle & Thake, 2013).

A series of network comparisons found that networks of identity disturbance were not significantly different between adult and adolescent, clinical and community and complex and other diagnosis subsamples. There was, however, a statistically significant difference in network structures between the UK and Italy. Despite the

difference in strength, it is important to note that the overall pattern of node connections and centrality was similar. The difference between UK and Italian adults suggests that there is an impact of nationality on identity disturbance. This is keeping with our understanding that identity is not only an inner construct, but is influenced by society and culture (Jorgensen, 2006). Further, within the age comparison which consisted of Italian participants only, there were strong edges between PSQ2 (others' views) and PSQ3 (stable sense of self) in both adult and adolescent networks; this was not observed in other networks. This suggests that the views of others are of increased importance in identity within Italian culture. The increased strength of this edge within the adolescent subsample is consistent with the importance of peers during this developmental stage (Westen et al., 2011).

Despite there being no significant difference between adult and adolescent networks, differences in centrality emerged. These findings are in line with previous findings that identity disturbance in adolescents has a strong resemblance to identity disturbance in adults, but with some key differences (Westen et al., 2011). Results showed that there was no significant difference in the network structures between those with a complex diagnosis and those with other diagnoses. However, as we might expect, initial t-tests demonstrated that total PSQ scores were significantly higher in the complex subsample. This suggests that level of identity disturbance is higher in those with complex diagnoses but the structure of identity disturbance is consistent. This supports the use of the PSQ across clinical presentations.

Centrality was generally similar across subsample comparisons with some differences between adults and adolescents. PSQ3 (stable sense of self) was the most central in all subsamples with the exception of the adolescent subsample. Similarly,

PSQ2 (others' views) was the least central in all subsamples, with the exception of the adolescent subsample in which PSQ8 (regret) was least central. Regret has important links with decision-making in adulthood, yet the ability to experience regret begins relatively late in development (McCormack, Feeney & Beck, 2020). Thus, the findings around centrality are consistent with developmental theory.

The findings around centrality have clinical implications in that the most central nodes in a network are most likely to activate other nodes and may therefore be important in intervention (Fried et al., 2016). Predictability, one of the centrality measures used, has been identified to be an attempt to characterise the controllability of a network; if predictability is high, nodes or symptoms can be controlled by their neighbouring nodes in the network, whilst low predictability means that nodes should be addressed directly (Haslbeck & Fried, 2017). Within this study, substantial variability remained unexplained by neighbouring nodes. Thus, whilst some targeting of neighbouring nodes may impact on other nodes, direct intervention on central nodes may be more effective.

Network stability analysis showed that the majority of the individually estimated networks were robust. However, the adolescent subsample network was below the recommended level in terms of its CS coefficient. Given our understanding that stability of networks can be influenced by sample size (Borsboom et al., 2018) and the comparatively small adolescent subsample size, the fact that this is the least stable network is unsurprising. The lack of stability in this network should be held in mind when interpreting results. However, it is noted that the recommendations around CS coefficient cut-offs are not definitive guidelines (Epskamp et al., 2018) and thus, the network should not be discounted entirely.

Clinical implications

The findings support the use of the PSQ as a clinical measure of identity disturbance in both UK and non-UK settings. This is beneficial to clinicians given the brevity of the PSQ, which makes it desirable in comparison to longer measures of identity disturbance, such as the 30-item SCIM (Berrios et al., 2016). The present study therefore addresses the need for brief measures of identity disturbance for use in routine clinical practice (Ryle, 1995) and suggests that the PSQ could track state-shifting over time if used as a sessional outcome measure. The findings support the continued use of the PSQ by clinicians using CAT, in which the PSQ is central to assessment (Ryle & Kerr, 2002).

Understanding network structures behind an individuals' symptom interactions enable us to select interventions that make a healthy state accessible (Borsboom, 2017). The findings of this research suggest that interventions that aim to create a more stable sense of self and that target changing mood are likely to be effective in treating identity disturbance. This would indicate that interventions that create patterns of behavioural consistency in patients' lives are of particular benefit. This could be a behavioural thread running through any psychological intervention with the aim of reducing state-shifting. Given the impact of others' views in the adolescent subsample, interventions with young people may benefit from an additional focus on the system around the individual. However, this recommendation is caveated by the concerns around reliability with this subsample due to its small size.

It is crucial to highlight that interventions should not be developed around centrality alone. Firstly, nodes may differ in the extent to which they are susceptible to

change through intervention (Robinaugh et al., 2016); less central nodes may therefore have greater potential for clinical change when targeted. Additionally, nodes with the lowest centrality may still be highly important clinical features that may be linked to risk (Fried et al., 2018) and so the constructs represented by the least central PSQ items should not be disregarded. It is also possible that the targeting of edges as opposed to individual nodes may be a successful intervention (Robinaugh et al., 2016). Thus, the findings around centrality do not provide concrete guidance on intervention in isolation but rather should be held in mind by clinicians when making clinical judgements.

Theoretical implications

The present study addresses gaps in the evidence base for both the PSQ and network approaches to identity disturbance by providing a network analysis utilising a large multi-site, cross-cultural sample. This study has shifted the perspective typically applied to the PSQ; rather than assuming the PSQ captures a singular underlying process common to all features of identity disturbance, the current study has explored how features of identity disturbance interact. The observed strong edges between PSQ items in networks of identity disturbance are consistent with literature regarding the MSSM model of CAT, which finds deficits in interpersonal regulation (e.g. self-to other interactions), continuity (e.g. state shifts) and coherence (e.g. mood variability) to occur along a continuum of severity (Pollock et al., 2001).

The investigation of network models across multiple datasets is a key strength of this study as it addresses concerns around the replicability crisis in network analysis literature (Borsboom et al., 2017). Additionally, the use of both clinical and community populations increases the generalisability of findings and address concerns highlighted by Fried et al. (2018) about the use of singular, often non-clinical samples in network

analyses. The even representation of males and females in the research further improves generalisability. Previously, studies exploring the PSQ have utilised adult samples and so the inclusion of adolescents in this research makes an important contribution to the evidence.

Limitations

It must be acknowledged that there are limitations to this study and results should be interpreted with these in mind. Whilst this research adds to the evidence-base by exploring the use of the PSQ outside the UK, this is limited to an Italian subsample only. Despite the presence of participants in the full sample from countries other than the UK and Italy, the small subsamples did not allow for network comparisons between other nationalities. The comparison between nationality is therefore only reflective of Western Europe. Additionally, due to the nature of the data available, all adolescent participants were Italian and hence, the age comparison was limited to Italian data. Furthermore, the adolescent subsample was small in comparison to other subsamples and therefore may have lacked power to detect differences between networks.

There are also limitations with the use of multi-site secondary data. Given the different sources, it is not possible to ensure control over how participants are categorised. For example, it is possible that definitions of community participants may differ between sites. Whilst all community data was collected in community settings, it may be that some community participants had previous clinical input or existing mental health disorders. Additionally, the way in which diagnoses were recorded may differ between sites. It is important to note that diagnosis was defined by the primary diagnosis or presenting problem. Therefore, a patient may have a diagnosis of personality disorder but be presenting for treatment due to low mood; their presenting

problem may therefore be categorised as depression, placing them in the 'other' subsample, despite the presence of a complex diagnosis. Given the use of secondary data, further control around defining diagnosis subsamples was not possible. However, it must be considered that the nature of this multi-site study using secondary data is such that there may be unidentified confounding variables.

Future research

The present study makes an important step to assess the transcultural validity of the PSQ. However, further exploration utilising data from other countries is important. Given that a person's identity can be influenced by culture or society (Jorgensen, 2006), it is important for future research to explore the use of the PSQ with individuals from countries and cultures extending beyond the UK and Italy and indeed Western Europe.

Additionally, the use of larger, more appropriately matched sample sizes would improve the robustness of comparisons, particularly with regard to the adolescent subsample which was comparatively small. Further, age comparisons across nationalities would be of interest, given that the age comparison consisted of Italian data only. The inclusion of specific older adult clinical and community samples would be a useful addition to future methodologies. Comparing intake and discharge PSQ networks in patients that demonstrate clinical and reliable change on the PSQ would also be of interest.

Conclusions

This is the first known study to explore identity disturbance, and indeed the PSQ, utilising a network approach. The nature of this research as a cross-national, multi-site study increases generalisability and addresses concerns about replicability in network

analysis. This contribution to the evidence-base for the PSQ addresses the highlighted need for validated measures assessing identity disturbance for use in clinical practice (Kaufman et al., 2015). The findings have clinical implications for assessment and intervention, particularly for those utilising CAT, in which the PSQ has been identified as central (Ryle & Kerr, 2002). The findings can enhance assessment but also provide direction for intervention, given that the study has been a step forward in identifying potential treatment targets. Future work needs to evaluate if such targeting is clinically effective. Whilst there are limitations to the research, it provides direction for future research to develop the evidence base for the assessment of identity disturbance. The PSQ is emerging as a useful brief measure for this endeavour.

References

Adler, J. M., Chin, E. D., Kolisetty, A. P., & Oltmanns, T. F. (2012). The distinguishing characteristics of narrative identity in adults with features of borderline personality disorder: an empirical investigation. *Journal of Personality Disorders, 26,* 498-512. Doi: 10.1521/pedi.2012.26.4.498

Beard, A., Millner, A.J., Forgeard, M. J., Fried, E. I., Hsu, K. J., Treadway, M. T. . . . Bjorgvinsson, T. (2016). Network analysis of depression and anxiety symptom relationships in a psychiatric sample. *Psychological Medicine, 46,* 3359-3369. Doi: 10.1017/S0033291716002300

Bedford, A., Davies, F., & Tibbles, J. (2009). The personality structure questionnaire (PSQ): a cross-validation with a large clinical sample. *Clinical Psychology and Psychotherapy, 16,* 77-81. Doi: 10.1002/cpp.603

Berrios, R., Kellett, S., Fiorani, C., & Poggioli, M. (2016). Assessment of identity disturbance: factor structure and validation of the personality structure questionnaire in an Italian sample. *Psychological Assessment, 28,* 27-35. Doi: 10.1037/pas0000143

Borgatti, S. P., Carley, K. M., & Krackhardt, D. (2006). On the robustness of centrality measures under conditions of imperfect data. *Social Networks, 28,* 124-136. Doi: 10.1016/j.socnet.2005.05.001

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

Borsboom, D., & Cramer, A. O. (2013). Network analysis: an integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology, 9,* 91-121. Doi: 10.1146/annurev-clinpsy-050212-185608

Borsboom, D., Fried, E. I., Epskamp, S., Waldorp, L. J., Van Borkulo, C. D., Van Der Maas, H. L., & Cramer, A. O. (2017). False alarm? A comprehensive reanalysis of "evidence that psychopathology symptom networks have limited replicability" by Forbes, Wright, Markon, and Krueger (2017). *Journal of Abnormal Psychology, 126,* 989-999. Doi: 10.1037/abn0000306

Borsboom, D., Robinaugh, D. J., Rhemtulla, M., & Cramer, A. O. (2018). Robustness and replicability of psychopathology networks. *World Psychiatry, 17,* 143-144. Doi: 10.1002/wps.20515

Boyson, G. A., & VanBergen, A. (2013). A review of published research on adult dissociative identity disorder: 2000-2010. *The Journal of Nervous and Mental Disease, 201,* 5-11. Doi: 10.1097/NMD.0b013e31827aaf81

Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., … Tuerlinckx, F. (2013). A network approach to psychopathology: new insights into clinical longitudinal data. *PLOS ONE, 8,* p.e60188. Doi: 10.1371/journal.pone.0060188

Costantini, G., & Epskamp, S. (2017). EstimateGroupNetwork: Perform the joint graphical lasso and selects tuning parameters. R package (Version 0.1.2) [Computer software]. Retrieved from https://cran.r-project.org/web/packages/EstimateGroupNetwork/index.html

Danaher, P., Wang, P., & Witten, D. M. (2014). The joint graphical lasso for inverse covariance estimation across multiple classes. *Journal of the Royal Statistical Society, 76,* 373-397. Doi: 10.1111/rssb.12033

Deaux, K. (2000). Identity. In A. E. Kazdin (Ed.), *Encyclopaedia of Psychology* (pp. 222-225). Oxford: Oxford University Press.

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

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

Epskamp, S., & Fried, E. I. (2018). A tutorial on regularized partial correlation networks. *Psychological Methods, 23,* 617-634. Doi: 10.1037/met0000167

Erikson, E. H. (1968). *Identity, youth and crisis.* New York: Norton.

Fonseca-Pedrero, E., Ortuño, J., Debanné, M., Chan, R. C., Cicero, D., Zhang, L. C., . . . Fried, E. I. (2018). The network structure of schizotypal personality traits. *Schizophrenia Bulletin, 44,* 468-479. Doi: 10.1093/schbul/sby044

Forbes, M. K., Wright, A. G., Markon, K. E., & Krueger, R. F. (2017). Evidence that psychopathology symptom networks have limited replicability. *Journal of Abnormal Psychology, 126,* 969-988. Doi: 10.1037/abn0000276

Fried, E. I., Eidhof, M. B., Palic, S., Costantini, G., Huisman-van Dijk, H. M., Bockting, C. L., . . . Karstoft, K. (2018). Replicability and generalizability of

posttraumatic stress disorder (PTSD) networks: a cross-cultural multisite study of PTSD symptoms in four trauma patient samples. *Clinical Psychological Science, 6,* 335-351. Doi: 10.1177/2167702617745092

Fried, E. I., Epskamp, S., Nesse, R. M., Tuerlinckx, F., & Borsboom, D. (2016). What are 'good' depression symptoms? Comparing the centrality of DSM and non-DSM symptoms of depression in a network analysis. *Journal of Affective Disorders, 189,* 314-320. Doi: 10.1016/j.jad.2015.09.005

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,* 1-10. Doi: 10.1007/s00127-016-1319-z

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics, 9,* 432-441. Doi: 10.1093/biostatistics/kxm045

Goth, K., Foelsch, P., Schluter-Muller, S., Birkholzer, M., Jung, E., Pick, O., & Schmeck, K. (2012). Assessment of identity development and identity diffusion in adolescence – theoretical basis and psychometric properties of the self-report questionnaire AIDA. *Child and Adolescent Psychiatry and Mental Health, 6,* 27. Doi: 10.1186/1753-2000-6-27

Haslbeck, J. M., & Fried, E. I. (2017). How predictable are symptoms in psychopathological networks? A reanalysis of 18 published datasets. *Psychological Medicine, 47,* 2767-2776. Doi: 10.1017/S0033291717001258

Horz-Sagstetter, S., Caligor, E., Preti, E., Stern, B. L., De Panfilis, C., & Clarkin, J. F. (2018). Clinician-guided assessment of personality using the structural interview and the structured interview of personality organization (STIPO). *Journal of Personality Assessment, 100,* 30-42. Doi: 10.1080/00223891.2017.1298115

Isvoranu, A., Guloksuz, S., Epskamp, S., van Os, J., & Borsboom, D. (2019). Toward incorporating genetic risk scores into symptom networks of psychosis. *Psychological medicine, 50*(4)*,* 1-8. Doi: 10.1017/s003329171900045x

Jorgensen, C. R. (2006). Disturbed sense of identity in borderline personality disorder. *Journal of Personality Disorders, 20,* 618-644. Doi: 10.1521/pedi.2006.20.6.618

Kaufman, E., Cundiff, J., & Crowell, S. (2015). The development, factor structure, and validation of the self-concept and identity measure (SCIM): A self-report assessment of clinical identity disturbance. *Journal of Psychopathology and Behavioral Assessment, 37,* 122-133. Doi: 10.1007/s10862-014-9441-2

Kellett, S., Bennett, D., Ryle, T., & Thake, A. (2013). Cognitive analytic therapy for borderline personality disorder: therapist competence and therapeutic effectiveness in routine practice. *Clinical Psychology and Psychotherapy, 20,* 216-225. Doi: 10.1002/cpp.796

Kernberg, O. F. (2006). Identity: recent findings and clinical implications. *The Psychoanalytic Quarterly, 75,* 969-1004. Doi: 10.1002/j.2167-4086.2006.tb00065.x

Lauritzen, S. L. (1996). *Graphical models.* Gloucestershire, England: Clarendon Press.

McCormack, T., Feeney, A., & Beck, S. R. (2020). Regret and decision-making: a developmental perspective. *Current Directions in Psychological Science: A Journal of the American Psychological Society, 29,* 346-350. Doi: 10.1177/0963721420917688

McNally, R. J., Robinaugh, D. J., Wu, G. W., Wang, L., Deserno, M. K., & Borsboom, D. (2015). Mental disorders as casual systems: a network approach to posttraumatic stress disorder. *Clinical Psychological Science, 3,* 836-849. Doi: 10.1177/2167702614553230

Modestin, J., Oberson, B., & Erni, T. (1998). Identity disturbance in personality disorders. *Comprehensive Psychiatry, 39,* 352-357. Doi: 10.1016/S0010-440X(98)90047-5

Neacsiu, A. D., Herr, N. R., Fang, C. M., Rodriguez, M. A., & Rosenthal, M. Z. (2015). Identity disturbance and problems with emotion regulation are related constructs across diagnoses: identity disturbance and emotion dysregulation. *Journal of Clinical Psychology, 71,* 346-361. Doi: 10.1002/jclp.22141

Pollock, P. H., Broadbent, M., Clarke, S., Dorrian, A., & Ryle, A. (2001). The personality structure questionnaire (PSQ): a measure of the multiple self states model of identity disturbance in cognitive analytic therapy. *Clinical Psychology and Psychotherapy, 8,* 59-72. doi:10.1002/cpp.250

Rhemtulla, M., Fried, E. I., Aggen, S. H., Tuerlinckx, F., Kendler, K. S., & Borsboom, D. (2016). Network analysis of substance abuse and dependence symptoms. *Drug and Alcohol Dependence, 161,* 230-237. Doi: 10.1016/j.drugalcdep.2016.02.005

Richetin, J., Preti, E., Costantini, G., & De Panfilis, C. (2017). The centrality of affective instability and identity in borderline personality disorder: evidence from network analysis. *PLOS ONE, 12,* p.e0186695. Doi: 10.1371/journal.pone.0186695

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

Ryle, A. (1995). *Cognitive analytic therapy: Developments in theory and practice.* Wiley: Chichester.

Ryle, A. (1997). *Cognitive analytic therapy and borderline personality disorder: The model and the method.* Wiley: Chichester.

Ryle, A., & Kerr, I. (2002). *Introducing cognitive analytic therapy: Principles and practice*. Chichester: John Wiley & Sons.

Southward, M. W., & Cheavens, J. S. (2018). Identifying core deficits in a dimensional model of borderline personality disorder features: a network analysis. *Clinical Psychological Science, 6,* 685-703. Doi: 10.1177/2167702618769560

Taylor, S., & Goritsas, E. (1994). Dimensions of identity diffusion. *Journal of Personality Disorders, 8,* 229-239. Doi: 10.1521/pedi.1994.8.3.229

Thorlund, K., & Mills, E. J. (2012). Sample size and power considerations in network meta-analysis. *Systematic Reviews*, *1*, 41. Doi: 10.1186/2046-4053-1-41

van Borkulo, C., Boschloo, L., Kossakowski, J., Tio, P., Schoevers, R., Borsboom, D., . . . Waldorp, L. (2017). Comparing network structures on three aspects: a permutation test. Unpublished manuscript. Doi: 10.13140/RG.2.2.29455.38569

Walter, M., Berth, H., Selinger, J., Gerhard, U., Kuchenhoff, J., Frommer, J., & Dammann, G. (2009). The lack of negative affects as an indicator for identity disturbance in borderline personality disorder: a preliminary report. *Psychopathology, 42,* 399-404. Doi: 10.1159/000241196

Westen, D., Betan, E., & Defife, J. A. (2011). Identity disturbance in adolescence: associations with borderline personality disorder. *Development and Psychopathology, 23*, 305-313. Doi: 10.1017/S0954579410000817

Wilkinson-Ryan, T., & Westen, D. (2000). Identity disturbance in borderline personality disorder: an empirical investigation. *American Journal of Psychiatry, 157,* 528-541. Doi: 10.1176/appi.ajp.157.4.528

Zandersen, M., & Parnas, J. (2019). Identity disturbance, feelings of emptiness, and the boundaries of the schizophrenia spectrum. *Schizophrenia Bulletin, 45,* 106-113. Doi: 10.1093/schbul/sbx183

Appendices

Appendix A – Ethical Approval Letter

Downloaded: 05/06/2020
Approved: 01/06/2020

Georgia Mangion
Registration number: 170149374
Psychology
Programme: Doctorate of Clinical Psychology

Dear Georgia

PROJECT TITLE: A Network Analysis of the Personality Structure Questionnaire
APPLICATION: Reference Number 034903

On behalf of the University ethics reviewers who reviewed your project, I am pleased to inform you that on 01/06/2020 the above-named project was **approved** on ethics grounds, on the basis that you will adhere to the following documentation that you submitted for ethics review:

- University research ethics application form 034903 (form submission date: 19/05/2020); (expected project end date: 30/11/2020).

If during the course of the project you need to deviate significantly from the above-approved documentation please inform me since written approval will be required.

Your responsibilities in delivering this research project are set out at the end of this letter.

Yours sincerely

Thomas Webb
Ethics Administrator
Psychology

Please note the following responsibilities of the researcher in delivering the research project:

- The project must abide by the University's Research Ethics Policy: https://www.sheffield.ac.uk/rs/ethicsandintegrity/ethicspolicy/approval-procedure
- The project must abide by the University's Good Research & Innovation Practices Policy: https://www.sheffield.ac.uk/polopoly_fs/1.671066!/file/GRIPPolicy.pdf
- The researcher must inform their supervisor (in the case of a student) or Ethics Administrator (in the case of a member of staff) of any significant changes to the project or the approved documentation.
- The researcher must comply with the requirements of the law and relevant guidelines relating to security and confidentiality of personal data.
- The researcher is responsible for effectively managing the data collected both during and after the end of the project in line with best practice, and any relevant legislative, regulatory or contractual requirements.

Appendix B – Description of Samples

The overall sample (*N* = 1553) was made up of smaller, multi-site samples. These samples are broken down below:

1) A UK sample of adults with a history of two or more experiences of self-harm (*N* = 96) and a control group with no experience of self-harm (*N* = 110)
2) A UK clinical sample of adults - private practice (*N* = 73)
3) A sample of adults presenting with psychosis (*N* = 182) and a non-clinical control group (*N* = 295)
4) A UK adult clinical sample - NHS (*N* = 22)
5) An Italian adult sample consisting of a clinical population (*N* = 237) and a community control (*N* = 296)
6) An Italian sample of both adolescents and adults consisting of a clinical population (*N* = 152) and community control (*N* = 90)

Appendix C – Personality Structure Questionnaire

Personality Structure Questionnaire

The aim of this questionnaire is to obtain an account of certain aspects of your personality. People vary greatly in all sorts of ways: the aim of this form is to find out how far your feel yourself to be constant and 'all of a piece' or variable and made up of a number of distinct 'sub-personalities' or liable to experience yourself as shifting between two or more quite distinct and sharply differentiated states of mind.

Most of us experience ourselves as somewhere between these contrasted ways. A state of mind is recognised by a typical mood, a particular sense of oneself and of others and by how for one is in touch with, and in control of, feelings. Such states are definite, recognizable ways of being; one is wither clearly in a given state or one is not. They often affect one quite suddenly; they may be of brief duration or they last for days, Sometimes, but not always, changes of state happen because of change in circumstances or an event of some kind.

Please indicated which description applies to you most closely by shading the appropriate circle.
Please complete ALL questions.

Shade circles like this: ● **Not** like this: ⊘ or ✓
Shade one circle per question only.

THANK YOU FOR YOUR HELP. ALL INFORMATION WILL BE TREATED AS CONFIDENTIAL.

	1 Very True	2 True	3 May or may not be True	4 True	5 Very True	
1. My sense of self is always the same	○	○	○	○	○	How I act or feel is constantly changing
2. The various people in my life see me in much the same way.	○	○	○	○	○	The various people in my life has different views of me as if I were not the same person
3. I have a stable and unchanging sense of myself	○	○	○	○	○	I am so different at some times that I wonder who I really am
4. I have no sense of opposed sides to my nature	○	○	○	○	○	I feel I am split between two (or more) ways of being, sharply differentiated from each other
5. My mood and sense of self seldom change suddenly	○	○	○	○	○	My mood can change abruptly in ways which make me feel unreal or out of control
6. My mood changes are always understandable	○	○	○	○	○	I am often confused by my mood changes which seem either unprovoked or quite out of scale with what provoked them
7. I never lose control	○	○	○	○	○	I get into states in which I lose control and do harm to myself and / or others
8. I never regret what I have said or done	○	○	○	○	○	I get into states in which I do and say things which I later regret

ACAT Administration
Association for Cognitive Analytic Therapy

Association for Cognitive Analytic Therapy, 3rd Floor, South Wing, Division of Academic Psychiatry
St Thomas' Hospital, Lambeth Palace Road, London SE1 7EH
enquiries@acat.org.uk www.acat.org.uk 0207 928 9292 ext 3769