This is a repository copy of *Cognitive analytic therapy: A review of the outcome evidence base for treatment*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/102724/

Version: Accepted Version

Article:

Calvert, R. and Kellett, S. (2014) Cognitive analytic therapy: A review of the outcome evidence base for treatment. Psychology and Psychotherapy: Theory, Research and Practice, 87 (3). pp. 253-277. ISSN 1476-0835

https://doi.org/10.1111/papt.12020

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Purpose. This study reviews the quality of the extant outcome evidence for Cognitive Analytic Therapy (CAT) in order to inform decisions of when to use the model with patients and to stimulate a future CAT research strategy.

Method. An electronic search identified CAT efficacy and effectiveness studies and these were subject to systematic review. The methodological quality of studies meeting inclusion criteria were appraised using two validated research study quality checklists and studies were fitted to an established model of psychotherapy evaluation.

Results. Twenty-five outcome studies met the inclusion criteria, including five randomised controlled trials. The CAT evidence-base is predominated by small-scale practice-based studies, in typically complex and severe clinical populations - 44% were focal to the treatment of personality disorder. Whilst the quality of extant CAT evidence is generally sound (52% of studies were high quality), the depth and breadth of the evidence-base is currently limited. Where comparisons with other modalities were available, CAT appeared largely equivocal.

Conclusions. CAT is a popular and promising intervention for complex presentations. However, the evidence-base currently lacks wider credibility due to having largely bypassed the rigors of the controlled phase of the hourglass model of psychotherapy evaluation. There is a particular need for further CAT outcome research with common mental health problems.

Practitioner points

- CAT can be an effective intervention across a range of mental health difficulties
- Consider a 24 session CAT contract for those patients presenting with complex and severe difficulties

- Practice research networks could make a significant contribution to the CAT evidence-base

Cognitive Analytic Therapy (CAT) was developed as a time-limited, integrative and researchable psychotherapy specifically for the needs of the public sector (Ryle, 1995). Despite increasing numbers of clinicians engaging in training and CAT growing in popularity for a range of psychological difficulties (Ryle, Kellett, Hepple & Calvert, in press), the evidence-base for CAT is relatively scant in comparison with some other psychotherapies (Margison, 2000, Llewelyn, 2003). In the language of CAT, this has been termed the 'uptake versus credibility' dilemma (Marriott & Kellett, 2009). The consideration of the evidence-base for any psychotherapy is a complex endeavour, as it requires the critical evaluation and assimilation of a typically diverse range of evidence across a range of outcome methodologies (Barkham, Stiles, Lambert & Mellor-Clark, 2010).

The 'hourglass model' (Salkovskis, 1995) is a widely accepted conceptualisation for guiding the evolution of a psychotherapy evidence-base through a cyclical three-stage evaluation process; 1) an emerging psychotherapy approach is initially tested under controlled conditions on small numbers of patients, 2) findings then stimulate larger and more stringent randomised controlled trials (RCTs) to assess efficacy and isolate mechanisms of change, then 3) promising findings are transported back into larger practice-based effectiveness studies to assess wider clinical utility, before reverting to testing new developments/iterations under stage one conditions again. To consider whether any psychotherapy is useful and safe requires the critical integration of evidence from each stage of the hourglass. This means adopting position of equipoise regarding externally valid practice-based evidence (PBE) from effectiveness

studies situated in routine clinical services to that of internally valid evidence-based practice (EBP) style research trials (Barkham, Stiles, Lambert & Mellor-Clark, 2010). No systematic review has been conducted to explore the state of the outcome evidence-base for CAT. Therefore, the central aims of this review were to (1) describe and evaluate the quality of extant CAT outcome studies, (2) fit this evidence to the 'hourglass model' to assess congruence with an established model of psychotherapy development, (3) summarise findings and finally (4) stimulate CAT outcome research.

CAT: theory & practice

CAT integrates analytic and cognitive models to offer a time-limited (usually either 16 or 24 sessions, plus follow-up), collaborative and relational approach to therapeutic change (see Kerr, 2005; Ryle & Kerr, 2002). CAT draws on personal construct theory (Kelly, 1956) and object relations theory (Ogden, 1983; Ryle, 1985) to state that representations of self, others and the world are socially formed by early reciprocal interactions with significant others (Ryle & Kerr, 2002). These representations are internalised as '*reciprocal roles'* and problematic patterns of interactions with others and limited/damaging self-care are maintained by '*target problem procedures'* (commonly termed 'traps', 'snags' and 'dilemmas'; Ryle & Kerr, 2002). If severe neglect or abuse occurs, then such reciprocal roles can become dissociated into separate self-states; conceptualised as the CAT multiple self-states model (Ryle, 1997). CAT adopts a phased approach to change of *reformulating*, *recognising* and then *revising* target problems. Specific CAT tools (i.e. psychotherapy file, narrative reformulation, sequential diagrammatic reformulation and goodbye letters) are matched to demands of the phases of therapy (see Ryle & Kerr, 2002).

Over the past decade in CAT, the influence of Leiman's introduction of the ideas of Vygotsky and Bakhtin (Leiman, 1994) and further evidence examining the interplay of biological and social influences on psychopathology have established a firm dialogical perspective on the self. Human biological evolution has progressed in an evolving social context and created a marked readiness for social formation (Ryle, 2001). Aitken and Trevarthen (1997) stated "*the dependence of the child on co-operative understanding and cultural learning is part of human genetic inheritance*" and this is *"firmly grounded in the developmental neurobiology of the infant."* CAT's cornerstone of the dialogical-self has two key implications for therapy; (1) learning takes place in the zone of proximal development and (2) learning takes place through the development, use and internalization of cultural signs and tools.

Method

Search terms and inclusion/exclusion criteria

An electronic literature search of PsycInfo, Medline and CINAHL was conducted (using the search term "*cognitive analytic**") that identified 250 papers published between 1960 and 2013. Studies were selected based on the following criteria: 1) individual or group CAT delivered, 2) use of psychometrically sound outcome measures, 3) at least pre-post outcome scores available, 4) written in English, 5) published or accepted for publication in a peer-reviewed journal and 6) independent datasets reported. Accordingly, the following papers were excluded; 9 non-English language papers, 4 unpublished theses, 93 books/book chapters/book reviews, 30 papers did not cite CAT and 89 CAT papers reported

insufficient psychometric outcomes and/or used wholly qualitative methodologies. A final sample of *N*=25 studies was retrieved for inclusion in the review.

Quality ratings

Two quality ratings were made on each study to consider the quality of the CAT evidence-base as a whole (Downs & Black, 1998) and according to individual study methodology (CASP, 2010). The Downs and Black (1998) tool is a checklist to assess the reliability and internal/external validity of an outcome study that is suitable for the both randomised and non-randomised methodologies. This checklist enables calculation of an overall quality score for each paper (0-32) and facilitated a systematic comparison of the methodological quality across CAT outcome studies. Studies were therefore compared against a published mean score of 14 (SD=6.39) for randomised and 11.7 (SD=4.64) for non-randomised studies (Downs & Black, 1998), with study scores ≥17 points deemed to be of high methodological quality (Brouwers, Johnston, Charette, Hanna, Jadad & Browman, 2005). Secondly, the Critical Appraisal Skills Programme (CASP, 2010) was used to apply targeted criteria to evaluate the methodological quality within the differing categories of outcome study that comprise the CAT evidence base. The CASP therefore provides assessment tools that grade the applicability, reliability and validity of outcome study according to their original methodology (e.g. RCTs, case controlled studies and so on). The intraclass correlation coefficient for the Downs & Black (1998) scale was 0.98 (95% C.I.=0.67 to 1.00) and 0.96 (95% C.I.=0.49 to 1.00) for the CASB. This evidenced excellent inter-rater reliability between total quality rating scores of three randomly selected papers across two raters (Field, 2005).

The N=25 CAT studies meeting the inclusion criteria are summarised in Table 1,
reporting total quality scores for each rating scale. Studies are clustered by
research methodology consistent with each stage of the 'hourglass' model
(Salkovskis, 1995) and consisted of 5 RCTs, 4 SCEDs, 11 effectiveness studies and 5
case studies. CAT case studies were of poor methodological quality (M=8.16, range
4-12 and 0/5 studies being rated as high quality). The N=4 SCEDs had a mean of
16 (range 13-20 and 2/4 rated as high quality), the N=5 RCTs had a mean score of
22 (range 20-26 and 5/5 studies rated as high quality) and the remaining N=11
quasi-experimental/effectiveness studies had a mean of 16 (range 11-24 and 6/11
rated as high quality). Overall, 52% (13/25) of the CAT outcome studies met the
criteria for high quality outcome research. This was a pattern largely reflected in
the CASP scoring, although there were some discrepancies between quality scores
on some specific studies. Table 1 highlights that the CAT outcome evidence to date
is predominated by small *n,* uncontrolled practice-based methodologies, with
eleven studies (42.30%) focal to the treatment of Personality Disorder (PD).
Clustering studies according to the hourglass model (Salkovskis, 1995) highlighted
the lack of chronologically coherent and co-ordinated research endeavours both
across and within diagnostic categories.

Insert table 1 here please

Personality Disorder (PD)

An early (low quality) paper reported clinical outcomes in the context of a
case description of a patient with BPD, suggesting that CAT was associated with

improvements in interpersonal functioning, reduced global distress and dissociation, with changes maintained at follow-up (Ryle & Beard, 1993). However, the only reliable conclusion that can be drawn from this paper is that CAT appeared helpful for that patient, at that time. More relevant (high quality) evidence came from Duignan and Mitzman's study (1994) of a combination of individual and group CAT. The study had high external validity and demonstrated statistically significant change across a range of outcome measures between start and 1-month follow-up (N=7), although no rates of reliable and clinically significant change were reported (Jacobson & Truax, 1991). However, selection bias may have feasibly influenced clinical outcomes, and the study's methodology made it impossible to extrapolate unique effects of either individual or group intervention. Ryle and Golynkina's (2000) study (high quality) provided further encouraging evidence; 52% (n=14) of BPD patients 'improved' and 22% (n=6) achieved some level of change at 6-months follow-up. Again the study had high external validity, recruiting from a population in clinical practice. However, although consideration was given to diagnostic validity within the study, the allocation process was poorly detailed and confidence in the results undermined as therapists' expertise or competence was not sufficiently controlled for. Furthermore, confidence in attributing any changes to the intervention was significantly reduced as measures were only completed prior to assessment and again at follow-ups. Wildgoose et al.'s (2001) BPD case series (high quality) measured dissociation, personality fragmentation, global distress and interpersonal functioning (N=5). At 9-month follow-up, all participants had reduced BPD severity to the extent that four patients were considered 'recovered.'

Kellett et al. (2013) used a mixed-method repeated measures design (high quality) to evaluate CAT with N=17 BPD patients. Four patients experienced

clinically significant and reliable change, three patients a reliable improvement, and one patient reliably deteriorated. Analysing outcomes at the group level showed statistically significant reductions in risk, dissociation and psychological distress, with psychological improvements occurring early in treatment. This study benefited from assessing treatment fidelity using the CAT competency measure (CCAT; Bennett & Parry, 2004) indicating that 93% of sessions (N=70) were competently delivered. Furthermore, patients qualitatively attributed various personal changes to their CAT therapy. However, the lack of a contemporaneous control condition, restricted diagnostic certainty, selection bias and too few BPD-specific outcome measures, all compromised the internal validity of the study. Nonetheless, this study provides the most rigorous and relevant evidence to date to suggest that CAT delivered under routine clinical conditions can be effective for BPD.

A controlled (but low quality) study compared CAT (N=17) with Brief Psychodynamic Therapy (BPT N=17) delivered by trainee therapists (Mace et al., 2006). Patients were allocated to treatment condition following independent assessment and matched on a range of variables. CAT and BPT produced similar statistically significant improvements. Six CAT patients achieved clinically significant improvement compared with 13 of the BPT patients, although the CAT patients were significantly more distressed at assessment and twice as many patients allocated to CAT were diagnosed with PD. The study design limits the confidence in results, as although an attempt was made to control for therapist effects, valid comparisons between conditions is not feasible given the group differences at baseline and the lack of clarity regarding the 'dose' of therapy in each condition. Measures were only taken at assessment and 3-month follow-up thereby reducing the validity of attributing change to interventions. Furthermore,

although the paper reports rates of clinically significant change, this criterion can inflate recovery rate and does not take into account inherent measurement error. Overall the study contributes further evidence that CAT can be an effective intervention, although the quality does not convincingly demonstrate the benefits of CAT over and above BPT.

More recently Chanen et al., (2008) completed a high quality RCT comparing CAT (N=41) with manualised 'good clinical care' (N=37) for adolescents displaying BPD features.   Outcomes were collected at baseline, 6, 12 and 24-month follow-up, with 92% (N=72) of participants completing outcomes for at least three time-points. Results indicate that both CAT and GCC were efficacious. However, it was difficult to attribute change to either intervention, as both were delivered as an adjunct to a comprehensive treatment package.   Therefore, Chanen et al. (2009) reanalysed outcomes for the CAT and GCC with adolescents who received 'Historical Treatment As Usual' (H-TAU, N=32). At 2-year follow-up, CAT produced the most marked improvement in externalising difficulties and parasuicidal behaviour and had the fastest rate of improvement in internalising and externalising difficulties. There was no difference between conditions in the rate of improvement observed in borderline psychopathology or parasuicidal behaviour, nor the frequency of service utilisation during treatment. The study had high external validity, controlled for therapist effects and standardised treatment packages, thereby providing strong evidence for CAT as a helpful early intervention for BPD. Contemporaneously randomising participants to a 'treatment-as-usual' condition and randomly allocating patients to therapists would have enhanced the validity of findings. Replication on a larger-scale with longer-term follow-up is certainly warranted, and confidence in findings would be enhanced with reporting of rates of reliable and clinically significant change.

Clarke et al., (2013) conducted a (high quality) RCT comparing the efficacy of 24 sessions of CAT for patients with a broader range of personality disorders (N=38) with TAU (N=40). Participants were randomised according to whether the patient met diagnostic criteria for each PD clusters (A = 0, B = 18, C = 28 and mixed = 55). All patients in the trial had received at least one previous episode of therapy. Post-therapy 33% (9/27) of the patients completing CAT no longer met diagnostic criteria for any personality disorder. In TAU, all patients (100%, 33/33) continued to meet criteria for at least one personality disorder, with evidence of continuing personality deterioration in 53% (16/33). Reliable change scores in the CAT group noted that 42% (15/35) had either improved or recovered. Limitations of the Clarke et al, (2013) trial are the small sample size, the absence of Cluster A personality disorder patients, the exclusion of patients exhibiting self-harm behaviours and lack of systematic data collection about TAU. However, the study was well designed and provides further evidence that a structured psychotherapy, such as CAT, is superior to standard care for treating a broad range of personality difficulties seen in community settings.

Single-case experimental designs (SCED) can contribute important clinically relevant data to the CAT evidence-base. Such studies entail the collection of time series data over various phases of treatment (and at times treatment withdrawal and re-introduction) and follow-up which are then compared against the patient baseline functioning (Kazdin, 2010). Kellett (2007) completed a (low quality) A/B with extended follow-up SCED, with a patient diagnosed with Histrionic Personality Disorder (HPD). The DSM-IV (APA, 2000) matched experimental variables were measured daily across baseline (21 days), treatment (182 days) and follow-up (154 days) phases. All HPD measures significantly reduced, with more than a 40% reduction in histrionic tendencies during intervention. Clinically

significant pre/post improvements were observed. However, the study baseline also formed the assessment 'pre-reformulation' phase of CAT and can therefore be criticised for failing to achieve a technically neutral baseline. Although CAT was evidently helpful for this patient, any N=1 study has limited generalizability.

Kellett & Hardy (2013) completed a (high quality) mixed-method SCED concerning the treatment of a patient with Paranoid Personality Disorder (PPD). The patient kept a daily diary of six key paranoia measures throughout assessment (3 sessions), treatment (21 sessions) and follow-up (4 sessions). Psychometric outcome measures were collected at assessment, termination and follow-up. Analysis of the time series daily data illustrated significant reductions in suspiciousness and anxiety, with a significant increase in problem solving during treatment sessions. Graphing of the daily variables noted that 5 out of the 6 SCED variables had extinguished during treatment. The patient was also independently interviewed with the 'Change Interview' (Elliott, 2002), providing subjective evidence that changes achieved were attributed to CAT. Results suggest that CAT was an effective intervention in this case of PPD, although again the study suffers from methodological limitations commonly associated with SCED; the questionable reliability of self-report measures, unique patient and therapist characteristics and the nature of SCED being an overly individualised evaluation.

CAT PD studies detail evidence that suggests that CAT can produce good outcomes for patients with personality disorder both in routine clinical practice and under trial conditions. Eight of the eleven CAT PD studies were high quality. Indeed, CAT has been included as a potential treatment in NICE guidelines for Borderline Personality Disorder (BPD; NICE, 2009a).

Anxiety and depression

Bennett's (1994) early case description (low quality) indicated that 16-session CAT had a positive impact on depression, global functioning and interpersonal difficulties, with change maintained at 3-month follow-up. A further (low quality) case description by Hamill and Mahoney (2011) also suggested that 16-session CAT with follow-up was helpful in reducing depression, anxiety and physical health complaints in a person caring for a relative with dementia. However, both of these papers lack any internal rigour and offer only weak contributions to the CAT evidence-base for treating affective difficulties.

An evaluation (low quality) of referrals to a CAT clinic (Dunn, Golynkina, Ryle and Watson, 1997) noted that 58% of referrals were for minor depression and anxiety. Of those patients who attended for follow-up, highly significant pre-post improvements in interpersonal and psychological functioning were reported, but no significant improvements were observed in social functioning. This was not a methodologically rigorous study, but rather a description of outcomes in routine clinical practice. It is important to recognise that results are from a biased follow-up sample, with limited diagnostic validity, with no attempt was made to analyse outcomes for those prematurely ending therapy. Despite these limitations, the results are from a relatively large sample with high external validity. The relevance to other settings is unclear given the poor detailing of service and therapist characteristics.

A high quality study by Brockman et al. (1987) randomly allocated 48 patients to either CAT (N=30) or 'interpretative therapy' (INT, N =18; Mann & Goldman, 1982), delivered by trainee therapists. Both therapies produced improvements in depression and general mental health, with CAT patients also experiencing a significantly improved self-attitude. There was no difference

between conditions in participants' subjective ratings of change experienced during therapy. At follow-up, CAT patients demonstrated further improvements in depression scores, but deterioration in general well-being between end of therapy and follow-up. It is important to note that the duration of follow-up varied widely. Furthermore, it is difficult to make valid comparisons between treatments as although diagnostic groupings appeared similar at intake, CAT patients demonstrated greater distress at baseline. Also, the more complex cases were allocated to more experienced therapists, with this not taken into account in subsequent analysis. The lack of measurement of treatment adherence and the fact that an experienced CAT practitioner provided weekly group supervision across modalities weakens the internal validity of the study. Birtchnell et al's study (2004) also explored CAT delivered in routine clinical practice (low quality rating). Within the constraints of severe methodological and reporting limitaions, the study illustrated that 16-session CAT produced a significant improvement on two measures of interpersonal functioning - with change maintained at 3-month follow-up. The study failed to report patient and service characteristics, or the validity of the primary outcome measure and so it is exceedingly difficult to generalise findings.

In a high quality study, Marriott and Kellett (2009) benchmarked short-term (median sessions=16; N=38) and medium-term (median sessions=24, N=27) CAT against short and medium-term CBT and person centred therapy (PCT) in routine clinical practice. Despite more distressed patients being allocated to CAT, patients were matched across modalities on a score of global distress at intake and number of sessions received. Results indicate that all modalities were effective in reducing distress across a number of measures. Recovery rates were significantly higher during short-term CBT, but equivocal across medium-term therapies.

Those patients receiving longer-term CAT were more likely to achieve recovery than those receiving less than 16-sessions of CAT. The study suffers from the range of methodological limitations inherent to practice-based evidence; no randomisation to modality or therapist, lack of diagnostic validity and no measure of treatment adherence.

Table 1 indicates a paucity of evidence for CAT targeting common mental health problems, despite the impetus to evaluate non-CBT therapies for anxiety and depression (Care Services Improvement Partnership Choice & Access Team, 2008; NICE, 2009b; Department of Health, 2011). Of the six studies completed, four were low quality. Results across studies can only suggest that CAT offers some promise as an effective intervention for this population, with a more detailed understanding clearly indicated.

Eating Disorders (ED)

A high quality pilot RCT by Treasure et al. (1995) randomised participants to either CAT (n=14) or 'educational behavioural treatment' (EBT; n=16). No significant differences in outcomes were observed between treatments at 1-year follow-up; both resulted in an average weight gain of 6.8kg with 30% of treatment completers maintaining weight gain, although participants in the CAT condition subjectively reported greater improvement. Poorer outcomes were predicted by a greater proportion of pre-treatment weight loss. Treasure et al., (1995) concluded that both outpatient CAT and EBT can be effective for adult onset AN. It is worth considering that the results may be artefact of having the same clinicians deliver both treatments; although both treatments were manualised, no measure of model adherence was employed. Furthermore, the study had a small sample size and therefore suffered from limited power to differentiate between treatments.

The other CAT high quality RCT with an ED population randomised patients to either 'focal psychoanalytic psychotherapy' (FPP, n=12), 'family therapy' (FT, N=16), CAT (N=13) or treatment as usual (TAU, N=13; Dare, Eisler, Russell, Treasure & Dodge, 2001). There were no significant differences in engagement rates between active therapies, and of the original sample, 64% completed treatment with significantly more non-completers in the TAU condition. Results at 1-year follow-up (controlling for initial weight) showed that across active therapies one third of patients no longer met diagnostic criteria - compared with 5% of the TAU sample. Intention-to-treat analysis revealed that 32% (N=7) of the CAT participants achieved a good outcome, compared with 52% (N=11) of those in FPP and 41% (N=9) in FT. Taken together, findings suggest that CAT is superior to TAU, but that FPP and FT appear to achieve better outcomes than CAT. Results should be interpreted with caution, as no measures of model fidelity were taken, sample sizes were small, there were inconsistent treatment contracts and therapist competence may have biased findings (FPP and FT were delivered by trained, experienced clinicians, whereas CAT was delivered by non-accredited therapists supervised by an experienced CAT clinician).

Table 1 indicates there is promising evidence for the use of CAT for ED from two controlled studies treating Anorexia Nervosa (AN), which contribute to the NICE guidelines for managing eating disorders (2004). Both trials were high quality.

Survivors of Childhood Sexual Abuse (CSA)

Two small scale and low quality practice-based studies report outcomes of 16-sessions of CAT for female (Clarke & Llewelyn, 1994; N=6) and male CSA survivors (Clarke & Pearson, 2000; N=4). Clarke and Llewelyn (1994) collected

pre, post and 3-month follow-up outcomes. Following CAT, five patients

demonstrated reliable improvement in global functioning, although only two

scored below the clinical cut-off at end of CAT. Scores also indicated

improvements in depression, self-esteem and reduced self-blaming beliefs/self-

harming behaviour. Broadly, change appeared maintained over 3-month follow-

up, although there was some indication of mood relapse.   Despite evident

symptomatic relief, repertory grid methodology demonstrated little change in the

interpersonal constructs of CAT patients. Clarke and Pearson (2000) replicated the

study with male survivors of CSA, with all participants demonstrating a reduction

in self-blaming beliefs about their early abuse experiences and reduced depression

scores. Overall, levels of global distress reduced, but two participants with BPD

reported increased psychological distress following termination of CAT. The

authors conclude that the results corroborate previous findings that 16-session

CAT maybe too brief to effect change for patients with highly complex emotional

difficulties (Mace et al., 2006; Wildgoose et al., 2001).

Both studies had significant methodological limitations; treatment was

individualised and therefore difficult to generalise from and no diagnostic validity

was employed in either study undermining validity, given the heterogeneous

population that survivors represent. Given the small sample sizes, it is only

possible to tentatively suggest that results indicate some initial evidence to suggest

that CAT may be a useful intervention for CSA survivors experiencing a range of

difficulties.

Dissociative Disorders

Graham and Thavasothy's (1995) case description (low quality) of a very

brief intervention, 5-session CAT, noted a reduction in the frequency and severity

of dissociative experiences associated with an episode of dissociative psychosis. Progress was sustained at 2-year follow-up. However, the CAT intervention was poorly described and arguably comprised part of a wider inpatient treatment package, making it impossible to conclusively attribute any changes to CAT. Higher quality evidence for CAT with gross dissociation comes from Kellett's (2005) SCED with a patient diagnosed with Dissociative Identity Disorder. Seven experimental dissociative measures rated daily throughout baseline (35 days), intervention (175 days) and follow-up (168 days) phases demonstrated a reduction to intensity of state dissociation and increasing awareness of identity shifts during treatment. Specific changes in dissociative variables were associated with specific CAT interventions (such as completion of the diagrammatic reformulation) and change was shown to be maintained over the follow-up. Reliable and clinically significant improvements in global functioning, depression and personality integration occurred between assessment and termination. However, insufficient attention was paid to possible confounding variables and the study suffers from methodological limitations inherent within SCED outlined earlier. There is therefore currently a lack of convincingly sound evidence for the utility of CAT for dissociative difficulties.

Morbid Jealousy

Kellett and Totterdell (2013) used matched (low quality) SCED designs (N=2) to compare CAT with Cognitive Behavioural Therapy (CBT) for the treatment of morbid jealousy (obsessive subtype). Five experimental jealousy measures were rated throughout baseline (CAT=35 days; CBT=44 days), intervention phase (CAT=98 days; CBT=51 days) and follow-up phases (84 days for both). Patients' partners also returned contemporaneous daily ratings of two

target difficulties. Reliable and clinically significant pre-post improvements were demonstrated in levels of global functioning, depression and interpersonal difficulties for the CAT patient. Change was maintained for the CAT patient at 3-month follow-up, compared with evidence of mood relapse in the CBT patient. Whilst there were significant improvements in jealousy, hyper-vigilance, anxiety and self-esteem, the observed improving trend in the baseline reduces the confidence of the results being attributed purely to CAT. The partner of the CAT patient did not report any significant subjective improvements in the patient's behaviour. Although the 'dose' of therapy was controlled for, the frequency of sessions was not and the same therapist delivered both interventions (thus controlling for therapist effects), but no measure of model adherence was employed. Such limitations impact on the internal validity of findings, but represent the reality of conducting practice-based N=1 SCED research. It is premature to draw any firm conclusions on the utility of CAT in the treatment of morbid jealousy.

Long-term physical health conditions

An uncontrolled (low quality) case description suggested that CAT (as an adjunct to concurrent cognitive rehabilitation) can be effective in reliably reducing anxiety and anger, and improving interpersonal functioning, approximately 2-years following an acquired brain injury (Yeates et al., 2008). However, the study design was poor as it did not detail the severity of the patient's brain injury, or the intervention delivered. Therefore, no wider valid inferences can be drawn on the utility of CAT for patients with dysexecutive difficulties.

Fosbury et al's (1997) high quality RCT randomly allocated diabetic patients to CAT (N=15) or a 'diabetes specialist nurse education' programme

(DSNE; N=17) with the aim of improving participants' self-care regimes. CAT produced significant change in patients' knowledge of diabetes, whereas DSNE was shown to be effective in reducing blood glucose levels. Of the original sample, 81% (N=26) completed measures at 9-month follow-up, with a greater rate of attrition in the CAT condition (33%; N=5). At follow-up, both interventions achieved significant improvement in glucose levels and diabetes knowledge, with CAT effecting greater interpersonal change. CAT appeared to produce more durable change as those in the DSNE condition demonstrated a relapsing trend following intervention. Although benefiting from randomisation of patients and high external validity, the study had a small sample size and 18 potential participants refused to take part, thus limiting the representativeness of the research sample.

It is difficult to draw any strong conclusions regarding the effectiveness of CAT with physical health conditions, because of the large differences in outcome methodologies employed and the disparate populations studied.

Discussion

This review suggests that there is growing evidence for the utility of CAT under routine clinical practice and trial conditions across a diverse range of presenting difficulties. The results indicate that the CAT treatment evidence-base is currently small, but notwithstanding the qualitative case descriptions, the evidence consists of relatively well-conducted outcome studies. Over half the outcome studies conducted were of high quality. It appears that the CAT evidence-base is yet to achieve scientific credibility when evidence-based practice and practice-based evidence sources are considered in equipoise, simply due to the lack of number/weight of studies across and within diagnostic categories. In no

one single diagnosis has the CAT evidence yet completely progressed successfully and appropriately through the hourglass. In terms of personality disorders, the accumulating evidence suggests that CAT has a major contribution to make in terms of the treatment of patients with personality pathology in front line clinical services (Mulder & Chanen, 2013). The challenge is now to benchmark the effectiveness of CAT for PD under routine care conditions, via large-scale service evaluations and clinical audits. This review has also highlighted the dearth of rigorous and reliable evidence for CAT with common mental health difficulties.

Margison (2000) commented that, *"CAT is unusual in being increasingly widely practised without following the full three-stage model* [hourglass model] *of development"* (p.146). Over one decade later, the development of CAT outcome research still remains somewhat incoherent, with little evidence of any strategic and chronological progression through an established framework of psychotherapy development and evaluation. The majority of studies comprising the evidence-base utilised practice-based designs, with small sample sizes. The paucity of controlled CAT studies means that it is therefore difficult to infer any firm conclusions with true confidence. Despite RCTs often being promoted as the 'gold standard' of outcome research, the methodology does have some inherent weaknesses (Williams, 2010) and the future CAT evidence base needs to balance the dual development of PBE and EBP approaches. The CAT evidence is also marked by inadequate standards of reporting of patient and therapist characteristics in the extant PBE studies. Failure to initially generate more evidence has been due to choosing to expand the CAT workforce and associated training and supervision endeavours, which has been prioritised at the expense of establishing an academic base (Ryle et al., in press). The lack of a substantive evidence-base should however *in no way* be equated with ineffectiveness. Rather

CAT outcome research should be viewed as a nascent endeavour, clearly requiring urgent support, attention and energy.

Whilst CAT was originally developed as a trans-diagnostic treatment (Margison, 2000), evidence has highlighted that CAT is being often selected to treat patients with more complex and severe difficulties (e.g. Denman, 2001). The current systematic review substantiates this finding in two additional ways; (a) the majority of published CAT outcome studies have been completed with patients with more severe difficulties (e.g. PD), and (b) in routine practice, patients with PD are more likely to be allocated to CAT by independent assessors (Brockman et al., 1987; Mace et al., 2006; Marriott & Kellett, 2009). The logic for such assessment outcomes appears the matching of a relational therapy for patients with chronic problems with relating.

The CAT approach is avowedly relational (Ryle et al., in press) and therefore to enable patients to reflect on (and then change) their interpersonal patterns and roles both within and outside of the therapeutic relationship, this demands a high degree of interpersonal skill. The bedrock of such skills lays in the formation and maintenance of effective therapeutic alliances (from screening to discharge) and the common factors that facilitate engagement, trust and durable change. The CCAT competency measure (Bennett & Parry, 2004) is reflective of the centrality of common factors with the ten domains of competency being: (1) phase-specific tasks (such as engagement skills in early CAT sessions), (2) making theory-practice links, (3) CAT tools (such as narrative and diagrammatic reformulation), (4) managing boundaries, (5) common factor skills, (6) collaborative climate, (7) assimilation of warded-off or problematic states, (8) making links and hypotheses, (9) managing threats to the therapeutic alliance and (10) awareness and management of therapist's own reactions/feelings. Regular use of the CCAT is

indicated in terms of patient safety and also the personal/professional development of the therapist, via clinical supervision.

A recent adaptation of the CAT model is for the approach to be used to help clinical teams manage those patients that are unsuitable for individual or group talking treatments (Kerr, Dent-Brown & Parry, 2007). The development of the CAT consultancy approach enables CAT theory to (1) map the dysfunctional roles and procedures that often arise between patient and clinical team, (2) enable a 'common language' to describe such reciprocations to emerge in the team and (3) to identify new ways of working (e.g. noticing and reducing unhelpful and potentially iatrogenic negative interactions). The process for this type of CAT work has been manualised (Carradice, 2012) and recent RCT evidence indicates its positive impact on organisational processes and the individual clinical practice of team members (Kellett, Wilbram, Davis & Hardy, in press).

Future indicated research strategy

Co-ordination of future CAT outcome research should be corralled within single diagnostic categories to enable more depth and breadth of evidence to be generated (Barkham & Mellor-Clark, 2003). Such concerted efforts are particularly indicated, as a lack of compelling evidence-base may preclude future public investment. It is apparent that CAT is popular with therapists and is eagerly taken up in clinical practice (Ryle et al., in press). Therefore, future large-scale pragmatic trials (Goodyer et al., 2011) may offer a methodology matched to the aspirations of CAT in evaluating outcomes in clinical practice, incorporating longer-term follow-up and benchmarking outcomes against other modalities (Lueger & Barkham, 2010). Like all psychotherapies, CAT needs to demonstrate the health economic value of the approach. CAT case descriptions suffered from a

lack of rigour and were of poor methodological quality. Any further such case study research needs to be considered as only useful in under-represented clinical populations at the initiation of a diagnostic hourglass.

As CAT has specific interventions at specific points during therapy (e.g. early narrative reformulation), deconstruction trials/component analyses (Ahn & Wampold, 2001) would also usefully index the agency of such specific CAT tools. All future CAT outcome studies should routinely report rates of reliable and clinically significant change (Jacobson & Truax, 1991) and index fidelity to the CAT model (e.g. CCAT; Bennett & Parry, 2004). Studies should also explore the 'dose' response effect of CAT, why some patients deteriorate during CAT and what it is about the approach that appears to preclude the low attrition rates observed in other psychotherapies (Kellett et al, 2013). As CAT is being increasingly delivered via groups in services (Ryle et al., in press) the efficiency and efficacy of group CAT delivery demands evaluation. The CAT consultancy approach is novel and also demands further controlled evaluation (Kellett et al, in press). Both group CAT and CAT as a consultation tool need to development of specific competency/fidelity measures to support therapists in this work.

Future CAT SCED studies should attempt to measure problem frequency in a baseline that is established prior to the pre-reformulation sessions and adjudicated hermeneutic single case efficacy design will also be useful in defining CAT efficacy at the N=1 single case level (Elliott, 2002). Adjudicated hermeneutic designs entail adopting a legalistic approach to evaluation, with many strands of evidence being debated by opposing briefs (sceptical and affirmative) and a final opinion regarding treatment efficacy being assigned by 'judges' who are experienced psychotherapy researchers (Elliott, Partyka, Alperin, Dobrenski, Wagner, Messner, Watson & Castonguay, 2009). Within quantitative outcome

designs, there is also much room for process studies to identify how therapeutic change is brought about by CAT therapists. For example, studies defining how 'exits' (i.e. CAT change methods) are negotiated, co-constructed, practiced and evaluated with patients are required.

A finding of the current review is that the CAT evidence-base tends to be limited by two key factors, (1) small sample sizes and (2) lack of RCTs. In relation to the former, then aspiring CAT clinician-researchers should pay specific attention to the issue of sample size and the generation of adequate statistical power. Due to the popularity of CAT in routine practice, the generation of large-scale service evaluations of the effectiveness of the model in frontline clinical services is realistic. As noted, this is now particularly indicated in the area of PD, as accumulated RCT evidence does now attest to the efficacy of CAT with such patients. In relation to common mental health problems, there remains a curious absence of trials examining the efficacy of CAT. The expansion of IAPT and the ready provision, in particular, of cognitive-behaviour therapy in Primary Care for people with anxiety and depression should enable large-scale comparative trials to be a reality.

There is a clear need for the CAT community to establish systems for co-ordinating coherent research strategies that engage and encourage CAT therapists into research work. Using the CAT model, it may be the case that CAT therapists are caught in the dilemma of: *you either are a therapist or a researcher.'* The wide use of SCED methodologies with CAT patients provides examples of a 'research exit' from this dilemma, as it is achievable by the single-handed clinician within a scientist-practitioner framework (Kazdin, 2010). An example of an unhelpful research reciprocal role for CAT therapists might be '*overwhelming* to *anxious'* about the research process, which would fuel ambivalence at best - and active

avoidance at worst. Given the popularity of the approach, CAT practice research networks (Castonguay et al. 2010) are currently under-utilised and may provide an exit from the unhelpful reciprocal role, to that of one of '*including* to *energised*.' Collaborating with the wider psychotherapy research community would be useful, particularly in setting up valid treatment comparisons and learning about how to establish and nurture research/evaluation cultures in front line clinical services.

Conclusions

Outcome evidence to date suggests that CAT is a promising intervention across a range of diagnostic groups. Although the nascent CAT evidence-base lacks the additional weight and rigor of sufficient efficacy trials, the predominance of practice-based evidence does index the relevance and uptake of CAT. This review highlighted a trend for complex patients to be allocated to CAT in routine practice (Brockman et al., 1987; Mace et al., 2006; Marriott & Kellett, 2009). Denman (2001) noted that CAT therapists tend to specialise in treating more severe difficulties and this process is reflected in the populations and outcomes studied to date. More CAT attention should be focused on the needs and outcomes of patients with common mental health difficulties and the associated develop of lower intensity versions of the model are indicated. In the language of CAT the 'exit' to the identified 'uptake versus credibility' dilemma is a coordinated research strategy, capable of producing large-scale efficacy and effectiveness evidence within diagnostic groups, with research streams appropriately following existing psychotherapy evaluation criteria and processes.

References

Denotes studies included in the review in Table 1

Ahn, H., & Wampold, B.E. (2001). Where oh where are the specific ingredients? A meta-analysis of component studies in counselling and psychotherapy. *Journal of Counselling Psychology, 48* (3), 251-257. doi:10.1037/0022-0167.48.3.251

Achenbach, T. (1997) *Young adult self report*. Burlington, VT: University of Vermont, Department of Psychiatry.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). doi:10.1176/appi.books.9780890423349

Aitken, K.J., & Trevarthen, C. (1997). Self/other organisation in human psychological development. *Development and Psychopathology, 9,* 653-677.

Barkham, M., Hardy, G. E., & 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

Barkham, M., & Mellor-Clark, J. (2003). Bridging evidence-based practice and practice-based evidence: Developing a rigorous and relevant knowledge for the psychological therapies. *Clinical Psychology and Psychotherapy,* 10, 319–327. doi: 10.1002/cpp.379

Barkham, M., Stiles, W. B., Lambert, M. J., & Mellor-Clark, J. (2010). Building a rigorous and relevant knowledge base for psychological therapies. In M. Barkham, G. E. Hardy & J. Mellor-Clare (Eds.), *Developing and delivering practice-based evidence*. Chichester: John Wiley & Sons.

Beck, A. T., Ward, C. M., Mendelson, M., Mock, J. E., & Erdaugh, J. C. (1961). An inventory for measuring depression. *Archives of General Psychiatry,* 4, 561-571.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology,* 56, 893-897. doi: 10.1037/0022-006X.56.6.893

Beck, A. T., Steer, R. A., & Brown, G. K. (1995). *Beck depression inventory II manual.* San Antonio, CA: The Psychological Corporation, Harcourt Brace & Co.

Beckett, R., Tarrier, N., Intili, R., & Beech, A. (1992). *The prestwich jealousy questionnaire.* Unpublished questionnaire, Manchester.

Bell, L. (2001). Psychological therapies in anorexia nervosa. *The British Journal of Psychiatry,* 179. 78-79. doi: 10.1192/bjp.179.1.78

*Bennett, D. (1994). Readiness to change - the impact of reformulation: A case example of cognitive analytic therapy. *International Journal of Short-Term Psychotherapy,* 9, 83-91.

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

Berstein, E. M., & Putnam, F. W. (1986). Development, reliability and validity of a dissociative scale. *Journal of Nervous and Mental Disease,* 12, 727-732.

*Birtchnell, J., Denman, C., & Okhai, F. (2004). Cognitive analytic therapy: Comparing two measures of improvement. *Psychology and Psychotherapy: Theory, Research and Practice,* 77, 479-492. doi: 10.1348/1476083042555398

Birtchnell, J., & Evans, C. (2004). The person's relating to others questionnaire

    (PROQ2). *Personality and Individual Differences,* 36, 125-140. doi:

    10.1016/S0191-8869(03)00073-4

Brouwers, M. C., Johnston, M. E., Charette, M. L., Hanna, S. E., Jadad, A. R.,&

    Browman, G. P. (2005). Evaluating the role of quality assessment of primary

    studies in systematic reviews of cancer practice guidelines. *BMC Medical*

    Research Methodology, 5-8. doi: 10.1186/1471-2288-5-8

*Brockman, B., Poynton, A., Ryle, A., & Watson, J. P. (1987). Effectiveness of time-

    limited therapy carried out by trainees: Comparison of two methods. *British*

    Journal of Psychiatry, 151, 602-610.

CASP UK | Making sense of evidence. (2010). Critical Appraisal Skills Programme:

    making sense of evidence about clinical effectiveness: 11 questions to help

    you make sense of a trial. [ONLINE] Available at: http://www.casp-uk.net.

    [Accessed 24 April 2013].

Castonguay, L.G., Boswell, J.F., Zack, S., Baker, S., Boutselis, M., Chiswick, N., Damer,

    D., Hemmelstein, N., Jackson, J., Morford, M., Ragusea, S., Roper, G., Spayd, C.,

    Weiszer, T., Borkovec, T.D., & Grosse Holtforth,, M. (2010). Helpful and

    hindering events in psychotherapy: A practice research network study.

    Psychotherapy: Theory, Research, Practice, and Training, 47, 327-344. doi:

    10.1037/a0021164.

Care Services and Improvement Partnership Choice and Access Team (2008)

    Improving access to psychological therapies (IAPT) commissioning toolkit.

    London: Department of Health.

Carlson, B.E., & Putnam, F. W. (1993). An update on the dissociative experiences

    scale. *Dissociation,* 6, 16-27.

Carradice, A. (2012). 5-session CAT consultancy: using CAT to guide care planning with people diagnosed with personality disorderwithin community mental health teams. *Clinical Psychology and Psychotherapy, 20,* 359-367.

*Chanen, A. M., Jackson, H. J., McCutcheon, L. K., Jovev, M., Dudgeon, P., Yuen, H. P., Gernano, D., McDougall, E. & McGorry, P.D (2008). Early intervention for adolescents with borderline personality disorder using cognitive analytic therapy: Randomised controlled trial. *British Journal of Psychiatry,* 193, 477-484. doi: 10.1192/bjp.bp.107.048934

*Chanen, A. M., Jackson, H. J., McCutcheon, L. K., Jovev, M., Dudgeon, P., Yuen, H. P., Gernano, D., Nistico, H. McDougall, E., Weinstein, C., Clarkson, V. & McGorry, P.D. (2009). Early intervention for adolescents with borderline personality disorder: Quasi-experimental comparison with treatment as usual. *Australian and New Zealand Journal of Psychiatry,* 43, 397-408. doi: 10.1080/00048670902817711

*Clarke, S., & Llewelyn, S. (1994). Personal constructs of survivors of childhood sexual abuse receiving cognitive analytic therapy. *British Journal of Medical Psychology,* 67, 273-289. doi: 10.1111/j.2044-8341.1994.tb01796.x

*Clarke, S., & Pearson, C. (2000). Personal constructs of male survivors of childhood sexual abuse receiving cognitive analytic therapy. *British Journal of Medical Psychology,* 73, 169-177. doi: 10.1348/000711200160408

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

Corney, R. H., Clare, A. W., & Fry, J. (1982). The development of a self-report questionnaire to identfy social problems - a pilot study. *Psychological Medicine,* 12, 903-909. doi: 10.1017/S0033291700049199

Crown, S., & Crisp, A. H. (1979). *Manual of Crown-Crisp experimental index*. London: Hodder and Stoughton.

*Dare, C., Eisler, I., Russell, G., Treasure, J., & Dodge, L. (2001). Psychological therapies for adults with anorexia nervosa: Randomised controlled trial of out-patient treatments. *British Journal of Psychiatry,* 178, 216-221. doi: 10.1192/bjp.178.3.216

Denman, C. (2001). Cognitive-analytic therapy. *Advances in Psychiatric Treatment,* 7, 243-252. doi: 10.1192/apt.7.4.243

Derogatis, L. (1993). *Brief symptom inventory: Adminstering, scoring and procedures manual* (3rd ed.). Minneapolis: National Computer Systems.

Derogatis, L., Richels, K., & Rock, A. (1976). The SCL 90 and the MMPI: A step in the validation of a new self report scale. *British Journal of Psychiatry,* 128, 280-289. doi: 10.1192/bjp.128.3.280

Department of Health (2011). Talking therapies: A four-year plan of action. *A supporting document to* No health without mental health: A cross-government mental health outcomes strategy for people of all ages. London: Department of Health.

Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology and Community Health,* 52, 377-384. doi: 10.1136/jech.52.6.377

*Duignan, I., & Mitzman, S. (1994). Measuring individual change in patients receiving time-limited cognitive analytic group therapy. *International Journal of Short-Term Psychotherapy,* 9, 151-160.

*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. (2002). Hermeneutic single case efficacy design. *Psychotherapy Research,
12*, 1-20. doi: 10.1080/713869614

Elliott, R., Partyka, R., Alperin, R., Dobrenski, R., Wagner, J., Messner, S.B., Watson,
J.C. & Castonguay, L.G. (2009). An adjudicated hermeneutic single-case
efficacy design study of the experiential therapy of panic/phobia.
Psychotherapy Research, 19, 543-557.

Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J., &
Audin, K. (2002). Towards a standardised brief outcome measure:
Psychometric properties and utility. *British Journal of Psychiatry,* 180, 51-
60. doi: 10.1192/bjp.180.1.51

Field, A. P. (2005) Intraclass correlation. In B.S. Everitt, & D.C. Howell *Encyclopedia
of statistics in behavioral sciences*. Chichester: Wiley.

First, M.B., Gibbon, M., Spitzer, R.L., & Williams, J.B.W. (1996). User's guide for the
structured clinical interview for DSM-IV Axis II disorders- research version.
Biometrics Research.

*Fosbury, J. A., Bosley, C. M., Ryle, A., Sonksen, P. H., & Judd, S. L. (1997). A trial of
cognitive analytic therapy in poorly controlled type I patients. *Diabetes
Care,* 20, 959-964. doi: 10.2337/diacare.20.6.959

Goldberg, D. P. (1972). *The detection of psychiatric illness by questionnaire*.
Maudsley Monograph No. 21: Oxford University Press.

Goldberg, D. P., & Williams, P. (1988). *A user's guide to the GHQ*. Windsor: Nfer-
Nelson.

Goldman, H., Skodol, A., & Lave, T. (1992). Revising axis V for DSM-IV: a review of

measures of social functioning. *American Journal of Psychiatry,* 149, 1148-

1156.

Goodyer, I.M., Tsancheva, S., Byford. S., Dubicka, B., Hill, J., Kelvin, R., Reynolds, S.,

Roberts, C., Senior, R., Suckling, J., Wilkinson, P., Target, M. & Fonagy, P.

(2011) Improving mood with psychoanalytic and cognitive therapies

(IMPACT): A pragmatic effectiveness superiority trial to investigate

whether specialised psychological treatment reduces the risk for relapse in

adolescents with moderate to severe unipolar depression: study protocol

for a randomised controlled trial. *Trials,* 12, 175. doi: 10.1186/1745-6215-

12-175.

*Graham, C., & Thavasothy, R. (1995). Dissociative psychosis: An atypical

presentation and response to cognitive-analytic therapy. *Irish Journal of

Psychological Medicine,* 12, 109-111.

*Hamill, M. & Mahony, K. (2011). 'The long goodbye': Cognitive analytic therapy

with carers of people with dementia. *British Journal of Psychotherapy*, 27,

292-304. doi: 10.1111/j.1752-0118.2011.01243.x

Horowitz, L. M., Rosenberg, S. E., Baer, B. A., Ureno, G., & Villesenor, V. S. (1988).

Inventory of interpersonal problems: Psychometric properties and clinical

applications. *Journal of Consulting and Clinical Psychology,* 56, 885-892. doi:

10.1037/0022-006X.56.6.885

Jacobson, N. S., & Truax, P. (1991). Clinical significance: a statistical approach to

defining meaningful change in psychotherapy research. *Journal of

Consulting and Clinical Psychology,* 59, 12-19. doi: 10.1037/0022-

006X.59.1.12

Jehu, D. (1988). *Beyond sexual abuse: therapy with women who were victims in childhood.* Chichester: Wiley.

Kazdin, A.E. (2010). *Single case research designs: Methods for clinical and apllied settings* (2nd edition). NY, Oxford Univerity Press.

*Kellett, S. (2005). The treatment of Dissociative Identity Disorder with cognitive analytic therapy: Experimental evidence of sudden gains. *Journal of Trauma & 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: Theory, Research and Practice,* 80, 389-405. doi: 10.1348/147608306X161421

*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: Theory, Research and Practice*, 86, 52-69. doi: 10.1111/j.2044-8341.2011.02045.x

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

*Kellett, S. & Hardy, G. (2013). Treatment of Paranoid Personality Disorder with cognitive analytic therapy: A mixed methods single case experimental design. *Clinical Psychology and Psychotherapy.* doi: 10.1002/cpp.1845.

Kellett, S., Wilbram, M., Davis, C., & Hardy, G. (in press). Team consultancy using cognitive analytic therapy: A controlled study in Assertive Outreach. *Journal of Psychiatric and Mental Health Nursing.*

Kelly, G. (1956). *The psychology of personal constructs.* New York: Norton.

Kerr, I.B. (2005). Cognitive analytic therapy. *Psychiatry,* 4, 28-33. doi:

    10.1383/psyt.4.5.28.65105

Kerr, I.B., Dent-Brown, K., & Parry, G. (2007). Psychotherapy and mental health

    teams. International Review of Psychiatry, 19, 63-80.

Krüger, C., & Mace, C.J. (2002). Psychometric validation of the state scale of

    dissociation (SSD). *Psychology and Psychotherapy: Theory, Research and*

    Practice, 75, 33-51. doi: 10.1348/147608302169535

Leiman, M. (1984). The concept of sign in the work of Vygotsky, Winnicott and

    Bakhtin; further integration of object relations theory and activity theory.

    British Journal of Medical Psychology, 65, 209-221.

Llewelyn, S.P. (2003). Cognitive analytic therapy: Time and process.

    Psychodynamic Practice, 9, 501-520. doi:

    10.1080/13533330310001616759

Lueger, R. J., & Barkham, M. (2010). *Using benchmarks and benchmarking to*

    improve quality of practice and service. Chichester: John Wiley & Sons.

*Mace, C., Beeken, S., & Embleton, J. (2006). Beginning therapy: Clinical outcomes

    in brief treatments by psychiatric trainees. *Psychiatric Bulletin,* 30, 7-10.

    doi: 10.1192/pb.30.1.7

Mann, J. & Goldman, R. (1982). *A casebook in time limited psychotherapy*. New York:

    McGraw-Hill.

Margison, F. (2000). Cognitive analytic therapy: a case study in treatment

    development. *British Journal of Medical Psychology,* 73, 145-149. doi:

    10.1348/000711200160354

*Marriott, M., & Kellett, S. (2009). Evaluating a cognitive analytic therapy service;

    practice-based outcomes and comparisons with person-centred and

cognitive-behavioural therapies. *Psychology and Psychotherapy: Theory, Research and Practice,* 82, 57-72. doi: 10.1348/147608308X336100

Morgan, H. G., & Russell, G. F. M. (1975). Value of family background and clinical features as predictors of long-term outcome in anorexia-nervosa: Four-year follow-up study of 41 patients. *Psychological Medicine,* 5, 355-371. doi: 10.1017/S0033291700056981

Morris, J. (2001). Psychological therapies in anorexia nervosa. *The British Journal of Psychiatry,* 179, 77-78. doi: 10.1192/bjp.179.1.77

Mulder, R., & Chanen, A. (2013). Effectiveness of cognitive analytic therapy for personality disorders. *British Journal of Psychiatry, 202,* 89-90. doi: 10.1192/bjp.bp.112.113571

NICE clinical guideline 9 (2004). *Eating disorders: Core interventions in the treatment and management of anorexia nervosa, bulimia nervosa and related eating disorder.* London: National Institute for Health and Clinical Excellence.

NICE clinical guideline 78 (2009a). *Borderline personality disorder: treatment and management.* London: National Institute for Health and Clinical Excellence.

NICE clinical guideline 90 (2009b). *Depression: the treatment and management of depression in adults (update)*. London: National Institute for Health and Clinical Excellence.

Ogden, T. (1983). The concept of internal object relations. *International Journal of Psychanalysis,* 64, 227-241.

Pines, A. (1992). *Romantic jealousy: Understanding and conquering the shadow of love*. New York: St. Martin's Press.

Pollock, P. H., Broadbent, M., Clarke, S., Dorrian, A., & Ryle, A. (2001). The personality structure questionnaire (PSQ): A measure of the mutliple self-

states model of identity confusion in cognitive analytic therapy. *Clinical Psychology and Psychotherapy,* 8, 59-72. doi: 10.1002/cpp.250

Rosenberg, M. (1989). *Society and the adolescent self-image*. Revised edition. Middletown, CT: Wesleyan University Press.

Ryle, A. (1985). Cognitive theory, object relations and the self. *British Journal of Medical Psychology,* 58, 1-7. doi: 10.1111/j.2044-8341.1985.tb02608.x

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

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

*Ryle, A., & Beard, H. (1993). The integrative effect of reformulation: Cognitive analytic therapy with a patient with borderline personality disorder. *British Journal of Medical Psychology,* 66, 249-258. doi: 10.1111/j.2044-8341.1993.tb01748.x

*Ryle, A., & Golynkina, K. (2000). Effectiveness of time-limited cognitive analytic therapy of borderline personality disorder: Factors associated with outcome. *British Journal of Medical Psychology,* 73, 197-210. doi: 10.1348/000711200160426

Ryle, A. (2001). CAT's dialogic perspective on the self. *Reformulation*, ACAT News Autumn.

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

Ryle, A., Kellett, S., Hepple, J., & Calvert, R. (in press). Cognitive analytic therapy at thirty. *Advances in Psychiatric Treatment.*

Salkovskis, P. M. (1995). Demonstrating specific effects in cognitive and

behavioural therapy. In M. Aveline & D. A. Shapiro (Eds.), *Research*

foundations for psychotherapy practice (pp. 191-228). Chichester: Wiley.

Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the

quality of marriage and other dyads. *Journal of Marriage and the Family,* 38,

15-28. doi: 10.2307/350547

Spielberger, C. D. (1999). *Manual for the state-trait anger expression inventory-2*.

Odessa, FL: Psychological Assessment Resources.

*Treasure, J., Todd, G., Brolly, M., Tiller, J., Nehmed, A., & Denman, F. (1995). A pilot

study of randomised trial of cognitive analytical therapy vs educational

behavioral therapy for adult anorexia nervosa. *Behaviour Research and*

Therapy, 33, 363-367. doi: 10.1016/0005-7967(94)00070-Z

Vanderlinden, J., Dyck, R. van., Vandereycken, W., Vertommen, H. & Verkes, R. jan.,

(1993) The dissociation questionnaire (DIS-Q): Development and

characteristics of a new self-report questionnaire. *Clinical Psychology &*

Psychotherapy, 1, 21-27. doi: 10.1002/cpp.5640010105

*Wildgoose, A., Clarke, S., & Waller, G. (2001). Treating personality fragmentation

and dissociation in borderline personality disorder: A pilot study of the

impact of cognitive analytic therapy. *British Journal of Medical Psychology,*

74, 47-55. doi: 10.1348/000711201160795

Williams, B.A. (2010). Perils of evidence-based medicine. *Perspectives on Biology*

and Medicine, 53, 106-120.

*Yeates, G., Hamill, M., Sutton, L., Psaila, K., Gracey, F., Mohamed, S., & O'Dell, J.

(2008). Dysexecutive problems and interpersonal relating following frontal

brain injury: Reformulation and compensation in cognitive analytic therapy

(CAT). *Neuro-Psychoanalysis,* 10, 43-58.

Young, J. E. (1998). The Young Schema Questionnaire: Short Form [On-line].

Available: www.home.sprynet.com/sprynet/schema/ysqs1.htm.

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale.

Acta Psychiatrica Scandinavica, 67, 361-370. doi: 10.1111/j.1600-

0447.1983.tb09716.x

Table 1; *CAT outcome studies grouped by methodological approach consistent with stages of the 'hourglass' model (Salkovskis, 1995)*

Author(s)	Number of sessions (Completers sample size)	Sample	Comparison Condition (Completers sample size)	Design	Standardised outcome measures	CASP Score	Downs and Black Score[a]
STAGE 1							
Case Description							
Ryle & Beard (1993)	26-sessions (N=1)	Borderline Personality Disorder Adult		Uncontrolled (pre-post)	IIP-127[1] SCL-90R[2] DES[3]	8[†]	10[†]
Bennett (1994)	16-sessions (N=1)	Depression, Anxiety Adult		Uncontrolled (pre-post)	BDI-II[4] SCL-90R IIP-127	6[†]	6[†]
Graham & Thavasothy (1995)	5-sessions (N=1)	Dissociative psychosis Adult		Uncontrolled (pre-post)	DES	6[†] 5[‡]	6[†] 4[‡]

---

[a] Total Downs and Black (1998) quality rating score

[†] First author's rating score; [‡] Second author's rating score
[1] Inventory of interpersonal problems-127 (Horowitz, Rosenberg, Baer, Ureno, & Villesenor, 1988)
[2] Symptom Checklist 90 Revised (Derogatis, Richels, & Rock, 1976)
[3] Dissociative Experiences Scale (Berstein & Putnam, 1986)
[4] Beck Depression Inventory-II (Beck, Steer, & Brown, 1995)

Yeates et al. (2008)	11-sessions (N=1)	Acquired Brain Injury Adult	Uncontrolled (pre-post)	BAI[5] HADS[6] STAXI-2[7] DAS[8]	7[†]	12[†]
Hamill & Mahoney (2011)	16-sessions (N=1)	Depression, Anxiety Older Adult	Uncontrolled (pre-post)	HADS GHQ-12[9]	5[‡]	11[†]
Repeated Measures Design						
Kellett (2005)	24-sessions (N=1)	Dissociative Identity Disorder Adult	Single Case Experimental Design	BSI[10] BDI-II IIP-32[11] DES[12] SSD[13] PSQ[14]	13[‡]	13[†]

[5] Beck Anxiety Inventory (Beck, Epstein, Brown, & Steer, 1988)
[6] Hospital Anxiety and Depression Scale (Zigmond & Snaith, 1983)
[7] State-Trait Anger Expression Inventory (Spielberger, 1999)
[8] Dyadic Adjustment Scale (Spanier, 1976)
[9] General Health Questionnaire-12 (Goldberg & Williams, 1988)

[10] Brief Symptom Inventory (Derogatis, 1993)
[11] Inventory of Interpersonal Problems-32 (Barkham, Hardy, & Startup, 1996)
[12] Dissociative Experiences Scale (Carlson & Putnam, 1993)
[13] State Scale of Dissociation (Krüger & Mace, 2002)
[14] Personality Structure Questionnaire (Pollock, Broadbent, Clarke, Dorrian, & Ryle, 2001)

Kellett (2007)	24-sessions (N=1)	Histrionic Personality Disorder Adult		Single Case Experimental Design	BSI BDI IIP-32 PSQ YSQ-SV[15]	13‡	15†
Kellett & Totterdell (2013)	16-sessions (N=1)	Morbid Jealousy Adult	16 sessions CBT (N=1)	Single Case Experimental Design	BSI BDI-II IIP-32 RJQ[16] PJQ[17]	15‡	13†
Kellett, Bennett, Ryle & Thake (2013)	24-sessions (N=17)	Borderline Personality Disorder Adult		Repeated measures	PSQ DES CORE[18] Borderline Index of Severity	15‡	20†
Kellett & Hardy (in press)	24-sessions (N=1)	Paranoid Personality Disorder		Single Case Experimental Design	BSI BDI-II IIP-32 PSQ	13‡	17†

---

[15] Young Schema Questionnaire-Short Version (Young, 1998)

[16] Romantic Jealousy Questionnaire (Pines, 1992)
[17] Prestwich Jealousy Scale (Beckett, Tarrier, Intili & Beech, 1992)

[18] Clinical Outcomes in Routine Evaluation (Evans, Connell, Barkham, Margison, McGrath, Mellor-Clark, & Audin, 2002)

STAGE 2 **Randomised Controlled Trials (RCT)**							
Treasure et al. (1995)	20-sessions (N=10);	Anorexia Nervosa or Bulimia Nervosa Adult	20 sessions Educational Behavioural Treatment (N=10)	RCT	MRS[19]	15‡	20†
Fosbury, Bosley, Ryle, Sonksen & Judd (1997)	16-sessions (n=10);	Diabetes Adult	Diabetes Specialist Nurse Education (DSNE) 14-18 sessions (N=16)	RCT	IIP-127	15†	23†
Dare, Eisler, Russell, Treasure & Dodge (2001)	13-sessions (n=22)	Anorexia Nervosa or Anorexia Nervosa/ Bulimia Nervosa Adult	25-sessions Focal Psychodynamic Psychotherapy (N=21) 14-sessions Family Therapy (N=22) 11-sessions 'Routine' treatment as usual (N=19)	RCT	MRS	14† 15‡	20† 21‡

---

[19] Morgan and Russell Scale (Morgan & Russell, 1975)

| Clarke, Thomas & James (2013) | 24-sessions (n=38) | Personality Disorder

Adult | Treatment as Usual (N = 40) | RCT | SCID-II[20]
IIP-32
CORE
DisQ[21]
DES
SCL-90R | 14‡ | 26† |

STAGE 3

Quasi-Experimental studies

| Chanen et al. (2008*; 2009) | 13-sessions (N=41) | Borderline Personality Disorder

Adolescents (aged 15-18) | 11-sessions, Good Clinical Care (N=37)

15-sessions Historical Treatment as Usual (N=32) | Quasi-experimental design or RCT* | YSR[22]/YASR
SOFAS[23] | 15‡ | 24† |

Practice-based studies of effectiveness

| Brockman, Poynton, Ryle & | 12 sessions (N=30) | Depression, Anxiety | 12-sessions Interpretive Therapy (N=18) | Uncontrolled | BDI[24]
GHQ-60[25] | 9† | 17† |

---

[20] Structured Clinical Interview for DSM-IV Axis II (First, Gibbon, Spitzer & Williams, 1996)
[21] Dissociative Questionnaire (Vanderlinden, Van Dyck, Vandereycken, Vertommen & Verkes, 1993)

[22] Young Adult Self-Report (Achenbach, 1997)
[23] Social and Occupational Functioning Assessment Scale (Goldman, Skodol, & Lave, 1992)

[24] Beck Depression Inventory (Beck, Ward, Mendelson, Mock, & Erdaugh, 1961)
[25] General Health Questionnaire-60 (Goldberg, 1972)

Watson (1987)		Adult			CCI[26]		
Duignan & Mitzman (1994)	4 individual sessions and 12-sessions Group CAT (n=7)	Personality Disorders, Depression, Panic Disorder, Psychosis Adult	12-sessions individual CAT (n=30; Brockman et al., 1987)	Uncontrolled	BDI GHQ-60 CCI	8[†]	17[†]
Clarke & Llewelyn (1994)	16-sessions (N=6)	Female Survivors of Childhood Sexual Abuse Adult		Uncontrolled, pre/post	SCL90-R JBI[27] BDI-II RSES[28]	10[†]	15[†]
Dunn, Golynkina, Ryle & Watson (1997)	16-sessions (N=86)	Depression, Anxiety Adult		Uncontrolled, pre/post	SC[29] BDI SCL90-R IIP	10[†]	15[†]
Clarke & Pearson (2000)	16-sessions (N=4)	Male Survivors of Childhood Sexual Abuse (BPD & NPD[30]) Adult		Uncontrolled, pre/post	SCL90-R JBI BDI-II RSES	9[†]	15[†]

---

[26] Crown Crisp Inventory (Crown & Crisp, 1979)
[27] Jehu Belief Inventory (Jehu, 1988)
[28] Rosenberg Self Esteem Scale (Rosenberg, 1989)
[29] Social Questionnaire (Corney, Clare & Fry, 1982)
[30] Narcissistic Personality Disorder

Study	Intervention	Diagnosis	Comparison	Design	Measures		
Ryle & Golynkina (2000)	24-sessions (N=27)	Borderline Personality Disorder Adult		Uncontrolled, pre-post	BDI SCL90-R IIP-127 SQ Borderline Index of Severity	14‡	19†
Wildgoose, Clarke & Waller (2001)	16-sessions (N=5)	Borderline Personality Disorder Adult		Uncontrolled, pre-post	PSQ DIS-Q[31] SCL-90-R IIP-127	13‡	18†
Birtchnell, Denman & Okhai (2004)	16-sessions (N=32)	Not reported. (classified as common mental health difficulties) Adult		Uncontrolled, pre-post	PROQ2[32] CORE[33]	11‡	11†
Mace, Beeken & Embleton (2006)	16-sessions (N=17)	Personality Disorders, Affective Disorders Adult	Brief Psychodynamic Therapy (n=17)	Uncontrolled, pre-post	GHQ-12	14† 12‡	13† 11‡
Marriott & Kellett (2009)	Short-term CAT 16-sessions (n=38) Medium-term CAT 24-sessions (n=27)	Depression, Anxiety, Phobias, Obsessive Compulsive Disorder, Post Traumatic Stress Disorder, Personality Disorders Adult	Short-term CBT 16-sessions (n=38) Medium-term CBT 20-sessions (n=27)	Uncontrolled, pre-post	BSI BDI-II IIP-32	16‡	17†

---

[31] Dissociation Questionnaire (Vanderlinden, Dyck, Vandereycken, Vertommen, & Verkes, 1993)

[32] Person's Relating to Others Questionnaire (Birtchnell & Evans, 2004)

[33] Clinical Outcomes in Routine Evaluation (Evans et al., 2002)

Short-term PCT
9-sessions
(n=38)

Medium-term PCT
17-sessions (n=25)