

The Content, Quality and Impact of Cognitive Behavioural Case Formulation During Treatment of Obsessive Compulsive Disorder

Anna Nattrass

University of Sheffield, UK

Stephen Kellett

University of Sheffield, and Sheffield Health and Social Care NHS Foundation Trust, UK

Gillian E. Hardy

University of Sheffield, UK

Tom Ricketts

Sheffield Health and Social Care NHS Foundation Trust, UK

Background: Case formulation (CF) is considered a core component of cognitive behavioural therapy (CBT). Despite this, few studies have explored the content and quality of CFs in routine practice and any associated impact on treatment outcome. **Aims:** This study investigated (1) the content, timing and quality of CF for patients ($N = 29$) with Obsessive Compulsive Disorder; (2) the impact of CF on symptoms, distress and therapeutic alliance; and (3) the relationship between CF quality and treatment outcome. **Method:** CF content and quality were assessed from audiotapes of routine CBT sessions using a validated coding manual and evaluated against treatment outcomes at different stages of therapy. **Results:** CFs were developed early during treatment and contained strong behavioural and cognitive components, with a focus on symptoms and maintaining factors. CF quality ratings ranged from rudimentary to excellent. A significant improvement in distress and the therapeutic alliance occurred following the introduction of CF, but no significant relationship was found between CF quality and outcome at any stage of treatment. **Conclusion:** CF may be valuable in reducing patient attrition, due to its impact early in treatment. Further research is needed to explore the most important components of CF. Theoretically sound and disorder specific measurement tools for evaluating CFs are required.

Keywords: Case formulation, CBT, obsessive compulsive disorder.

Reprint requests to Anna Nattrass / Stephen Kellett, Clinical Psychology Unit, Department of Psychology, The University of Sheffield, S10 2TN, Sheffield. E-mail: annanattrass@yahoo.com/s.kellett@sheffield.ac.uk

© British Association for Behavioural and Cognitive Psychotherapies 2014

Introduction

Obsessive Compulsive Disorder (OCD) is a common and chronic mental health condition (APA, DSM-V, 2013). Sufferers report significant associated disability due to loss of income and decreased quality of life (Bobes et al., 2001). Historically, the most common psychological treatment for OCD has been behaviour therapy in the form of exposure with response prevention (ERP), which is effective for a large proportion of patients (e.g. Foa et al., 2005). However, up to 30% of patients report no benefit from ERP, with a 15–20% attrition rate during the early stages of treatment (de Haan et al., 1997). Attempts to improve treatment outcomes have focused on the addition of cognitive techniques (Zivor, Salkovskis and Oldfield, 2011). Evidence for the added benefit of a cognitive element to treatment is mixed (McManus, Grey and Shafran, 2008), but nevertheless CBT is now the recommended treatment for patients with mild to moderate OCD (NICE, 2005). Maximizing the effectiveness of CBT treatment for OCD therefore represents a major and on-going challenge.

Case formulation (CF) is a core component of CBT (Eells, 2007). It aims to synthesize patient experience with theory and research in a patient-centred manner (Kuyken, Padesky and Dudley, 2009). CF should be collaboratively co-created by the patient and therapist, evolve over the course of therapy, and emphasize patient strengths (Kuyken et al., 2009). It is claimed that a good CF enhances the process and outcome of therapy by guiding the selection and sequence of appropriate intervention, enriching the therapeutic alliance and relieving patient distress (Beiling and Kuyken, 2003; Kuyken, Padesky and Dudley, 2008). Whilst strongly advocated in CBT theory, few studies have investigated the claimed benefits of CF and little is known about how CF is used in routine practice (Beiling and Kuyken, 2003; Eells, 2007; Kuyken, 2006).

Zivor, Salkovskis and Kushnir (2013) investigated the quality of OCD CFs developed by CBT clinicians and found the average CF to be “poor”, increasing to “adequate” following a CF training workshop. Kuyken, Fothergill, Musa and Chadwick (2003) rated clinicians’ CBT CFs for depressed patients and found fewer than half (44%) to be “good enough”. CF quality has also been found to vary according to depth of clinical expertise (Eells, Lombart, Kendjelic, Turner and Lucas, 2005b; Kuyken et al., 2003; Zivor, Salkovskis and Oldfield, 2013). A limitation of these studies however is that they evaluated CFs generated from case vignettes, which fails to replicate the process of developing a CF with a patient in real time (Beiling and Kuyken, 2003).

Research investigating the effectiveness of CF has largely relied upon randomized control trials (RCTs) comparing outcomes of CBT with and without patient-centred CF; and has found CBT with patient-centred CF is no different or only marginally better (Emmelkamp, Bouman and Blaauw, 1994; Schulte, Kunzel, Pepping and Schulte-Bahrenberg, 1992). The reliability of RCTs in detecting difference has however been questioned due to the many factors that blur the distinction between standardized and individualized CBT (Wilson, 1997) and the failure to systematically evaluate CF quality (Schulte et al., 1992). Gower and Kuyken (2011) found therapists’ competence in CF was associated with better treatment outcomes among depressed patients, suggesting that CF may improve treatment outcome, but CF quality needs to be considered.

Chadwick, Williams and Mackenzie (2003) used a mixed methodology to investigate the immediate impact of a two-session CF phase on symptoms and therapeutic alliance, for

patients with psychosis. Patients reported feeling reassured and optimistic following the CF, but also worried due to perceiving problems to be complex and longstanding. Therapists (but not patients) reported a positive impact on the alliance, but no symptomatic gains were reported. Findings cannot, however, be generalized to OCD patients, due to the nature of the sample. Limited research to date suggests CF may have an impact on therapy process, but the impact on treatment outcome is unclear.

In summary, it is claimed that CF enhances the process and outcome of CBT. However, few studies have explored the quality and clinical impact of CF in routine practice and there is little disorder-specific evidence. Evidence concerning the manner in which CFs are developed in routine practice and how this relates to treatment process and outcome in a disorder-specific sample is therefore at a premium. This is significant for OCD patients, where there is high attrition rate and non response to standard CBT treatment. The aims of the current study were: (1) to define the timing, content and quality of CF in routine practice; (2) to use a single diagnosis group of OCD; and (3) to assess the immediate and overall impact of CF.

Method

Design and sample

The study used archived data of CBT for OCD offered in a psychotherapy service within the NHS, UK ($N = 37$; Houghton, Saxon, Bradburn, Ricketts and Hardy, 2010). The current study received ethical approval from a local research ethics committee. Data comprised session-by-session outcome measurement and audiotapes of each therapy session. Twenty-nine patients (13 males, 16 females) were included in the current study (78% of the original 37). All patients met the International Classification of Diseases, 10th edition criteria for OCD (ICD-10, F42; World Health Organization, 1992).

All therapists ($N = 8$) met British Association for Behavioural and Cognitive Psychotherapy (BABCP) accreditation criteria for registration with the UK Council for Psychotherapy and had between 2–10 years of post-qualification CBT experience. Each therapist treated between 2–6 clients. Prior to commencement of the original study, therapists received two refresher workshops on CBT for OCD. During the study, therapists attended a 1-hour supervision group on a fortnightly basis.

Measures

Yale-Brown Obsessive Compulsive Scale (YBOCS; Goodman et al., 1989). The self-completed version of the Y-BOCS consists of 10 items corresponding to the personally relevant OCD symptoms identified in the clinician administered YBOCS interview (Steketee, Frost and Bogart, 1996). This measure was completed prior to each CBT session.

The Clinical Outcomes in Routine Evaluation-Short Form (CORE-SF; Evans et al., 2002). This is an abbreviated version of the full 34-item version (Barkham et al., 2001) and measures generic psychological distress. This measure was completed prior to each CBT session.

The Agnew-Davies Relationship Measure (ARM-12; Cahill et al., 2012). This is an abbreviated version of the ARM-28 (Cahill et al., 2012) and was used to measure the therapeutic alliance. The client version was used for the present study, completed following each CBT session.

Table 1. Inclusion and exclusion criteria for CF content and quality coding

Inclusion criteria

1. Explicit or implicit reference to formulation by therapist or client (e.g. “now I’d like to talk to you about your formulation/what’s keeping this problem going”) - include discussion that follows
2. Summary of symptoms/problems (includes obsessions - thoughts, images, impulses - feelings, emotions, physiological symptoms, difficulties in day-to-day functioning)
3. Summary of descriptive information (history of mental health problem, family psychiatric history, medical history, developmental history, relationships)
4. Links made between problem and predisposing factors (including childhood, adulthood, personality, skills deficits, biological factors); or maintenance factors (including triggers, thoughts, images, compulsions, safety behaviours, social/cultural factors, social factors, role conflict)
5. Links made between problem and theory, research, diagnosis or case examples
6. Links made between problem and client’s strengths/positive treatment indicators (including adaptive views of others, adaptive wishes, hopes, goals, psychosocial support, motivation for treatment)
7. Reference to therapy interfering events (items identified as possible obstacles to successful treatment outcome)

Exclusion criteria

1. Discussion of session practicalities (e.g. session planning)
 2. Review of the homework
 3. Assessment (refers to therapist enquiring about the client’s difficulties and client responding with descriptions or examples - includes process of therapist checking or clarifying what had been shared)
 4. Discussion about treatment
-

Intervention

Therapists provided CBT for OCD and treatment duration was not predetermined. Treatment was not manualized, but supervision and training ensured that all treatments included cognitive and behavioural approaches. Session content was recorded (by the therapist) at the end of each session on a Therapy Process Record (TPR).

Procedure

Step 1: Identifying CF sessions. The TPR identified CF sessions. Some therapists ($n = 9$) reported the development and presentation of the CF early and again later in treatment (indicating re-formulation). Only initial CF sessions were included.

Step 2: Identifying formulation narrative. For over half the cases ($n = 17$) development and presentation of CF were reported as occurring within the same sessions. When this was not the case, development of CF always preceded presentation of CF. Sessions during which CF was presented were selected for analysis ($n = 70$).

Audiotapes of the selected CF sessions were listened to by the first author and passages relating to CF were identified and transcribed. Exclusion and inclusion criteria were developed using the Eells, Lombart, Kendjelic, Turner and Lucas (2005a) guidelines in their Case Formulation Content Coding Manual (CFCCM v2). Table 1 lists exclusion and inclusion criteria for the study. For cases where the start of the CF presentation was not clear ($n = 12$), the audiotape from the previous session was also listened to and passages meeting

the inclusion criteria included. An additional researcher independently analyzed three cases, using the same criteria. There was full agreement on the CF passages corresponding to these patients.

Step 3: Preparation of the transcripts. Transcripts were segmented into “idea units” (IUs) using the Stinson, Milbrath, Reidbord and Bucci (1994) method, which divides text into small portions (usually a sentence or less) to represent an IU. Stinson et al. (1994) reported mean reliability estimates of 0.87 using this method. The number of IUs per CF ranged from 27 to 207 (mean = 108).¹

Step 4: Identifying formulation sessions. A maximum of two sessions per patient were selected for content and quality coding, as two was the modal number of CF sessions across the sample. For those patients where the therapist reported CF across more than two sessions ($n = 13$), sessions were ranked according to the number of IUs pertaining to CF. The two highest-ranked sessions were included in the analysis.

Step 5: Content and quality analysis of formulations. CF content and quality was assessed via the CFCCM v2 (Eells et al., 2005a), which has excellent interrater reliability (Eells et al., 2005b). The CFCCM v2 categorizes CF content according to three main categories: (1) descriptive; (2) diagnostic; and (3) inferential - with each category containing several specific CF elements. For the purpose of this disorder-specific research, coding categories remained static, with additional OCD descriptors added. Specifically, the dysfunctional thoughts category was elaborated to include unhelpful beliefs about symptoms (e.g. negative appraisals); the safety behaviours category was elaborated to include behaviours to prevent symptoms from occurring (e.g. avoidance) and the affect regulation category was elaborated to include behaviours adopted in response to OCD symptoms (e.g. checking). Therapy interfering events referred to possible obstacles to treatment; and positive treatment indicators referred to patient’s strengths.

The CFCCM quality coding system has six indices and includes: (1) formulation elaboration; (2) precision of language; (3) complexity; (4) coherence; (5) systematic process; and (6) comprehensiveness. The first five criteria are measured on a 5-point scale (0 = not present; 1 = rudimentary; 2 = adequate; 3 = good; and 4 = excellent) whilst the sixth criterion (comprehensiveness) was measured on a 9-point scale. For the purpose of analysis, this was transformed into a 0–4 scale. A global quality rating was achieved by summing each individual quality rating, resulting in a score between 0 (no quality) to 24 (maximum quality). This was transformed into a scale of 0 – 4 and assigned the same category labels as the independent criteria.² Evidence of written CFs was available for only four cases, so not included in the study.

Step 6: Recruiting and training raters. Three independent raters were used, each with prior experience of treating OCD with CBT and were blind to treatment outcome. Raters attended a 3-hour training session on OCD-CBT CF and treatment and received 2 days of CFCCM training. Raters were required to achieve a level of reliability of $>.75$ against the content codes on the example CFs (developed by the research team). Raters’ understanding of the quality codes was assessed via discussion about the quality of the example CFs. The Kappa statistic was used to establish interrater agreement for content coding (Siegel and Castellan, 1988) and the intraclass correlation coefficient (ICC) was used for interrater agreement for quality

¹ Example transcripts are available on request from the first author.

² The adapted manual is available on request from the first author.

coding (ICC; Shrout and Fleiss, 1979). The goal of the content coding step was to achieve a set of reliable consensus codes for each IU in a CF. All three raters coded the same CF session independently, and then immediately held a consensus meeting. Consensus was defined as at least two raters assigning the same code. Interrater reliability for content and quality across every CF obtained a multi-rater kappa coefficient 0.64, suggesting good agreement across content coding categories (defined as .60 – .80; Cohen, 1968). The ICC yielded an overall reliability estimate of 0.75 across quality ratings. Independent quality ratings ranged from 0.92 (complexity), 0.88 (systematic process), 0.78 (elaboration), 0.76 (precision of language) and 0.73 (coherence), indicating excellent interrater reliability across quality coding categories (defined as $>.74$; Fleiss, 1981).

Data analysis

Pre–post YBOCS and CORE-SF scores were analyzed to identify individual rates of both reliable and clinically significant change (Jacobson and Truax, 1991) and associated group effect sizes (Cohen, 1988). To consider clinical impact of CF, contact with patients (and associated sessional outcomes) were divided into three phases: (1) an “assessment phase” defined as all sessions prior to the first selected CF session; (2) a “formulation phase” defined as the two selected CF sessions (where this was more than two, the two sessions with the highest number of IUs pertaining to CF were selected); (3) a “post-formulation phase” defined as the two sessions immediately following the two CF sessions. Whilst it is acknowledged that CF may have continued to be discussed and reviewed, phases were designed to capture any initial impact of CF introduction. CF impact was tested by comparing phase mean outcome scores and post hoc pairwise comparisons identified any significant phase differences. Relationship between CF quality and overall outcome was tested via correlational analysis at three different time points – the end of the formulation phase, the end of the post-formulation phase, and at the end of treatment.

Results

Pre–post outcomes showed a large effect size, $d = .89$ (Cohen, 1988) on the Y-BOCS, with 31% of the sample achieving both a reliable and clinically significant reduction in OCD symptoms. Sample mean score was in the moderate-severe range at assessment ($M = 23.9$; $SD = 7.54$), falling to within the mild-moderate at termination ($M = 17.14$, $SD = 8.74$). A medium to large effect size, $d = .72$ (Cohen, 1988) was found on the CORE-SF, with 20% experiencing both a reliable and clinically significant change. ANOVA found no effect of therapist on treatment outcome on YBOCS $F(7, 21) = .47$, $p = .85$; or on CORE-SF $F(7, 21) = .69$, $p = .68$.

Every therapy contained a CF, with treatment duration ranging from 6 to 31 sessions ($M = 13.52$, $SD = 6.02$). CF development began between sessions 1 and 5 ($M = 2.48$, $SD = 1.30$) and took between 1 and 5 sessions ($M = 2.34$, $SD = 1.11$). CF presentation began between sessions 2 and 5 ($M = 3.07$, $SD = 1.25$) and typically lasted between 1 and 3 sessions ($M = 2.41$, $SD = .71$). Table 2 summarizes the content included in the CFs. Symptoms (mean 23 idea units per CF) and maintaining factors (mean 40 idea units per CF) were omnipresent. Affect regulation/compulsions (97 % of CFs), dysfunctional thoughts (93%) and triggers (79%) were the most frequently occurring maintaining factors. Predisposing factors were

Table 2. Content of OCD case formulation in routine practice

Formulation content code	Rated as present within the formulation			
	Total (<i>n</i> = 29)	%	No. idea units coded per formulation (Range)	Mean
1. Predisposing factors (total) and predisposing experiences (total)	19	65	0–34	8
Predisposing (childhood)	15	48	0–15	2
Predisposing (adulthood)	19	65	0–22	4
Problematic traits of self	11	38	0–9	1
Problematic aspects of relating	1	4	0–1	0
Skills deficit	0	0	0	0
Biological factors	3	10	0–5	0
2. Symptoms/problems (total)	29	100	1–53	23
Obsessions	27	93	0–23	8
Problems in global Functioning	14	48	0–6	1
Feelings/emotion	25	82	0–19	9
Physiological symptoms	19	65	0–35	5
3. Maintaining factors (total)	29	100	7–72	40
Triggers	21	79	0–26	7
Dysfunctional thoughts/beliefs	27	93	0–29	12
Affect regulation (e.g. checking/ neutralizing behaviour) ^a	28	97	0–45	19
Safety behaviours (e.g. avoidance) ^a	17	59	0–20	3
Social/cultural factors (total)	3	10	0–1	0
Poor psychosocial support	1	3	0–1	0
Demographics	2	7	0–1	0
Role conflict	0	0	0	0
4. Diagnosis (total)	15	52	0–6	2
5. Descriptive information (total)	16	55	0–13	2
Identifying information	3	11	0–2	0
History of mental health problem	15	52	0–10	2
Family psychiatric history	2	7	0–1	0
Medical history (adulthood)	2	7	0–1	0
Medical history in family	0	0	0	0
Developmental history	7	21	0–3	1
Adult life history (relationships)	3	10	0–1	0
Mental status/appearance	0	0	0	0
Other descriptive data	0	0	0	0
6. Positive treatment indicators (total)	5	17	0–2	0
Strengths/skills	5	17	0–2	0
Adaptive views of others	0	0	0	0
Positive motivation for treatment	0	0	0	0
Adaptive wishes, hopes, goals	1	3	0–1	0
Good psychosocial support	2	7	0–1	0
7. Therapy interfering events	3	10	0–1	0

^a behaviour also commonly labelled “compulsions”

^b total no. of idea units per formulation ranged from 27–207 (mean = 108)

Table 3. Mean scores on outcome measures during therapy phases

	Assessment phase mean (<i>SD</i>)	Formulation phase mean (<i>SD</i>)	Post-formulation phase mean (<i>SD</i>)	<i>F</i> value
YBOCS	23.9 (7.54)	22.83 (7.54)	20.93 (8.19)	3.83*
CORE-SF	18.20 (7.2)**	17.20 (9.5)	14.80 (8.6)**	9.56*
ARM-12	70.28 (8.43)**	73.31 (9.32)	75.41 (8.07)**	6.62*

* $p < .05$; **Mean scores significantly differ ($p < .05$) on pairwise comparisons with Bonferroni correction applied

Table 4. Partial correlations between formulation quality and outcome during study phases

Phase and outcome measure	Correlation with formulation quality*
YBOCS mean score - formulation phase	.36
YBOCS mean score - post formulation phase	.37
YBOCS termination score	.12
CORE-SF mean score – formulation phase	–.24
CORE-SF mean score –post formulation phase	.05
CORE-SF termination score	.09
ARM-12 mean score - formulation phase	.08
ARM-12 mean score - post formulation phase	.05
ARM-12 termination score	.22

*None of the correlations were significant at a level of $p < .05$.

present within 65% of CFs, typically describing negative life events. Descriptive information was evident within 55% of CFs, with history of mental health problems included in 52% (although with a mean of only 2 idea units per CF). Positive treatment indicators (17%) and potential therapy interfering events (10%) were the least frequently occurring content in the CFs.

The distribution of global quality ratings ranged from 9 to 23 with a mean of 14.83. Formulation quality scores were distributed as follows: rudimentary ($n = 4$; 14%), adequate ($n = 11$; 38%), good ($n = 9$; 31%) and excellent ($n = 5$; 17%). Elaboration achieved the highest mean quality rating ($M = 2.72$, $SD = .80$), followed by complexity ($M = 2.53$, $SD = 1.02$), coherence ($M = 2.41$, $SD = .87$), precision of language ($M = 2.31$, $SD = .76$), systematic process ($M = 2.21$, $SD = 1.05$) and comprehensiveness ($M = 1.90$, $SD = 3.55$).

In terms of CF impact, Table 3 presents the outcome mean scores according to study phase. Whilst there was a significant decrease in OCD symptoms (YBOCS; $F(2, 56) = 3.83$, $p < .05$), post hoc comparisons indicated no significant difference between phase pairs following Bonferroni correction (Abdi, 2007). There was a significant improvement in psychological distress (CORE-SF; $F(2, 56) = 9.56$, $p = .001$) and an increase in the alliance (ARM-12; $F(2, 56) = 6.62$, $p < .001$). Psychological distress during the post-formulation phase ($M = 14.81$, 95% CI [11.55, 18.08]) was significantly lower ($p < .01$) than during assessment ($M = 18.23$, 95% CI [15.48, 20.99]) and formulation phases ($M = 17.20$, 95% CI [13.61, 20.80]).

The alliance at the end of the post formulation phase ($M = 75.41$, 95% CI [72.34, 78.49]) was significantly higher ($p < .05$) than the alliance during the assessment phase ($M = 70.28$, 95% CI [67.07, 73.48]). The relationship between CF quality and outcome was tested across the formulation phase, post-formulation phase and end of treatment, with partial correlations controlling for intake scores. Table 4 illustrates that CF quality did not correlate with outcome at any of the study phases.

Discussion

This is the first attempt to assess the timing, content and quality of CF in an OCD specific sample during routine CBT practice and to simultaneously assess the relationship with outcome. Findings indicate that therapists formulate early in the course of OCD treatment, always include symptoms and maintaining factors, and that CF reduces patient distress and improves the alliance. The most frequently noted aspect of OCD CFs were affect regulation (which includes compulsions) followed by dysfunctional thoughts/beliefs. This suggests behavioural and cognitive balance to OCD CFs in routine practice. Over half of CFs (65%) included predisposing factors, indicating that therapists often paid attention to historical factors. Positive treatment indicators such as patient strengths were rarely included, despite the prominence given to strength/resilience factors in CF theory (Kuyken et al., 2009).

There was a range in overall CF quality, with less than half (48%) falling within the good to excellent range, despite specific training and on-going supervision. Findings are consistent with the notion that CF is one of the most complex skills in CBT (Kuyken et al., 2008), suggesting a need for specialist training (Zivov et al., 2013). However, simple descriptive CFs may in fact be appropriate during early stages of treatment (Kuyken et al., 2009), with therapists limiting content to prevent patient and therapist feeling overwhelmed (Chadwick et al., 2003). In the current study, the CFs received the lowest rating for comprehensiveness (the number of different CF categories included) and highest for elaboration (the extent to which one single element was explained), supporting the idea that OCD CFs are kept relatively simple early in treatment. A limitation of the current study was the failure to capture attempts at reformulation later in treatment, as CF quality may vary according to stage of treatment (Kuyken et al., 2009; Zivov et al., 2013).

In terms of the immediate impact of CF, results indicate a reduction to distress and an increase in the alliance following CF. Sharing CFs may usefully “contain” patient distress by helping make sense of a potentially confusing array of symptoms (Eells, 2007). It may take time for patients to effectively assimilate their CF (Stiles, 2002) due to the mixed feelings CF can provoke (Chadwick et al., 2003). The alliance during the post-formulation phase significantly improved, suggesting that CF may be an important aspect of alliance formation. These findings would support claims made in theory that CF enhances the process of therapy (Bieling and Kuyken, 2003; Kuyken et al., 2008). This is particularly important in OCD given the high rates of attrition early in treatment (de Haan et al., 1997). Findings are, however, tentative given there was no control group to measure whether change would have occurred without CF. The study also lacked patient feedback, which would have helped establish the extent to which patients attributed change to CF.

Findings suggest that quality of the CF was not related to treatment outcome at any stage of treatment. It is possible that common factors within the CF process, such as the experience of being attentively listened to, are more important. Gower and Kuyken (2011) however did find

a positive association between therapists' competence in CF and treatment outcome. A major distinction between these studies was the measure of CF quality employed. The CFCCM focuses on the comprehensiveness and structure of the CF, whereas the Collaborative Case Conceptualization Rating Scale (CCC-RS; Padesky, Kuyken and Dudley, 2011, as used in Gower and Kuyken, 2011) places greater emphasis on collaboration and extent to which CF is matched to patient need, at different stages of treatment. It is possible that factors associated with the content and structure of CF are less important than the extent to which a "good enough" CF is co-created in a collaborative manner. Furthermore, the CFCCM assumes quality indicators are static and of equal weight (i.e. it is an additive scale), but the current findings suggest this might not be the case. Further research is needed to identify aspects of CF that are important in routine practice and how this may alter over the course treatment. Such research could be used to inform the development of simple, theoretically sound, disorder-specific measures of CF quality for research, audit and training purposes. Existing research offers examples of such measures that have been developed for use in specific studies e.g. Zivov et al. (2013) developed a brief, OCD-specific CF quality measure for their research. Further research could establish the applicability of these measures more widely.

For the purpose of this research, treatment was divided into distinct phases – assessment, formulation and post-formulation. Given the nature of this research, CF sessions could not be considered to have been concerned with CF exclusively, so the "impact" of CF may have been due to other factors. This criticism is especially cogent for the post-formulation phase, during which outcomes may have been related to other processes, such as application of evidence-based interventions. Such limitations reflect the complexity of conducting CF research with high internal validity in routine practice and may highlight the reason why practice-based CF research is so limited. The current study investigated the impact of CF on patient distress, symptoms and alliance. Given CF is thought to impact on these factors indirectly (Bieling and Kuyken, 2003), further research could explore the direct impact of CF, including its impact on patients' understanding of their problems and the selection of interventions.

In conclusion, this research evidence suggests that CBT therapists are routinely using CF early in the treatment of OCD and this appears to have an immediate impact on reducing patient distress and on improving the therapeutic alliance. Such CF factors may therefore play an important role in reducing early attrition from OCD treatment. However, the mechanism by which these effects occur remains ill-defined and under researched. The relationship between CF quality and treatment outcome is less clear and this study has raised important questions about how CF quality is defined and measured and how this might vary depending on the presenting problem and stage of treatment. The prominent role that formulation plays in clinical practice has not yet been justified by the current evidence base.

References

- Abdi, H.** (2007). Bonferroni and Sidak corrections for multiple comparisons. In N.J. Salkind (Eds.), *Encyclopaedia of Measurement and Statistics* (pp. 103–107). Thousand Oaks, CA: Sage.
- American Psychiatric Association (APA)** (2013). *Diagnostic and Statistical Manual of Mental Disorders (5th ed.)*. Washington, DC: Author.

- Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., et al.** (2001). Service profiling and outcomes benchmarking using the CORE-OM: towards practice-based evidence in the psychological therapies. *Journal of Consulting and Clinical Psychology*, *69*, 184–196.
- Bieling, P. J. and Kuyken, W.** (2003). Is cognitive case formulation science or science fiction? *Clinical Psychology-Science and Practice*, *10*, 52–69.
- Bobes, J., Gonzalez, M.P., Bascaran, M.T., Arango, C., Siaz, PA. and Bousono, M.** (2001). Quality of life and disability in patients with obsessive-compulsive disorder. *European Psychiatry*, *16*, 239–245.
- Cahill, J., Stiles, W. B., Barkham, M., Hardy, G. E., Stone, G., Agnew-Davies, R., et al.** (2012). Two short forms of the Agnew Relationship Measure: the ARM-5 and ARM-12. *Psychotherapy Research*, *22*, 241–255.
- Chadwick, P., Williams, C. and Mackenzie, J.** (2003). Impact of case formulation in cognitive behaviour therapy for psychosis. *Behaviour Research and Therapy*, *41*, 671–680.
- Cohen, J.** (1968). Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, *70*, 213–220.
- Cohen, J.** (1988). *Statistical Power Analysis for the Behavioral Sciences (2nd ed.)*. Hillsdale, NJ: Erlbaum.
- de Haan, E., van Oppen, P., van Balkom, A. J. L., Spinhoven, P., Hoogduin, K. A. L. and Van Dyck, R.** (1997). Prediction of outcome and early vs late improvement in OCD patients treated with cognitive behaviour therapy and pharmacotherapy. *Acta Psychiatrica Scandinavica*, *96*, 354–361.
- Eells, T.** (2007). *Handbook of Psychotherapy Case Formulation (2nd ed.)*. New York: Guilford Press.
- Eells, T., Lombart, K. G., Kendjelic, E. M., Turner, C. and Lucas, C. P.** (2005a). *Manual for Case Formulation and Treatment Plan Coding*. Unpublished manuscript.
- Eells, T., Lombart, K. G., Kendjelic, E. M., Turner, C. and Lucas, C. P.** (2005b). The quality of psychotherapy case formulations: a comparison of expert, experienced, and novice cognitive behavioral and psychodynamic therapists. *Journal of Consulting and Clinical Psychology*, *73*, 579–589.
- Emmelkamp, P., Bouman, T. and Blauw, E.** (1994). Individualized versus standardized therapy: a comparative evaluation with obsessive-compulsive patients. *Clinical Psychology and Psychotherapy*, *1*, 95–100.
- Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J., et al.** (2002). Towards a standardised brief outcome measure: psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry*, *180*, 51–60.
- Fleiss, J. L.** (1981). *Statistical Methods for Rates and Proportions (2nd ed.)* (pp.38–46). New York: John Wiley.
- Foa, E. B., Liebowitz, M. R., Kozak, M. J., Davies, S. O., Campeas, R., Franklin, M. E., et al.** (2005). Treatment of obsessive compulsive disorder by exposure and ritual prevention, clomipramine, and their combination: a randomized, placebo-controlled trial. *American Journal of Psychiatry*, *162*, 151–161.
- Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischmann, R. L., Hill, C. L., et al.** (1989). The Yale-Brown Obsessive Compulsive Scale I: development, use and reliability. *Archives of General Psychiatry*, *46*, 1006–1011.
- Gower, P. and Kuyken, W.** (2011). *Therapist Competence, Case Conceptualisation and Therapy Outcome in Cognitive Behavioural Therapy*. Paper presented at the BABCP Annual conference, Guildford, July.
- Houghton, S., Saxon, D., Bradburn, G., Ricketts, T. and Hardy, G.** (2010). The effectiveness of routinely delivered cognitive behavioural therapy for obsessive compulsive disorder: a benchmarking study. *British Journal of Clinical Psychology*, *49*, 473–489.
- Jacobson, N. S. and Truax, P.** (1991). Clinical significance: a statistical approach to defining change in psychotherapy research. *Journal of Abnormal Psychology*, *59*, 12–19.

- Kuyken, W.** (2006). Evidence-based case formulation: is the emperor clothed? In N. Tarrrier (Ed.), *Case Formulation in Cognitive Behaviour Therapy: the treatment of challenging and complex cases*. New York: Routledge.
- Kuyken, W., Fothergill, C. D., Musa, M. and Chadwick, P.** (2003). The reliability and quality of cognitive case formulation. *Behaviour Research and Therapy*, *43*, 1187–1201.
- Kuyken, W., Padesky, C. and Dudley, R.** (2008). Process issues: the science and practice of case conceptualization. *Behavioural and Cognitive Psychotherapy*, *36*, 757–768.
- Kuyken, W., Padesky, C. and Dudley, R.** (2009). *Collaborative Case Conceptualization: working effectively with clients in cognitive-behavioural therapy*. New York: Guilford.
- McManus, F., Grey, N. and Shafran, R.** (2008). Cognitive therapy for anxiety disorders: current status and future directions. *Behavioural and Cognitive Psychotherapy*, *36*, 695–704.
- National Institute for Health and Clinical Excellence (NICE)** (2005). Obsessive compulsive disorder: core interventions in the treatment of obsessive compulsive disorder and body dysmorphic disorder. Retrieved from <http://www.nice.org.uk/Guidance/CG31>, November.
- Padesky, C. A., Kuyken, W. and Dudley, R.** (2011). *Collaborative Case Conceptualisation Rating Scale and Coding Manual*. Unpublished manual.
- Schulte, D., Kunzel, R., Pepping, G. and Schulte-Bahrenberg, T.** (1992). Tailor-made versus standardized therapy of phobic patients. *Advances in Behavior Research and Therapy*, *14*, 67–92.
- Shrout, P. E. and Fleiss, J. L.** (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, *86*, 420–428.
- Siegel, S. and Castellan, N. J.** (1988). *Nonparametric Statistics for the Behavioral Sciences (2nd ed.)*. New York: McGraw-Hill.
- Steketee, G. S., Frost, R. O. and Bogart, K.** (1996). The Yale-Brown Obsessive-Compulsive Scale: interview versus self-report. *Behaviour Research and Therapy*, *34*, 675–684.
- Stiles, W. B.** (2002). Assimilation of problematic experiences. In J. C. Norcross (Eds.), *Psychotherapy Relationships that Work: therapist contributions and responsiveness to patients* (pp. 357–365). New York: Oxford University Press.
- Stinson, C. H., Milbrath, C., Reidbord, S. P. and Bucci, W.** (1994). Thematic segmentation of psychotherapy transcripts for convergent analysis. *Psychotherapy*, *31*, 36–48.
- Wilson, K. G.** (1997). Science and treatment development: lessons from the history of behavior therapy. *Behavior Therapy*, *28*, 547–558.
- World Health Organization** (1992). *International Classification of Diseases (ICD-10)*. Geneva: World Health Organization.
- Zivor, M., Salkovskis, P. and Kushnir, J.** (2013). Formulation in cognitive behavior therapy for obsessive-compulsive disorder: aligning therapists, perceptions and practice. *Clinical Psychology: Science and Practice*, *20*, 143–151.
- Zivor, M., Salkovskis, P. and Oldfield, V.** (2011). Do we need to formulate in cognitive behavioural therapy for obsessive compulsive disorder? *Psicoterapia Cognitiva e Comportamentale*, *17*, 359–368.
- Zivor, M., Salkovskis, P. and Oldfield, V.** (2013). If formulation is the heart of cognitive behavioural therapy, does this heart rule the head of CBT therapists? *The Cognitive Behaviour Therapist*, *6*, 1–11.